THE STABILITY OF VARIETAL PERFORMANCE OVER YEARS

2. ANALYSING VARIETY TRIALS

D. G. PEDERSON*

*Faculty of Agriculture, University of Sydney, Sydney,
New South Wales, Australia*

Received 15.xi.73

* Present Address: Department of Genetics, University of Birmingham, P.O. Box 363, Birmingham B15 2TT.

SUMMARY

Associated with each environment in a variety trial is an environmental effect, locatable on some linear scale. The distribution $f(e)$ denotes the long-term distribution of effects, and the function $g(e)$ describes the expected response of a variety to a change in the environment. The value of practical interest for any one variety is

$$\mu_g = \int g(e)f(e)de,$$

the expected value of $g(e)$ over all environments. Methods of estimating μ_g are discussed, both when $f(e)$ is known and when only a limited set of effects is available. It is shown that in the latter case it is necessary to make restrictive assumptions concerning $g(e)$. An example involving wheat grain yield illustrates a situation in which the proposed analysis would be of advantage.

1. INTRODUCTION

THE general aim of plant breeding is to produce varieties with superior performance in all environments likely to be encountered in the future, and in most breeding programmes variety trials are the means by which future performance is predicted. Suppose that a set of a varieties are grown at b sites for c years, with r replicates per site per year in a completely random design. A model for the l th observation of the i th variety at the j th site in the k th year is

$$X_{ijkl} = \mu + v_i + s_j + y_k + (vs)_{ij} + (vy)_{ik} + (sy)_{jk} + (vsy)_{ijk} + \epsilon_{ijkl}.$$

If the site and year effects are considered to be fixed (Eisenhart, 1947) then the mean of the i th variety, in terms of the components of the model, is

$$\bar{x}_{i...} = \mu + v_i + \epsilon_{i...}$$

For two varieties included in the trial the variance of the difference in means is $2\sigma^2/rbc$, where $\sigma^2 = \text{var}(\epsilon_{ijkl})$. The accuracy of a comparison between varieties may therefore be increased by increasing any or all of r , b and c . However, it is more appropriate to assume that site and year effects are random (Eisenhart, 1947) since the interest lies in the performance of varieties in all environments likely to be encountered in the future. The mean of the i th variety then becomes

$$\bar{x}_{i...} = \mu + v_i + \bar{s}_j + \bar{y}_k + (\overline{vs})_{i.} + (\overline{vy})_{i.} + (\overline{sy})_{.j} + (\overline{vsy})_{..} + \bar{\epsilon}_{i...},$$

and the variance of the difference in means for two varieties is

$$2 \left(\frac{\sigma_{vs}^2}{b} + \frac{\sigma_{vy}^2}{c} + \frac{\sigma_{vsy}^2}{bc} + \frac{\sigma^2}{rbc} \right),$$

where, in general, σ_m^2 is the variance of component m . It is relatively easy to increase either b or r and so increase the accuracy of varietal comparisons, but an adequate sampling of years is usually not practicable. The lower limit of the variance is effectively $2\sigma_{vy}^2/c$ and this may be unacceptably high if variety $\times$ year interaction is an important source of variation.

It is also common to calculate for each variety some second-degree statistic which estimates the responsiveness to a change in environmental conditions. If the statistic is the variance over environments or the between-environment component of variance then, as is the case for the arithmetic mean, the accuracy of varietal comparisons will depend most critically on the number of years sampled. An alternative technique is to regress the values for each variety against a measure of effect for each environment (Yates and Cochran, 1938). The slope of the regression line and the deviations from the line are taken as measures of stability for each variety (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966). If the response is truly linear over the range of possible environments then the sample regression coefficient is an unbiased estimate of the population value regardless of the environments included in the sample, as has been noted by Breese (1969). The variance of varietal differences will depend more on the range of environmental effects than on the actual number of years sampled, for a given number of trials. But if, for example, it were concluded that a variety responds relatively well to improved environmental conditions then it would be necessary to know something of the expected frequency of good and poor environments to make full predictive use of this information (Thomson and Wright, 1971).

Further, when both the mean and a stability parameter have been estimated it may be difficult to objectively combine these two pieces of information and thus determine the overall worth of a variety. For example, would a relatively stable variety be preferred to a variety which is relatively unstable but has a higher average value?

The present study of variety trials and their analysis therefore has a two-fold aim:

- (1) to determine a single parameter which has elements of both the arithmetic mean and a stability parameter;
- (2) in determining this single parameter, to take account of the relative frequencies with which different environments, in particular good and poor years, are encountered by a released variety.

Thus a completely random model is assumed, and the attempt is made to increase the accuracy of varietal comparisons by the use of supplementary information.

2. GENERAL THEORY

The term "environment" will be used in a general sense, and may refer to the conditions encountered at a single location or the average conditions over a set of locations in a particular year. The gross effect of an environment will be denoted by e .

Different environments have different probabilities of reoccurrence, and in line with this we shall consider the function $f(e)$, the probability density function of e ,
i.e.

$$f(e) \geq 0$$

and

$$\int_{-\infty}^{\infty} f(e)de = 1.$$

Without loss of generality we can assume that the mean is zero,
i.e.

$$E[e] = 0.$$

Negative values of e are therefore associated with poor environments and positive values are associated with environments which are better than average.

The performance of a variety will supposedly depend on e and we can represent the relationship by the function $g(e)$. Then the average performance over all environments is given by the expected value of $g(e)$, viz.

$$\mu_g = \int_{-\infty}^{\infty} g(e)f(e)de. \quad (1)$$

It is suggested that this is the parameter most often of interest in a practical situation. To estimate μ_g it is necessary to obtain estimates of $f(e)$ and $g(e)$, and these will now be considered in turn.

3. THE DISTRIBUTION OF ENVIRONMENTAL EFFECTS

Long-term data are essential if a distribution of effects is to be built up. These will often be available in the case of a commercially grown crop plant, as has been demonstrated for grain yield in wheat by Pederson (1974). Ten

TABLE 1
Analysis of variance for data from n sites in each of m years, including a regression analysis of yearly means against estimated effects for the region

Source	d.f.	M.S.	Expectation of mean square
Years	$m-1$	M_Y	$\sigma_{YS}^2 + n\sigma_Y^2$
Regression	1	M_R	$\sigma_{YS}^2 + n(1-\rho^2)\sigma_Y^2 + n(m-1)\rho^2\sigma_Y^2$
Deviation	$m-2$	M_D	$\sigma_{YS}^2 + n(1-\rho^2)\sigma_Y^2$
Sites	$n-1$	M_S	$\sigma_{YS}^2 + m\sigma_S^2$
Years $\times$ Sites	$(m-1)(n-1)$	M_{YS}	σ_{YS}^2

regions of New South Wales were considered, each of some 3000 square kilometres, and for each region a polynomial curve was fitted to commercial wheat yields from 1922 to 1970. The deviation from the fitted yield for each year was divided by the fitted yield as a standardising procedure, and the standardised deviations were taken as estimates of e . For each region a lognormal distribution gave a good fit to the observed distribution of effects.

Since each region covered a wide area, the effects for any one region would hardly be relevant to a series of variety trials at only one site within

the region. However, there was some evidence that the mean yield of several trials correlated well with the estimated regional effect for a particular year. Looking at this point in more detail, suppose that one or more varieties have been grown at n sites within a region for m years. A two-factor analysis of variance may be carried out on the mn average yields, and a regression analysis of the mean for each year against the estimated yearly effect may also be performed. The form of the composite analysis and the expectations of the mean squares, where ρ is the correlation between the expected trial mean and the regional effect, are given in table 1.

Under the null hypothesis that $\rho = 1$ we have

$$M_D/M_{YS} \sim F[m-2, (m-1)(n-1)].$$

Tests of the null hypothesis for five sets of data are given in table 2, for four regions which have previously been considered by Pederson (1974). The

TABLE 2
*Correlation analyses of yearly means against estimated effects for five sets of data;
test of the null hypothesis that $\rho = 1$*

Region	Number of years	Number of sites	Number of varieties	Variance ratio
Coolamon	5	2	4	6.7*
Culcairn	6	2	4	0.3 NS
Peel	4	2	4	0.3 NS
Narrandera	8	2	3	1.9 NS
Narrandera	7	3	3	0.9 NS

NS = Not significant at 5 per cent level; * = Significant at 5 per cent level.

null hypothesis is rejected in only one case, in apparent disagreement with Davidson (1962), who found that an asymptotic regression function described the relationship between farm yield and experimental yield for a set of wheat trials in the State of Victoria.

In the analysis of variance (table 1) we have

$$r^2 = \frac{M_R}{(m-1)M_Y},$$

where r is the observed correlation between trial mean and regional effect. The estimated components of variance from an analysis may therefore be used to investigate the relationship between r and n . If there are n^* sites per year rather than n , the square of the correlation is estimated by

$$\hat{r}^2 = \frac{M_{YS} + (n^*/n)(M_R - M_{YS})}{(m-1)[M_{YS} + (n^*/n)(M_Y - M_{YS})]}.$$

Values for $n^* = 1$ to 5 and for an infinitely large n^* are shown in table 3. For the Peel data $\hat{r}^2$ is negative for an infinitely large n^* and so r is undetermined. The general impression is that at least four sites would be necessary per year to obtain a mean trial yield in agreement with the commercial yield for the region. Only if this condition were met would it be reasonable to apply the distribution of regional effects to the analysis of a set of varietal trials, with the overall mean for a year as the basic unit of analysis.

TABLE 3

Predicted correlations between the mean of a variety trial and the effect for a region, for various numbers of sites per trial

Region	Number of sites					
	1	2	3	4	5	∞
Coolamon	0.73	0.75	0.76	0.76	0.76	0.77
Culcairn	0.83	0.95	1.01	1.05	1.08	1.21
Peel	0.70	0.84	1.01	1.24	1.58	—
Narrandera	0.81	0.86	0.88	0.89	0.90	0.93
Narrandera	0.94	0.97	0.98	0.99	0.99	1.00
Average	0.80	0.87	0.93	0.98	1.06	0.98

4. VARIETAL RESPONSE AND APPLICATION OF THE THEORY

If $f(e)$ is known, then it may be used in the analysis of variety trial data, but in many cases suitable long-term data will not be available for the determination of $f(e)$ and more approximate analysis will be necessary. The two cases will be considered in turn.

(i) $f(e)$ known

Although $f(e)$ will invariably be only an estimate of the true distribution, any analysis is considerably simplified if it is assumed that both $f(e)$ and the individual values of e have been estimated without error.

Suppose that a variety is tested in m environments ($m \geq 2$), giving the

FIG. 1.—Hypothetical plot of observed variety yields against environmental effect, giving the function $g'(e)$.

means $g_1, g_2, \dots, g_m$. At least two years should preferably be represented if variety $\times$ year interaction is expected to be high. Let the corresponding effects for the environments be $e_1, e_2, \dots, e_m$, where we assume that $e_1 < e_2 < \dots < e_m$. These are the data from which the function $g(e)$ can be estimated, and hence μ_g by application of equation (1).

If g is plotted against e (fig. 1) the simplest procedure is to assume linearity between the observed points, giving the function $g'(e)$. An estimate of μ_g is then given by

$$\begin{aligned}\hat{\mu}_g &= \int_{e_1}^{e_m} g'(e)f(e)de / \int_{e_1}^{e_m} f(e)de \\ &= \sum_{i=1}^{m-1} \int_{e_i}^{e_{i+1}} g'(e)f(e)de / p_{1m},\end{aligned}$$

where, in general,

$$p_{jk} = \int_{e_j}^{e_k} f(e)de; \quad k > j = 1, 2, \dots, m-1.$$

But between e_1 and e_2

$$g'(e) = \frac{g_1 e_2 - g_2 e_1}{e_2 - e_1} + \frac{g_2 - g_1}{e_2 - e_1} e$$

so that

$$\int_{e_1}^{e_2} g'(e)f(e)de = \frac{g_1 e_2 - g_2 e_1}{e_2 - e_1} p_{12} + \frac{g_2 - g_1}{e_2 - e_1} m_{12},$$

where, in general,

$$m_{jk} = \int_{e_j}^{e_k} ef(e)de; \quad k > j = 1, 2, \dots, m-1.$$

It follows that

$$\hat{\mu}_g = (1/p_{1m}) \sum_{i=1}^{m-1} \left[\frac{g_i e_{i+1} - g_{i+1} e_i}{e_{i+1} - e_i} p_{i, i+1} + \frac{g_{i+1} - g_i}{e_{i+1} - e_i} m_{i, i+1} \right].$$

Since this is a linear function of the g_i , the variance of $\hat{\mu}_g$ may be estimated from the variance of the g_i .

(ii) $f(e)$ unknown

Suppose that a set of trials have been carried out within m environments for which no appropriate distribution of effects is available. To apply equation (1) it is still necessary to have estimates of e_i , and these will be available if

- (1) m is reasonably large, say 20 or more; or
- (2) if m is small, data are available for similar trials held at the same sites in other years.

Whichever is the case, it is important that more than one year be represented if the intention is to predict the future performance of varieties.

Let the number of "environments" represented be n , $n \geq m$, and denote the average trial yields by $Y_1, Y_2, \dots, Y_n$. Then an estimate of the effect for the i th environment is given by

$$e_i = \frac{Y_i - \bar{y}}{\bar{y}},$$

where

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n Y_i.$$

This linear transformation of the Y_i is consistent with the method suggested for estimating effects from long-term data. The assumption is that the genetic effect in each environment is the same, regardless of which varieties have been included.

To apply equation (1) the further assumption must be made that g is a linear function of e . Breese (1969) has reviewed the experimental evidence on this point and concluded that linearity is the rule, although Perkins and Jinks (1968) have detected deviations from linearity for 29 lines of *Nicotiana rustica*, attributable partly to the extreme genetic diversity. If we write

$$g = \alpha + \beta e \quad (2)$$

then

$$\begin{aligned} \mu_g &= \int (\alpha + \beta e) f(e) de \\ &= \alpha \end{aligned} \quad (3)$$

since, by definition, $E[e] = 0$. The estimate $\hat{\mu}_g$ is therefore given by the intercept at $e = 0$, when g is linearly regressed against e . In general terms, $\hat{\mu}_g$ is the predicted yield in the average environment. A similar measure has been proposed by Pinthus (1973), with a standard variety being used to position the average environment on the e -scale.

There may be genotypes for which a quadratic function is the best description of $g(e)$.

i.e.

$$g = \alpha + \beta_1 e + \beta_2 e^2. \quad (4)$$

Then

$$\begin{aligned} \mu_g &= \int (\alpha + \beta_1 e + \beta_2 e^2) f(e) de \\ &= \alpha + \beta_2 \text{var}(e) \end{aligned}$$

so that if the fitted curve is

$$\hat{g} = a^* + b_1^*(e - \bar{e}) + b_2^*(e^2 - \bar{e}^2),$$

where

$$\bar{e} = \frac{1}{m} \sum_{j=1}^m e_j$$

and

$$\bar{e}^2 = \frac{1}{m} \sum_{j=1}^m e_j^2,$$

then

$$\hat{\mu}_g = a^* - b_1^* \bar{e} - b_2^* (\bar{e}^2 - \widehat{\text{var}}(e)), \quad (5)$$

where

$$\widehat{\text{var}}(e) = \frac{1}{n-1} \left[\sum_{i=1}^n e_i^2 - \left(\sum_{i=1}^n e_i \right)^2 / n \right].$$

The variance of μ_g may be estimated from the variances and covariances of a^* , b_1^* and b_2^* , assuming for simplicity that the e_j and their variance have been estimated without error. The variance of μ_g will be the same for all varieties if a pooled error variance is used, since the variances and covariances of the regression estimates depend only on the e_j and the error variance.

It has been stated that if m is sufficiently large then the estimates of environmental effect may be obtained from the data intended for analysis. In this situation $\bar{e}$ will be zero since $m = n$, and equation (3) will give

$$\begin{aligned}\hat{\mu}_g &= \hat{\alpha} \\ &= \bar{g}.\end{aligned}$$

The estimated mean of a variety is therefore the arithmetic mean over all environments, and this will also be the case for a small number of trials for which, by chance, the average effect is close to zero. Under the quadratic assumption, equation (5) gives

$$\begin{aligned}\hat{\mu}_g &= a^* - b_2^*(\overline{e^2} - \widehat{\text{var}}(e)) \\ &= \bar{g} - b_2^*\left(\frac{1}{m} \sum_{j=1}^m e_j^2 - \frac{1}{m-1} \sum_{j=1}^m e_j^2\right) \\ &= \bar{g} + b_2^* \frac{1}{m(m-1)} \sum_j e_j^2 \\ &= \bar{g} + b_2^* \widehat{\text{var}}(e)/m \\ &\rightarrow \bar{g} \quad \text{as } m \rightarrow \infty.\end{aligned}$$

Although $\hat{\mu}_g$ may be different from $\bar{g}$ when m is small, for a large m there should be little difference even when $g(e)$ is quite non-linear.

The conclusion is that the usual analysis of variance, with calculation of the arithmetic means, should be carried out for a large set of trials. The most favourable circumstances for application of the present analysis would be when there were few trials, preferably involving at least two years, with a wide range of effects for which the average was removed from zero.

5. AN EXAMPLE

Data on wheat grain yield were obtained from three sites within the Narrandera region for the years 1965 to 1967. The data from one site in 1966 were discarded since the yields were inordinately low in comparison with the remaining two sites, leaving a total of eight sites for analysis.

Since there was a maximum of three sites per year, it was not possible to use the commercial yields for the region to estimate the distribution of effects. Again, eight sites were too few for the trial data to be used to estimate a set of effects, and use was made of 15 similar trials carried out at the three sites between 1963 and 1970. The average yield was determined for each trial in the composite set and the effects were calculated by the method given in

the previous section. The effects for the eight trials, plus the mean yields for four varieties, are shown in table 4.

TABLE 4
Grain yields (kg/ha) for four wheat varieties at eight sites, and the estimated effect for each site

Variety	Site Number								Mean
	1	2	3	4	5	6	7	8	
Falcon	1116	1211	1621	3087	3154	935	518	928	1571
Heron	1466	1527	1736	3262	3255	1015	673	881	1727
Olympic	1056	1184	1749	3558	3309	639	464	915	1609
Wren	565	679	1412	3874	3894	545	404	854	1528
Effect	-0.532	-0.440	-0.279	0.456	0.357	-0.665	-0.718	-0.639	-0.308

The data have been plotted in fig. 2, with the fitted least-squares regression lines. To test whether the intercepts at zero are significantly different, the model

$$Y_{ij} = \mu + \alpha_i + \beta_i(e_j - \bar{e}) + \varepsilon_{ij}$$

was first fitted. Y_{ij} is the yield of the i th variety at the j th site. The least-squares estimates of the effects are

$$m = (\sum_{ij} Y_{ij})/32$$

$$= \bar{y}..,$$

$$a_i = (\sum Y_{ij})/8 - \bar{y}..$$

$$= \bar{y}_{i.} - \bar{y}..,$$

and

$$b_i = \sum_j (Y_{ij} - \bar{y}_{i.})(e_j - \bar{e}) / \sum_j (e_j - \bar{e})^2.$$

The sum of squares due to these effects is

$$R(\mu, \alpha, \beta) = 32\bar{y}^2.. + 8 \sum_i (\bar{y}_{i.} - \bar{y}..)^2 + \sum_i [b_i \sum_j (Y_{ij} - \bar{y}_{i.})(e_j - \bar{e})].$$

The model

$$Y_{ij} = \mu' + \beta'_i e_j + \varepsilon'_{ij}$$

was then fitted, the least-squares estimates being

$$m' = \bar{y}.. - \bar{e} \cdot \sum_i b_i / 4$$

and

$$b'_i + (\sum_j Y_{ij} e_j / \sum_j e_j^2) - m' \sum_j e_j / \sum_j e_j^2.$$

The sum of squares due to these effects is

$$R(\mu, \beta) = m' Y.. + \sum_i (b'_i \sum_j Y_{ij} e_j).$$

The difference $R(\alpha | \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta)$ is attributable to including α in the model, and provides a test of the hypothesis that $(\mu + \alpha_i - \beta_i \bar{e})$ is the same for all i . The analysis of the data in table 4 is presented in table 5, the conclusion being that the intercepts are not significantly different.

FIG. 2.—Variety yields (kg/ha) plotted against environmental effect, showing the least-squares regression lines.

TABLE 5

A regression analysis of the data in table 4. The sums of squares and mean squares have been divided by 10^6

Source	d.f.	S.S.	M.S.	V.R.
$R(\mu, \alpha, \beta)$	8	1217.324	152.166	
$R(\mu, \beta)$	5	1215.438	243.088	
$R(\alpha \mu, \beta)$	3	1.886	0.629	1.81
Deviation	24	8.340	0.348	
Total	32	1225.664		

TABLE 6

Arithmetic mean yields and estimates of mean yield (kg/ha) on the linear and quadratic assumptions, for the data in table 4

Variety	Type of mean		
	Arithmetic	Linear	Quadratic
Falcon	1571	2234 $\pm$ 81	2234 $\pm$ 81
Heron	1727	2391 $\pm$ 81	2390 $\pm$ 81
Olympic	1609	2398 $\pm$ 81	2398 $\pm$ 81
Wren	1528	2497 $\pm$ 81	2501 $\pm$ 81

An estimate of the variance of $\hat{\mu}_g$ is
 (deviation mean square) $[1/8 + \bar{e}^2/\sum_j (e_j - \bar{e})^2]$.

The estimate of μ_g and its standard error for each variety are given in table 6.

From fig. 2 it can be seen that a quadratic function would give a better fit to the Wren data, and may be expected to reduce the deviation sum of squares of 5.58×10^5 which Wren contributes to the total of 8.34×10^5 under a linear model. A quadratic curve was therefore fitted for each variety, as shown in fig. 3. Application of equation (5) then gave the estimates of μ_g shown in table 6. The four deviation sums of squares, one for

FIG. 3.—Variety yields (kg/ha) plotted against environmental effect, showing the least-squares quadratic curves.

each variety, were pooled to give a deviation mean square of 0.281×10^5 with 20 degrees of freedom, from which the variance of μ_g was estimated to be 0.0654×10^5 .

The estimated yield under a linear model differs only slightly, if at all, from the estimated yield under a quadratic model (table 6). If this is a general result then it would argue in favour of the more simple linear model. The second, more important, point is that the ranking of varieties on the basis of the arithmetic mean differs from the ranking on the basis of μ_g -values. This is mainly due to the low yield of Wren in poor environments, which were over-represented among the eight sites compared with the long-term

expectation. It is not suggested that the change in ranking would be a feature of all variety trials since the present example was chosen specifically to illustrate the point, but it is probably as well to be aware of the circumstances in which such a result could occur, namely

- (1) the number of test environments is small;
- (2) there is a high variety $\times$ environment interaction;
- (3) The "average" of the test environments is either well above or well below the long term average.

Acknowledgments.—The author is grateful to the Division of Plant Industry, New South Wales Department of Agriculture, for permission to use the data from their annual wheat variety trials. Thanks are due to Professor J. L. Jinks for his valuable comments on a first draft of the paper.

6. REFERENCES

- BREESE, E. L. 1969. The measurement and significance of genotype-environment interactions in grasses. *Heredity*, 24, 27-44.
- DAVIDSON, B. 1962. Crop yields in experiments and on farms. *Nature*, 194, 458-459.
- EBERHART, S. A., AND RUSSELL, W. A. 1966. Stability parameters for comparing varieties. *Crop Sci.*, 6, 36-40.
- EISENHART, G. 1947. The assumptions underlying the analysis of variance. *Biometrics*, 3, 1-21.
- FINLAY, K. W., AND WILKINSON, G. N. 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. agric. Res.*, 14, 742-754.
- PEDERSON, D. G. 1974. The stability of varietal performance over years. 1. The distribution of seasonal effects for wheat grain yield. *Heredity* (in press).
- PERKINS, JEAN M., AND JINKS, J. L. 1968. Environmental and genotype-environmental components of variability. IV. Non-linear interactions for multiple inbred lines. *Heredity*, 23, 525-535.
- PINTHUS, MOSHE J. 1973. Estimate of genotypic value: a proposed method. *Euphytica*, 22, 121-123.
- THOMSON, A. J., AND WRIGHT, A. J. 1971. Principles and problems in grass breeding. Annual report, Plant Breeding Inst., Cambridge, 31-67.
- YATES, F., AND COCHRAN, W. G. 1938. The analysis of groups of experiments. *J. Agric. Sci. Camb.*, 28, 556-580.