Theoretical and Applied Genetics 47, 9-19 (1976)

The Use of Regression Techniques for Predicting the Response of Peas to Environment

B. Snoad and A.E. Arthur
John Innes Institute, Norwich (England)

Summary. Six pea varieties were grown in seven environments over two consecutive years and the results analysed using regression techniques. In both years a degree of linearity of response to environment was established for many of the characters recorded. Comparisons were made between the values observed in the second year and those predicted from the results of the first year, using 't' tests. As a result of these comparisons it appears that, due to poor linearity of response and/or different degrees of response in different years, accurate predictability of genotype behaviour was not possible.

Introduction

In an earlier series of experiments the presence of genotype-environment interaction was demonstrated in six pea cultivars grown in eight environments (Snoad and Arthur, 1974). With most of the characters scored, a linear relationship existed between genotype performance and measures of the environment, as originally demonstrated in cereals by Finlay and Wilkinson in 1963. There have been proposals that such regression techniques could be used for predictive purposes over years, sites and seasons (e.g. Breese, 1969). We have tried to evaluate the feasibility of using regressions as predictive tools for a range of characters in peas.

Materials and Methods

Six commonly grown pea varieties (DSP, Greenshaft, Jade, Puget, Scout and Sprite) were grown in 1973 in exactly the same way as in an earlier genotype-environment experiment (Snoad and Arthur, 1974). Thus four replicates of each of the six varieties were grown at four sites in Eastern England with an early and a late sowing at each site. Due to adverse soil conditions at the Peterborough site it was not possible to establish an early sowing there, so the total number of environments was seven in 1973.

Bedford (Site 1) Sowings 1 and 2
Peterborough (Site 2) Sowing 2
Stanfield Peat (Site 3) Sowings 1 and 2
Stanfield Sand (Site 4) Sowings 1 and 2

The sites used in 1973 were slightly different in location from those of the previous year, particularly the Stanfield Peat one which had a much lower water content in 1973. Fifty plants were removed for scoring from each replicate, and the same characters were recorded as in 1972 with the addition of mean dry weight of the seed harvested from each replicate.

Results

1. Analysis of variance

A random model situation has been assumed and the data were derived from the mean of each set of four replicates. The highest significant interaction mean square was always used as an error term in the F tests for the main effects. The significance of the interactions was decided by testing them against an error term derived from the total sums of squares of the means of the four replicates of each variety at each site and each sowing. In order to make a three way analysis, sites, sowings and varieties, site 2 had to be excluded since there were no data from a first sowing, and so the error term had only 108 degrees of freedom (Table 1).

It is immediately apparent that there are highly significant genotype-environment interactions for all the characters analysed. The main effects, apart from varieties, were generally non-significant when tested against the interactions; the exceptions were flowering time, where sites and sowings are important, and mean seed weight, where sites are influential. Broadly speaking the results from the two years are somewhat similar in that the most significant main effect was varietal, with sowing time being the least significant. Sites, however, which were marginally influential in 1972, were of little

Table 1. Degrees of freedom and mean squares derived in the analyses of variance of the data from three sites only (1973)

	d.f.	Days to first flower	AHUs to first flower	Node of first flower	Dist. bet 1st and 4th fl. nodes	No. pods at node 1	No. pods at node 2	No. pods at node 3	No. pods at node 4
Sites	2	97.32'	67083.49"	0.01	161.48	0.00	0.01	0.15	0.15
Sowings	1	10686.39'''	23793.06'	0.28	168.13	0.11	0.12	0.08	0.24
Varieties	5	101.61'''	11748.56'''	22.68'''	113.72'	0.85'''	0.43'	0.24'''	0.49'''
Si×So	2	4.24'''	429.95'''	0.01	76.77'''	0.11'''	0.25'''	0.44'''	0.34'''
Si×V	10	1.12'	257.51'''	0.04	3.46'''	0.02'''	0.02	0.01	0.01
So×V	5	1.06	371.21'''	0.16'''	21.24'''	0.02'	0.05'	0.04	0.01
Si×So×V	10	0.26"	33.47'''	0.04	1.81	0.01	0.01'	0.02'	0.01
Error	108	0.10	11.54	0.04	1.19	0.01	0.01	0.01	0.01

	d.f.	No. seeds/pod at node 1	No. seeds/pod at node 2	No. seeds/pod at node 3	No. seeds/pod at node 4
Sites	2	1.84	1.92	2.96	4.32
Sowings	1	0.03	0.01	0.69	1.01
Varieties	5	5.41'	4.60'	3.37'	3.87"
Si×So	2	1.52'''	1.82'''	2.16'	3.63'
Si ×V	10	0.08	0.06	0.11	0.46
So×V	5	0.82'''	0.56'	0.95	1.99
Si×So×V	10	0.13	0.15'	0.32'''	0.64'''
Error	108	0.08	0.07	0.12	0.19

	d.f.	No. seeds at node 1	No. seeds at node 2	No. seeds at node 3	No. seeds at node 4	Total No. seeds at nodes 1-4	Mean seed weight
Sites	2	2.85	3.35	4.65	6.10	62.55	3448.79'''
Sowings	1	5.02	6.24	0.18	0.27	80.16	129.58
Varieties	5	18.76'''	7.18	3.75	3.55	43.27	2333.49'''
Si ×So	2	4.25'''	11.54'''	11.23'''	6.88"	106.05'''	61.28
Si ×V	10	0.27	0.21	0.31	0.77	3.76	68.03
So×V	5	1.09'''	1.84'''	2.87'	3.72'	12.53'''	276.61
Si ×So×V	10	0.17	0.40	0.84'	0.86'	3.05	323.83'''
Error	108	0.33	0.27	0.38	0.40	2.84	53.87

' = 5.0 - 1.0%; " = 1.0 - 0.5%; ''' = < 0.5%

consequence in 1973. Significant genotype-environment interactions were demonstrated in both years.

2. Regression analyses

Again following the procedure outlined in our 1974 paper, the phenotypic behaviour of each variety in an environment was regressed upon the mean value obtained for all six varieties in that environment. These regression analyses can be related to the analysis of variance in order to determine how much of the genotype-environment interaction is due to the heterogeneity of the regressions and how much is unpredictable. In order to obtain as much information as possible, site 2 was included in these analyses. The genotype-environment interactions were determined from two-way analyses of variance, each sowing at a site being classed as an environment. This gave seven sites and six varieties and the error term was derived in the same way as for the three-way analysis but it had 126 degrees of freedom (Table 2).

In only 9 of the 18 characters analysed can the G×E interaction be partly accounted for by the heterogeneity of the regressions. In contrast, the residual or non-linear component was significant for all characters but one. When the data from the two years are compared a significant difference becomes appar-

Table 2. Degrees of freedom and mean squares for partitioning of genotype-environment interactions into heterogeneity of regression and residual components (1973)

	d.f.	Days to first flower	AHUs to first flower	Node of first flower	Dist. bet 1st and 4th fl. node	No. pods at node 1	No. pods at node 2	No. pods at node 3	No. pods at node 4
Heterogeneity of Reg.	5	2.80'''	1084.10'''	0.07	15.61'''	0.07'''	0.06'''	0.01	0.01
Residual	25	0.67'''	7.81	0.07'''	4.50'''	0.02'''	0.03'''	0.03'''	0.02'''
Error	126	0.11	12.98	0.03	1.24	0.01	0.01	0.01	0.01

	d.f.	No. seeds/pod at node 1	No. seeds/pod at node 2	No. seeds/pod at node 3	No. seeds/pod at node 4
Heterogeneity of Reg.	5	0.15	0.03	0.33'	0.37
Residual	25	0.30'''	0.22'''	0.32'''	0.91'''
Error	126	0.09	0.08	0.11	0.20

	d.f.	No. seeds at node 1	No. seeds at node 2	No. seeds at node 3	No. seeds at node 4	Total No. seeds at nodes 1-4	Mean seed weight
Heterogeneity of Reg.	5	1.22''	1.02'''	0.53	0.34	1.68	404.58'''
Residual	25	0.52'	0.70'''	0.99'''	1.60'''	6.85'''	315.19'''
Error	126	0.36	0.29	0.36	0.39	2.93	49.16

ent because in 1972 both the heterogeneity and the residual components were significant for all the characters. This suggests that although genotype-environment interactions were demonstrable in both years they were significantly less linear in 1973 than in 1972.

The differences between the two years are in fact quite striking when the interactions are displayed in the regressions. From one year to another there may be differences in ranking and/or differences in slope in many of the individual characters being analysed. These differences immediately indicate that it would be unwise to consider the regression lines from the 1972 data as necessarily being linearly related to similar lines from the 1973 data. It has been said (e.g. Breese, 1969) that the relative performances of populations and their hybrids can be predicted over seasons, years and locations using these regression techniques. The differences between two years with these pea data, however, throw doubt upon the validity of such an approach and it was decided to test the accuracy of prediction using a small number of important characters selected from those available.

The characters chosen were:

1. AHUs to day of first flower
2. Distance from 1st to 4th flowering node (internode length)
3. Number of pods at the second flowering node
4. Number of seeds per pod at the second flowering node
5. Total seeds at the second flowering node
6. Total seeds at the first four flowering nodes (yield).

The data involving these six characters in 1972 and 1973 were combined and the results of the three-way analyses of variance are given in Table 3. These analyses excluded the Peterborough sites in both years because of the missing first sowing in 1973, so there are only six sites and the error term had 216 degrees of freedom. It can be seen that only for internode length was the year component significant when tested

Table 3. Degrees of freedom and mean squares derived in the analyses of variance of the data from three sites only (1972 plus 1973)

	d.f.	AHUs to flower	Dist. bet 1st and 4th fl. nodes	No. Pods at node 2	No. Seeds/pod at node 2	Total seeds at node 2	Total yield at nodes 1-4
Year	1	5390.68	1579.59'	0.31	16.93	34.19	72.87
Env	5	46173.97'''	113.57	0.28	2.86	10.28	75.29
Variety	5	21548.82'''	227.43'''	0.48	9.77'''	15.77''	85.85'
Year × Env	5	2566.43'''	185.38'''	0.34'''	7.34'''	21.69'''	239.40'''
Year × Var	5	162.24	14.84	0.12'''	0.20	1.42'''	11.26
Env × Var	25	141.30	4.23	0.05'	0.28	1.26'''	8.28
Year × Env × Var	25	96.83'''	7.47'''	0.02'''	0.22'''	0.48'''	6.23'
Error	216	11.43	1.33	0.01	0.09	0.35	3.61

against the appropriate interaction. However, years are involved in significant interactions for all six characters and years are also significant as main effects or as interactions and the genotype-environment interactions are significant for all characters.

The regressions for the six characters for 1972 and 1973 were drawn separately and are shown side by side in Figs. 1 and 2. It is apparent that the data from 1973, despite the increased non-linear components demonstrated from the analyses of variance in relation to 1972, can still be portrayed in linear fashion. With the exception of the AHU character there is a noticeable increase in the standard deviations of the regression lines in 1973 as opposed to 1972. The regressions are based on data from eight environments in 1972 and seven environments in 1973.

AHUs to first flower

The ranking and interrelationships of the six varieties are broadly similar in the two years. However, the slopes of some of the regression lines are obviously different so that in 1973 there is an impression of the lines fanning out from the origin. The standard deviations tended to be lower in 1973 than in 1972 (Fig.1).

Internode length

The distances between the first and the fourth flowering nodes varied much more in 1972 than in 1973 for all six varieties. The relationships between the lines are different in the two years and the slopes of some of the lines also differ. In 1973 the data for Sprite were not linear and no regression line could be drawn. Standard deviations of the regressions were generally higher in 1973 (Fig.1).

Number of pods at the second flowering node

The ranking and the slopes of those lines that can be drawn do not differ very much between the two years. The major differences are that in 1972 Sprite was non-linear whereas in 1973 Jade and Scout were non-linear. Again there is a tendency for the standard deviations to be greater in 1973 (Fig.1).

Number of seeds per pod at the second flowering node

The general ranking of the lines remains similar over the two years but large differences in slope are noticeable with DSP and Greenshaft. The range of values observed in 1972 is extended considerably above and below that recorded in 1973. The data for Sprite in 1972 were not linear and were only marginally so in 1973 when the Puget data could not be used for regressions. Standard deviations were larger in 1973 (Fig.2).

Total seeds at the second flowering node

As might be expected from the results of observations on the previous character, the range of values observed in 1972 was greater than that in 1973. The slopes for DSP and Puget in particular were different in the two years and the standard deviations of all regression lines increased in 1973. Sprite was either border-line or non-significant as regards linearity in both years and Scout, which gave a good line in 1972, was non-linear the next year. One of the most striking differences concerns DSP and Puget which, because of the change in slope between the two years, rank very differently in one year as opposed to the other (Fig.2).

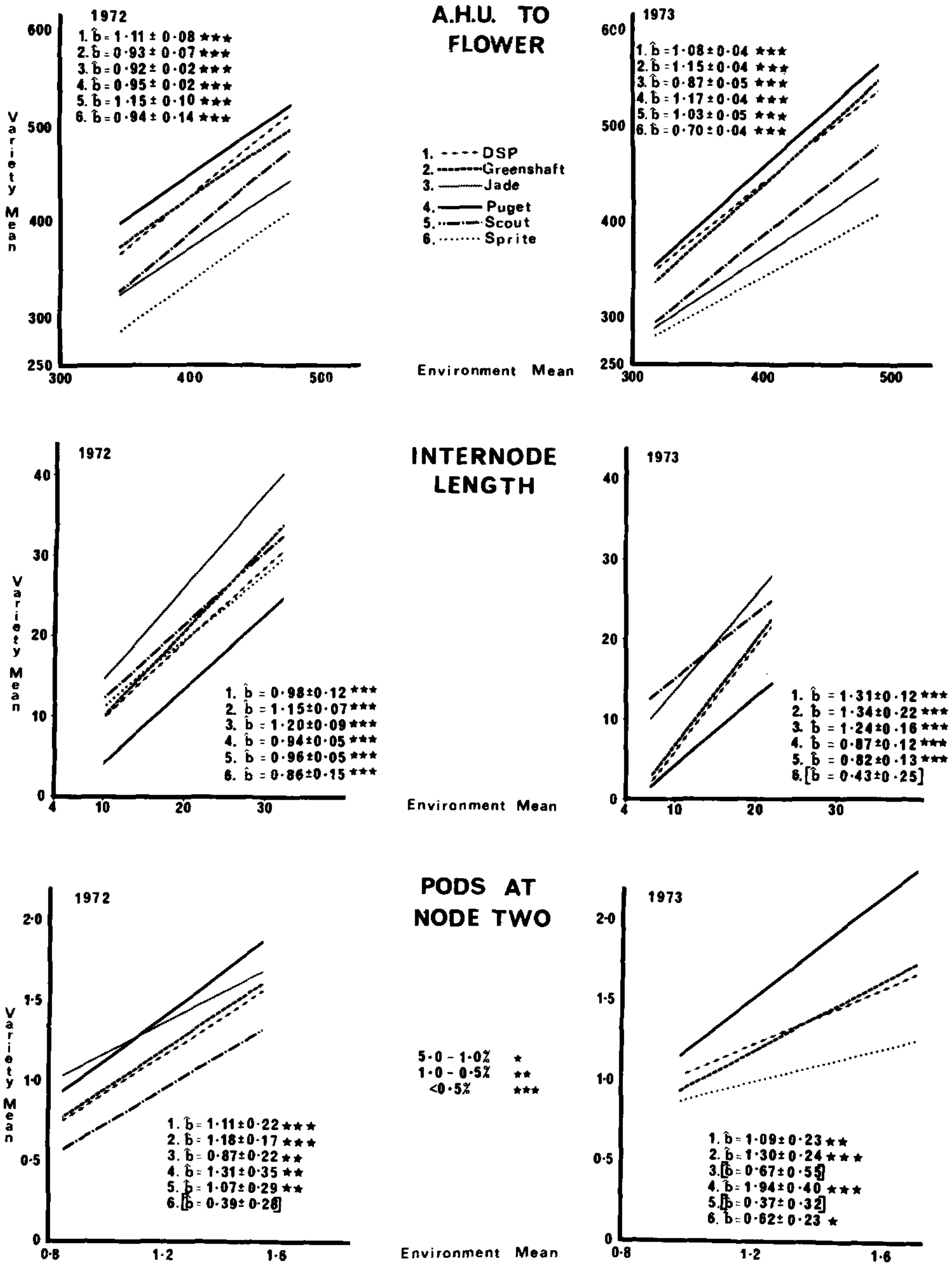

Fig.1. Regressions for flowering time (using accumulated heat units), internode length and the number of pods at the second flowering node in 1972 and 1973

Total seeds from the first four flowering nodes (Yield)

Differences in slope are particularly apparent for DSP and Sprite whilst differences in ranking can also be seen. The standard deviations were higher in 1973 and the Puget data were not linear in that year (Fig. 2).

Prediction

One of the most practical ways of testing the efficiency of using these regression data for prediction is to use the 1972 slopes to estimate the 1973 values by means of the formula,

$$y_j = \bar{y} + b(x_j - \bar{x})$$

where y_j = the predicted value of a genotype in the jth environment (1973)

$\bar{y}$ = the mean of y (1972)

b = the regression coefficient (1972)

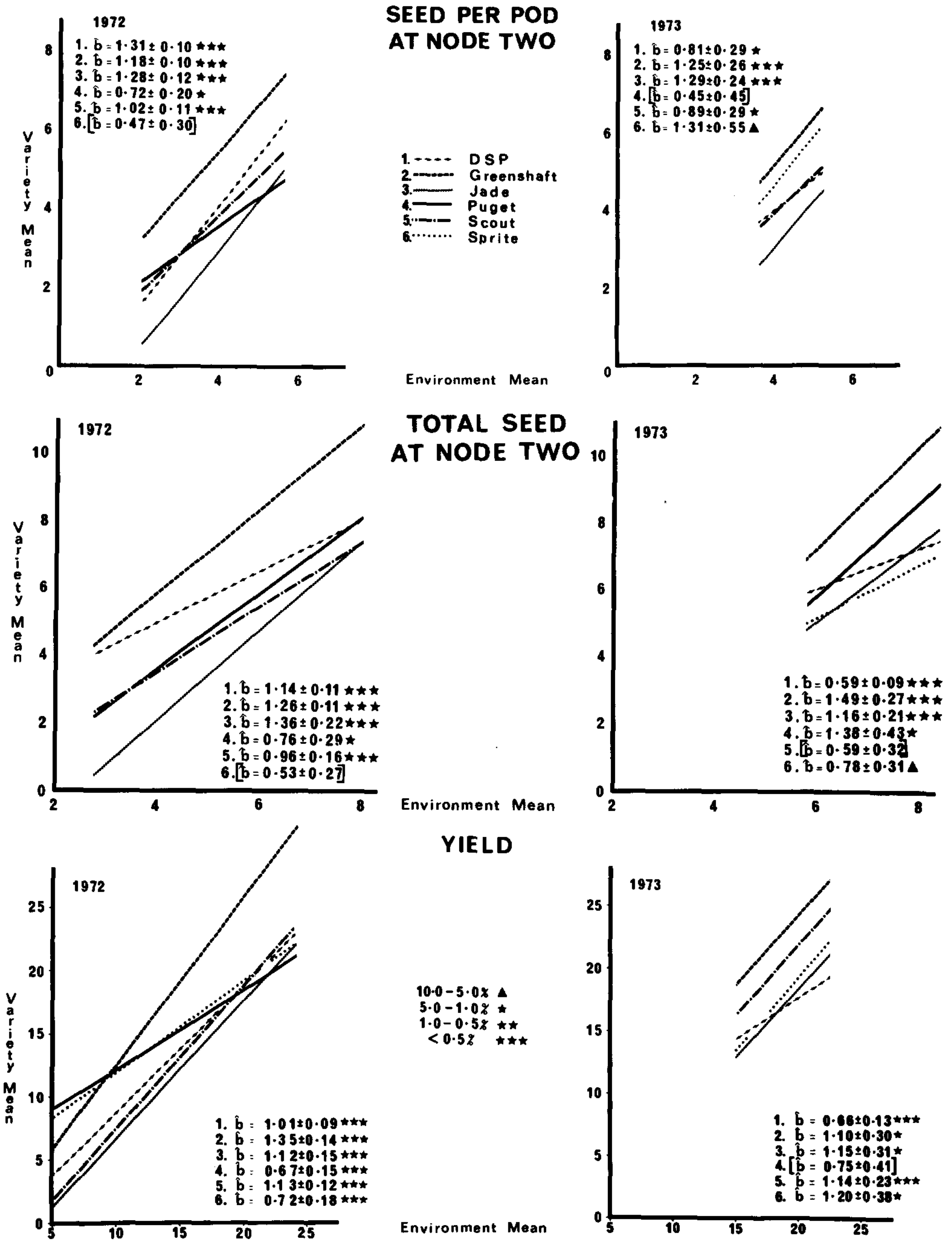

Fig.2. Regressions for the number of seeds per pod at the second flowering node, the total seed number at the second flowering node and the total seed number at the first four flowering nodes in 1972 and 1973

x_j = the environmental value in the jth environment (1973)

$\overline{x}$ = the mean of x (1973).

The observed and expected values may then be compared by means of 't' tests where the difference between the two values is divided by its standard error (Samuel, Hill, Breese and Davies, 1970). These calculations have been made for all six characters and the results are set out in Tables 4 and 5.

In none of the six characters is there a perfect fit for all the 42 pairs of observed and expected values. The number of pods at node two and the total seed yield at the first four nodes, however, have significant differences between only two and one, respectively, of the 42 values, and the seeds per pod at node two and the seed yield at node two have one and three discrepancies, respectively. AHUs to flower and internode length are the worst characters for fit in being significantly different ten and five times out of each 42.

It is obvious that poor regression data will not be accurate for predictive use and this applies particularly to the number of pods at node 2 in the present experiments and also, to a lesser extent, to those other characters involving components of yield. With data such as these the variances of the observed values will be large and so 't' tests will not demonstrate statistically significant differences even when there is considerable discrepancy between the observed and expected values. The results of the 't' tests in Tables 4 and 5 therefore have to be considered in relation to the quality of the data.

The two characters AHUs to flower and internode length are quite different, however, in providing good linear data in both years of these experiments and it is with these characters that the most stringent tests are made. The results show quite clearly that with these pea data it is not possible to predict with any great degree of accuracy from one year to the other because of differences in slope and/or ranking from year to year. A similar conclusion is reached if attempts are made to predict sowing 2 from sowing 1 data within or between years (Snoad unpubl.).

Two alternative methods of measuring the environmental values have been considered:

1. Regressing the mean value of replicates 1 and 2 on the mean values of replicate 3 and 4 of all six genotypes as measures of the environment thus ensuring that the environmental measures are related to, but not mathematically part of, the genotypic measures.
2. Regressing the mean value of the four replicates for each genotype on the mean of the remaining five genotypes, thus removing part of the correlation which is inbuilt when using the conventional Finlay and Wilkinson (1963), method.

Both these techniques raise problems for the measurement of genotype-environment interactions from the analyses of variance, so they have only been used for obtaining regressions in order to compare the slopes of the lines which can be drawn. The slopes of the regression lines obtained in these three ways for two characters in 1972 are given in Table 6 from which it can be seen that the differences generally speaking are not significant. This is in agreement with the findings of Perkins and Jinks (1973) who showed that heterogeneity of regression and remainder items and the ranking of inbred lines on the basis of their regression coefficients were not significantly changed when independent, as opposed to dependent, measures of the environment had been used. Williams (1975) also decided that the conclusions drawn from regression data were unaffected by the two methods he used to measure the environment.

Discussion and Conclusions

From a practical point of view it would obviously be valuable if plant response to a range of environments could be predicted from existing genotype-environment regression data. All attempts to do this with pea data from two consecutive years have, however, proved to be only marginally successful. Failure has resulted from two main causes: either the data have not been sufficiently linear to permit the necessary degree of accuracy or, when good linear data have been obtained, the degrees of response have differed from year to year.

Already there are a number of indications in the literature of curvilinearity of response to environment, of non-linearity and of different responses to environment in different years or seasons. As the range of plant material examined using these techniques is extended so the possibility of the Finlay and Wilkinson technique being universally applicable becomes less likely.

Table 4. Values for three characters observed in 1973 compared by 't' tests with the values expected from the 1972 data

	Envt.	1	2	3	4	5	6	7	8
AHUs to first flower									
Variety									
DSP	Obs.	487	540	-	530	346	420	345	416
	Exp.	493	534	-	518	339	403	335	401
Greenshaft	Obs.	496	543	-	537"	346	399	344	398
	Exp.	480	514	-	501	351	404	348	403
Jade	Obs.	425	448	-	419'''	293	341"	286'	338"
	Exp.	427	461	-	448	299	352	296	351
Puget	Obs.	514	566'''	-	540'	352"	434	348"	434
	Exp.	505	540	-	526	373	427	370	426
Scout	Obs.	435	478	-	474	304'''	346	300	346
	Exp.	454	497	-	480	294	360	291	359
Sprite	Obs.	392	401	-	389	279	324	279	324
	Exp.	392	428	-	414	262	316	260	315
Distance in cm. from 1st to 4th flowering node (Internode Length)									
Variety									
DSP	Obs.	17.80	14.49	-	19.24	16.79	21.78	3.12	13.30
	Exp.	17.84	16.18	-	19.59	15.79	20.82	7.07	15.81
Greenshaft	Obs.	14.33'	17.38	-	20.80	12.34	23.18	3.71	18.44
	Exp.	17.73	15.79	-	19.79	15.33	21.23	5.09	15.35
Jade	Obs.	20.39'	23.60	-	23.14	20.82	28.15	10.98	21.91
	Exp.	25.08	23.05	-	27.23	22.57	28.73	11.89	22.60
Puget	Obs.	10.40	8.40	-	14.81	8.76	14.51	2.77	11.09
	Exp.	12.37	10.78	-	14.05	10.41	15.23	2.04	10.43
Scout	Obs.	23.68'	20.26	-	22.32	21.34'	23.40	12.12	21.41'
	Exp.	20.25	18.63	-	21.97	18.24	23.17	9.70	18.26
Sprite	Obs.	22.25	18.18	-	17.87	19.82	19.66	13.77	13.87
	Exp.	19.01	17.55	-	20.54	17.21	21.62	9.55	17.22
Number of pods at the second flowering node									
Variety									
DSP	Obs.	1.70	1.16	-	1.08	1.48	1.36	1.53	1.39
	Exp.	1.73	1.25	-	1.14	1.51	1.46	1.38	1.52
Greenshaft	Obs.	1.73	1.16	-	0.96	1.42	1.34	1.53	1.46
	Exp.	1.79	1.27	-	1.16	1.56	1.50	1.42	1.57
Jade	Obs.	1.69	1.33	-	1.52	1.58	1.64	1.03'	1.74
	Exp.	1.82	1.43	-	1.35	1.64	1.60	1.54	1.65
Puget	Obs.	2.34	1.72	-	1.08	1.97	1.85	1.92	1.84
	Exp.	2.11	1.53	-	1.40	1.84	1.78	1.69	1.86
Scout	Obs.	1.45	1.31	-	1.42	1.57	1.60	1.36	1.71'
	Exp.	1.52	1.05	-	0.94	1.31	1.25	1.18	1.32
Sprite	Obs.	1.39	1.03	-	0.99	1.11	1.02	1.00	1.02
	Exp.	1.31	1.14	-	1.10	1.23	1.21	1.18	1.26

Table 5. Values for three characters observed in 1973 compared by 't' tests with the values expected from the 1972 data

	Envt.	1	2	3	4	5	6	7	8
Number of seeds per pod at the second flowering node									
Variety									
DSP	Obs.	3.99	4.63	-	4.36	5.13	4.50	3.64	4.45
	Exp.	4.83	4.29	-	4.71	5.68	4.98	3.76	4.90
Greenshaft	Obs.	6.05'	5.09	-	5.67	6.99	5.87	5.20	5.66
	Exp.	6.18	5.70	-	6.07	6.95	6.31	5.21	6.24
Jade	Obs.	3.98	3.13	-	3.56	4.89	3.59	3.02	4.05
	Exp.	3.70	3.18	-	3.59	4.54	3.84	2.65	3.77
Puget	Obs.	3.87	2.69	-	2.92	4.04	3.68	3.69	3.23
	Exp.	3.94	3.64	-	3.88	4.42	4.02	3.34	4.00
Scout	Obs.	4.18	4.10	-	4.98	4.86	4.48	3.50	4.58
	Exp.	4.41	3.99	-	4.32	5.07	4.52	3.58	4.46
Sprite	Obs.	5.40	5.37	-	5.42	5.48	6.03	3.53	5.83
	Exp.	4.51	4.32	-	4.47	4.82	4.56	4.13	4.53
Total number of seeds at the second flowering node									
Variety									
DSP	Obs.	7.19	5.98	-	6.19	7.81	6.47	6.08	6.65
	Exp.	8.49	5.63	-	6.05	8.45	6.94	5.73	7.25
Greenshaft	Obs.	11.22	6.90	-	6.85'	10.71	8.23'	8.30	8.66
	Exp.	11.36	8.19	-	8.66	11.31	9.64	8.31	9.98
Jade	Obs.	7.17	4.73	-	5.90	8.04	6.08	4.38	7.04
	Exp.	7.94	4.55	-	5.05	7.88	6.10	4.67	6.47
Puget	Obs.	9.84	5.36	-	4.75	8.77	7.36	7.49	6.84
	Exp.	8.23	6.37	-	6.65	8.25	7.25	6.44	7.45
Scout	Obs.	7.24	6.47	-	7.77'	8.38	7.58	5.62	8.13
	Exp.	7.66	5.26	-	5.61	7.63	6.36	5.34	6.62
Sprite	Obs.	7.53	5.74	-	5.90	6.29	6.33	3.84	6.35
	Exp.	6.46	5.13	-	5.32	6.43	5.73	5.17	5.88
Total seeds from flowering nodes 1-4 (Yield)									
Variety									
DSP	Obs.	18.19	14.51	-	17.46	20.86	17.80	14.48	17.96
	Exp.	20.79	14.72	-	17.42	21.55	21.06	13.79	19.94
Greenshaft	Obs.	28.76	18.09	-	19.67'	27.06	25.74	21.50	25.08
	Exp.	28.45	20.49	-	24.03	29.45	28.80	19.28	27.34
Jade	Obs.	17.17	13.65	-	19.86	21.37	20.76	11.03	21.47
	Exp.	20.47	13.52	-	16.61	21.34	20.78	12.46	19.50
Puget	Obs.	22.70	14.53	-	12.89	22.01	19.76	17.96	16.68
	Exp.	19.75	15.63	-	17.46	20.27	19.93	15.00	19.18
Scout	Obs.	21.34	16.57	-	22.39	24.35	24.92	15.48	24.59
	Exp.	20.85	14.72	-	17.44	21.61	21.12	13.78	19.99
Sprite	Obs.	22.40	17.83	-	18.69	19.37	23.14	9.30	19.85
	Exp.	20.23	16.17	-	17.98	20.74	20.41	15.55	19.67

Table 6. Comparison of regression slopes obtained with two characters having estimated the environment in three different ways

	Environment measured as		
AHUs	Mean of all genotypes	Mean of two replicates	Mean of remaining five genotypes
Variety 1	1.11 ± 0.08	1.11 ± 0.09	1.12 ± 0.10
Variety 2	0.93 ± 0.07	0.92 ± 0.07	0.92 ± 0.08
Variety 3	0.92 ± 0.02	0.91 ± 0.02	0.90 ± 0.02
Variety 4	0.95 ± 0.03	0.91 ± 0.07	0.94 ± 0.03
Variety 5	1.15 ± 0.10	1.13 ± 0.11	1.16 ± 0.13
Variety 6	0.94 ± 0.14	0.96 ± 0.14	0.90 ± 0.16
Internode length			
Variety 1	0.90 ± 0.12	0.99 ± 0.14	0.95 ± 0.15
Variety 2	1.15 ± 0.07	1.22 ± 0.08	1.17 ± 0.08
Variety 3	1.20 ± 0.09	1.34 ± 0.19	1.16 ± 0.12
Variety 4	0.94 ± 0.05	0.97 ± 0.06	0.93 ± 0.06
Variety 5	0.96 ± 0.05	0.85 ± 0.08	0.85 ± 0.08
Variety 6	0.86 ± 0.15	0.85 ± 0.17	0.80 ± 0.17

Five populations of cocksfoot were grown in two different locations over two years and regression techniques were applied by Breese (1969). Only in one of the years, 1965, is it possible to demonstrate good linearity of response to environment. In 1966 only two of the populations approach linearity but this might be due to data being available from only four environments in that year. Very good linear responses were obtained, however, when the data were combined from the two years but the degrees of response of the five populations, as measured by regression slopes, differ if the data are combined over years and sites rather than over years within sites.

Final plant height in *Nicotiana rustica* plants grown in all possible combinations of presence or absence of N,P and K fertilizers has been analysed using regression techniques (Hill and Perkins, 1969). The relationship between phosphatic and non-phosphatic groups was such that at the lower end of the scale a radical alteration of genotype response to environment had to be considered and three hypotheses were put forward. The relationship between environment and performance might break down altogether or a second, and different, relationship could then apply. Thirdly, the relationship over the complete environmental range might be curvilinear. It is pointed out that within the range normally experienced this curvilinear relationship would be, to all intents and purposes, linear.

A number of criticisms of genotype-environment regression techniques have been made by Knight (1970) using Breese's (1969) data. Knight pointed out the similarity between three of the five populations of cocksfoot used and indicated how big a contribution they were making to the environmental measures. One population, however, was strikingly different in origin, Cantal, with a different growth response and a high deviation from regression when sites and cutting-frequencies were distinguished. He concluded that extrapolation from these data would lead to serious errors in interpretation.

Regressions for two years have been compared for *Lolium perenne* grown as spaced plants and swards (Samuel, Hill, Breese and Davies, 1970). Significant differences in response were noted between the two years in both conditions of growth but, as one of the years was a drought year giving low yields, it was suggested that under these stress conditions the regressions become curvilinear. Comparisons of observed and predicted values of one population were not significantly different at three out of four nitrogen levels. In the fourth, high nitrogen, level the difference was significant, however, and this was attributed to nitrogen toxicity. One of the conclusions reached was that it would be unwise to generalise too far until these techniques have been extended to cover a wider range of genetic material.

Other genotype-environment interaction experiments have been reported in *Lolium perenne* grown for two years in contrasting nitrogen regimes (Hill and Samuel, 1971). When the environment was subdivided into its component parts, small and significant changes in slope of regression lines occurred from year to year and for the two nitrogen levels. Curvilinearity of the response lines was one of the reasons put forward to explain these differences in slope. Alternatively, it was suggested that light and temperature, which affect response within years, may have been operating at entirely different levels from those affecting response across nitrogen levels or from one year to another.

Methods for the statistical analysis of genotype-environment interactions have been reviewed extensively by Freeman (1973). In this review it is emphasized that extrapolation of genotype response to environment should be approached with caution, particularly when the new environments come from a population different from that from which the experimental environments are a sample and also if some environments are sub-optimal or super-optimal. The difficulties inherent in deciding what conditions will lead to linearity of regression are also pointed out together with an observation that, while one set of characters has frequently been found to give linear regressions, other characters measured on the same set of genotypes have not.

Wide deviations from linearity of response have been reported for *Brassica napus* from which it was concluded that regression techniques to characterise response of genotypes to environment are "an over-simplification" (Witcombe and Whittington, 1971).

The yield of strawberries has also been examined using regression techniques and it was concluded that, since different regression coefficients were obtained with the same material grown in a range of environments and years, predicting responses of cultivars to untried environments using these techniques is "a hazardous procedure" (Williams, 1975).

It seems inescapable that only occasionally can a truly linear response to a range of environments be adequately demonstrated, and this conclusion applies to a steadily widening range of characters and genera. Assuming that a curvilinear response sometimes exsists, there remains the problem of deciding at what point in the environmental range the change from linearity to curvilinearity takes place. An even more worrying problem is that raised by the demonstration of different degrees of linear response in different circumstances and it is this problem in particular which throws the most doubt upon accurate predictability of environmental responses using regression techniques.

Acknowledgements

We are again very much indebted to the staff of the Unilever Research Laboratories and the Processors and Growers Research Organisation for providing the Bedford and Peterborough sites, respectively, and for all the assistance given during the course of the pea season. We would also like to thank Ann Payne, Denise Caston and Sharon Negus for all the assistance which they gave in the recording and analysis of the data.

Literature

Breese, E.L.: The measurement and significance of genotype-environment interactions in grasses. Heredity 24, 27-44 (1969)

Finlay, K.W.; Wilkinson, G.N.: The analysis of adaptation in a plant breeding programme. Aust. J. Agric. Res. 14, 742-754 (1963)

Freeman, G.H.: Statistical methods for the analysis of genotype-environment interactions. Heredity 31, 339-354 (1973)

Hill, J.; Perkins, J.M.: The environmental induction of heritable changes in *Nicotiana rustica*. Effects of genotype-environment interactions. Genetics, 61, 661-675 (1969)

Hill, J.; Samuel, C.J.A.: Measurement and inheritance of environmental response amongst selected material of *Lolium perenne*. Heredity 27, 265-276 (1971)

Knight, R.: The measurement and interpretation of genotype-environment interactions. Euphytica 19, 225-235 (1970)

Perkins, J.M.; Jinks, J.L.: The assessment and specificity of environmental and genotype-environmental components of variability. Heredity 30, 111-126 (1973)

Samuel, C.J.A.; Hill, J.; Breese, E.L.; Davies, A.: Assessing and predicting environmental response in *Lolium perenne*. J. agrin. Sci., Cambridge 75, 1-9 (1970)

Snoad, B.; Arthur, A.E.: Genotype-environment interactions in peas. Theor. Appl. Genetics 44, 222-231 (1974)

Williams, H.: Genotype-environment interactions in strawberry cultivars. Hort. Res. 14, 81-88 (1975)

Witcombe, J.R.; Whittington, W.J.: A study of the genotype by environment interaction shown by germinating seeds of *Brassica napus*. Heredity 26, 397-411 (1971)

Received April 10, 1975
Communicated by R. Riley

B. Snoad
A.E. Arthur
John Innes Institute
Colney Lane
Norwich NR4 7UH (England)