A MODIFIED PROCEDURE FOR IDENTIFYING VARIETAL STABILITY

Nasrullah *)

Summary

An alternative method for identifying varietal stability was discussed. The procedure, based on a model similar to that of Eberhart and Russell (1966) $Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}$ where Y_{ij} the yield of the i^{th} variety ($i = 1,2, ..., v$) planted in the j^{th} environment ($j = 1, 2, ..., n$) is regressed on I, the environmental index, was developed by selecting an index of environment such that is is independent of Y_{ij} :

$$I_{ij} = \sum_{\substack{u=1 \\ u \neq i}}^{v} Y_{uj}/(v-1)$$

Numerical example was given using data from experiments conducted in tidal swamp area in Kalimantan (Soemartono, 1977). The two stability parameters are comparable and of identical ranks with those of Eberhart and Russell. A noted difference is that the sum of squares of deviation from the regression is consistently larger than that of Eberhart and Russell. It is understandably due to the exclusion of Y_{ij} from the environmental index.

Ringkasan

Dengan mendasarkan model serupa yang dipakai oleh Eberhart dan Russell (1966) $Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}$ di mana Y_{ij} produksi varitas ke i ($i = 1, 2, ..., v$) yang ditanam pada keadaan lingkungan ke j ($j = 1, 2, ..., n$) diregresikan terhadap index lingkungan I, suatu cara lain untuk meninjau stabilitas suatu varitas tanaman dikemukakan dalam tulisan ini. Prosedur ini meskipun masih menggunakan index lingkungan yang berdasarkan pada produksi varitas yang diuji memenuhi asumsi pada analisa regresi, yaitu bahwa index lingkungan dipilih sedemikian rupa sehingga bebas stokhastik terhadap Y_{ij} sebagai berikut :

$$I_{ij} = \sum_{\substack{u=1 \\ u \neq i}}^{v} Y_{uj}/(v-1)$$

Cara tersebut diilustrasikan dengan menggunakan data penelitian daerah pasang surut di Kalimantan (Soemartono, 1977). Kedua index stabilitas menunjukkan ranking yang sama dengan index serupa dari Eberhart dan Russell. Index stabilitas kwadrat penyimpangan terhadap garis regresi selalu lebih besar dari index yang sama dari Eberhart dan Russell. Hal ini disebabkan oleh tidak diikutsertakannya Y_{ij} dalam index lingkungan.

Introduction

It is often difficult to obtain valid conclusions from varietal trials carried out over a wide range of environment. The confusion comes from the fact

*) Departement of Agronomy, Faculty of Agriculture, Gadjah Mada University, P.O. Box 1, Yogyakarta, Indonesia.

that when varieties are compared over a series of environments, the relative rankings usually differ because varietal performance changes with the change in environment. These changes are known as genotype x environment interactions in biometrical analysis. Utz et al. (1978) presented a way to reduce those interactions in the analysis of varietal trials.

A number of authors (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966) have shown that in many such cases the performance of an individual variety can be expressed as a linear function of the environmental index. This regression provides two simple measures of the sensitivity of a variety to environmental change : a) the regression coefficient, and b) the deviation from regression mean squares.

The procedure violates the assumptions in regression analysis, because the yield of individual variety is also part of the environmental index, the regressor. The present paper were aimed to provide an alternative method for assessing varietal stability by selecting an environmental index which, though still as a linear combination of the yield of the varieties tested, is independent.

Empirical Model

Let the model be

$$Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}, i = 1, 2, ..., v; j = 1,2, ...,n \quad (1)$$

where Y_{ij} is the variety mean of the i^{th} variety at the j^{th} environment, μ_i is the mean of the i^{th} variety over all environments, β_i is the regression coefficient that measures the response of the i^{th} variety to varying environments, δ_{ij} is the deviation from regression of the i^{th} variety at the j^{th} environment, and I_j is the environmental index. If I_j is selected such that $\Sigma_j I_j = 0$, then the first stability parameter is a regression coefficient estimated in the usual manner

$$\hat{\beta}_i = \frac{\Sigma_j Y_{ij} I_j}{\Sigma_j I_j^2} \quad (2)$$

The sum of squares of deviation from the regression for the i^{th} variety is

$$\Sigma_j \hat{\delta}_{ij}^2 = (\Sigma_j Y_{ij}^2 - \frac{Y_i^2}{n}) - \frac{(\Sigma_j Y_{ij} I_j)^2}{\Sigma_j I_j^2} \quad (3)$$

with (n — 2) degrees of freedom and can be used to provide an estimate of another stability parameter ($\delta^2_{d_i}$) :

$$s^2_{d\,i} = \Sigma_j \delta^2_{i\,j} \;/(n-2) - s^2_e \;/r \tag{4}$$

where $s_e{}^2$ is the pooled error mean squares and r is the replication. Summing equation (3) over all varieties gives a pooled deviation sum of squares with v(n-2) d.f.

On imposing a common slope β in the model, equation (1) becomes

$$Y_{ij} = \mu_i + \beta I_j + (\beta_i - \beta)\, I_j + \delta_{ij}$$
$$= \mu_i + \beta I_j + \varepsilon_{ij} \text{ where } \varepsilon_{ij} = (\beta_i - \beta)\, I_j + \delta_{ij}$$

The estimate of β is

$$\hat{\beta} = \frac{1}{v} \frac{\Sigma_j Y_{.j} I_j}{\Sigma_j I_j^2} \tag{5}$$

and the sum square due to $\hat{\beta}$ is

$$\text{SS due to } \hat{\beta} = \frac{1}{v} \frac{(\Sigma_j Y_{.j} I_j)^2}{\Sigma_j I_j^2} \tag{6}$$

The sum of squares of deviations from the regression using the common slope is

$$\Sigma_{i,j} \varepsilon^2_{ij} = \Sigma_{i,j} \;(Y^2_{ij} - \frac{Y^2_{i.}}{n}) - \frac{1}{v} \frac{(\Sigma_j Y_{.j} I_j)^2}{\Sigma_j I_j^2}$$

with (vn - v - 1) degrees of freedom. The sum of squares due to $\beta_i \mid \beta$ = $\Sigma_{i,j} \; (\varepsilon^2_{ij} - \delta^2_{ij})$

$$= \frac{\Sigma_i (\Sigma_j \; Y_{ij} I_j)^2}{\Sigma_j I_j^2} - \frac{1}{v} \frac{(\Sigma_j \; Y_{.j} I_j)^2}{\Sigma_j I_j^2} \quad \text{with (v-1) d.f.} \tag{8}$$

The appropriate analysis of variance is given in Table 1.

In essence, Finlay and Wilkinson (1963) and Eberhart and Russell (1966) used site mean, mean yield of all varieties for each site, as an environmental index.

$$I_j = \Sigma_j \; Y_{ij}/v - \Sigma_{i,j} \; Y_{ij}/vn = \bar{Y}_{.j} - \bar{Y}_{..} \tag{9}$$

$$\text{Consequently } \hat{\beta} = \Sigma_j \; Y_{.j} I_j / (v \; \Sigma_j \; I_j^2)$$
$$= [\Sigma_j (vI_j + v\bar{Y}_{..}) I_j] / (v \; \Sigma_j \; I_j^2)$$
$$= 1$$

Note also that $\Sigma_i \hat{\beta}_i/v = (\Sigma_{i,j} Y_{ij}I_j)/(v \Sigma_j I_j^2)$

$$= [\Sigma_j (I_j \Sigma_i Y_{ij})]/(v \Sigma_j I_j^2)$$

$$= (\Sigma_j Y_{.j} I_j)/(v\Sigma_j I_j^2)$$

$$= \hat{\beta}$$

A desirable variety is therefore one with a high mean ($\bar{X}_{i.}$, unit regression coefficient ($\beta_i = 1.0$) and the deviation from regression as small as possible ($s^2_{d_i} = 0$).

The drawback of using site mean as a regressor is it violates regression methodology assumptions. When an individual variety is regressed on the site mean, the yield of that variety is the dependent variable, but is also part of the site mean, the independent variable. The error of measurement in the independent variable is correlated with dependent variable. Nor and Cady (1979) avoided that inconsistency using multivariate statistical method by taking the first principal component of independent measurements of the environment (Morrison, 1976). If we let $I_{ij} = \Sigma_{u=1, u \neq i}^{v} Y_{uj}/(v - 1)$ as an environmental index, I_{ij} is independent from Y_{ij}. The model now is

$$Y_{ij} = \mu_i + \beta_i I_{ij} + \delta_{ij} \tag{10}$$

and the estimates of regression coefficients

$$\hat{\beta}_i = \frac{\Sigma_j (Y_{ij} - \bar{Y}_{i.})(I_{ij} - \bar{I}_{i.})}{\Sigma_j (I_{ij} - \bar{I}_{i.})^2} \tag{11}$$

replace equation (2). Equations (3), (5), (6), and (8) are replaced by

$$\Sigma_j \hat{\delta}^2_{ij} = \Sigma_j (Y^2_{ij} - \frac{Y^2_{i.}}{n}) - \frac{[\Sigma_j (Y_{ij} - \bar{Y}_{i.})(I_{ij} - \bar{I}_{i.})]^2}{\Sigma_j (I_{ij} - \bar{I}_{i.})^2} \tag{12}$$

$$\hat{\beta} = \frac{\Sigma_{i,j}(Y_{ij} - \bar{Y}_{i.})(I_{ij} - \bar{I}_{i.})}{\Sigma_{i,j}(I_{ij} - \bar{I}_{i.})^2} \tag{13}$$

$$\text{SS due to } \hat{\beta} = \frac{[\Sigma_{i,j} (Y_{ij} - \bar{Y}_{i.})(I_{ij} - \bar{I}_{i.})]^2}{\Sigma_{i,j}(I_{ij} - \bar{I}_{i.})^2}$$

Table 1. **Analysis of Variance**

Source	d.f.	S.S.	M.S.
Total	nv-1	$\Sigma_{i,j}\ Y^2_{ij}$ - CF	
Varieties (V)	v-1	$\Sigma_i\ Y^2_{j.}/n$ - CF	MS_1
Environements (Env) V X Env	v(n-1)	$\Sigma_{i,j}\ Y^2_{ij} - \Sigma_i\ Y^2_{i.}/n$	
Env (linear)	1	$(1/v)(\Sigma_j\ Y_{.j}\ I_j)^2/\Sigma_j\ I^2_j$	
V x Eny (linear)	v-1	$\Sigma_i(\Sigma_j\ Y_{ij}\ I_j)^2 /\ j\ I^2_j$ - Env (linear) SS	MS_2
Pooled deviations	v(n-2)	$\Sigma_{i,j}\ \hat{\delta}^2_{ij}$	MS_3
Variety 1	n-2	$\Sigma_j\ Y^2_{1j} - \frac{Y^2_{1.}}{n} - (\Sigma_j Y_{1j} I_j)^2/\Sigma_j\ I^2_j$	
.	.	.	
.	.	.	
.	.	.	
Variety v	n-2	$\Sigma_j\ Y^2_{vj} - \frac{Y^2_{v.}}{n} - (\Sigma_j Y_{vj} I_j)^2/\Sigma_j\ I^2_j = \Sigma_j \hat{\delta}_{vj}$	
Pooled error	n(r-1)(v-1)		

Table 2. **Analysis of Variance And Coefficients of Regression for Each Variety.**

Source	d.f	Proposed S.S	Method $\hat{\beta}$	Eberhart and S.S	Russell $\hat{\beta}$
Total	77				
Variety (V)	10				
Env (linear)	1	3,152.67		3,450.81	
V x Env (linear)	10	513.52		433.09	
Pooled deviations	66	1,235.45		1,017.74	
IR 26	6	41.86	1.00	34.52	1.01
IR 28	6	117.65	.37	109.15	.42
IR 29	6	69.18	.72	60.05	.76
IR 30	6	97.58	.48	88.68	.53
IR 34	6	93.70	.60	78.87	.67
C4-63	6	236.68	1.05	192.58	1.11
Pelita I/1	6	171.03	1.32	132.89	1.33
Pelita I/2	6	78.06	1.11	71.14	1.11
PB 5	6	54.63	1.10	43.50	1.11
A7/PsJ/72K	6	68.41	1.55	51.22	1.49
A8/PsJ/72K	6	206.67	1.51	155.14	1.50
Pooled error					

$$\text{SS due to } \hat{\beta}_i | \hat{\beta} = \frac{\Sigma_i [\Sigma_j (Y_{ij} - \bar{Y}_{i.}) \; (I_{ij} - \bar{I}_{i.})]^2}{\Sigma_j (I_{ij} - \bar{I}_{i.})^2} - \text{SS due to } \hat{\beta}$$

The analysis of variance is like the one in Table 1 after appropriate substitutions.

Numerical Example

The proposed procedure is applied to data of varietal trials conducted in tidal swamp area in Kalimantan (Soemartono, 1977). Eleven high yielding varieties were tested in eight environments, a combination of location and year. The analysis of variance and the regression coefficients of each variety along with those obtained using Eberhart and Russell's approach are shown in Table 2. It is evident that the relative rankings for both coefficients of stability is comparable and identical. The sum of squares of deviations from the regression for the i^{th} variety of the proposed approach is consistently higher than that of Eberhart and Russell. It was expected as the exclusion of Y_{ij} from the environmental index reduces the fitness of the regression and eventually inflates the deviation.

References

EBERHART, S.A. and W.A. RUSSELL. (1966) Stability parameters for comparing varieties. **Crop Sci.** 6 : 36 - 40.

FINLAY, K.W. and G.N. WILKINSON. (1963) The analysis of adaptation in a plant breeding programme. **Aust. J.** Agric. Res. 14 : 742 - 754.

MORRISON, D.F. (1976) **Multivariate Statistical Methods.** 2nd ed. McGraw-Hill Kogakusha, Tokyo. 415 pp.

NOR, K.M. and F.B. CADY. (1979) Methodology for identifying wide adaptability in crops. **Agron. J.** 71 : 556 - 559.

SOEMARTONO, (1977) Prospek varietas padi daerah pasang surut di Indonesia. Kertas kerja pada **Simposium I Peranan Hasil Penelitian Padi dan Palawija dalam Pembangunan Pertanian, 26 — 29 September 1977 di Maros.**

UTZ, H.F., F.W. SCHNELL, and I.S. SINGH. (1978) Maturity corrected yield (MCY) as a way to reduce genotype environment interactions in maize trials. **Cereal Res. Comm.** 6 : 399 — 404.