Euphytica 39: 11–18 (1988)

547

Jn 647

Estimates of stability for comparing varieties

J.R. Witcombe
Pearl Millet Improvement Program, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Andhra Pradesh, India

Received 3 December 1986; accepted in revised form 1 July 1987

Key words: Pennisetum americanum, pearl millet, stability analysis, multilocational trials.

Summary

For purposes of comparison of analytical methods the yield data on pearl millet hybrids and varieties for five years from locations in India and Pakistan have been analysed using a regression analysis and a mean-standard deviation analysis. The results of the mean-standard deviation analysis and the regression analysis were similar whether carried out on all environments, or on the highest- and lowest-yielding sets of environments. This was substantiated by the remarkable correlation between the slopes from the regression analysis and the standard deviations whatever environmental set was considered. The validity of using a single year's across location data with the mean-standard deviation analysis, if choice-theoretic criteria are used, was examined. It is concluded that, although further confirmation is required, single-year data seem to suffice. The relative merits of the mean-standard deviation analysis and a regression analysis are discussed; the choice-theoretic framework of the mean-standard deviation analysis is advantageous and complementary to the regression analysis.

Introduction

Over five years and across many locations in India and Pakistan, open-pollinated varieties and hybrids have been tested in the International Pearl Millet Adaptation Trial (IPMAT) coordinated by ICRISAT and run by many cooperators. Yield data have been used to examine two contrasting approaches to the analysis of stability.

Several statistical techniques have been developed to analyse the interaction of genotypes with environments (G × E) and regression analyses have been extensively used. The first proposal of a regression analysis to study G × E interactions was by Yates & Cochran (1938). Modifications of this method have been used by Finlay & Wilkinson (1963), Eberhart & Russell (1966), and Perkins & Jinks (1968). In all these methods the mean squares for the major components of variation are identical or closely related.

In Finlay & Wilkinson (1963) the average yield across all the genotypes in an environment is used as an assessment of that environment. In other analyses the overall mean yield is subtracted from the individual environment mean, but whatever method is followed the end result of the analysis is the same. This biological assessment of the environment is generally termed an environmental index. The individual yield data of the genotypes are regressed on the environmental indices, giving estimates of three main parameters that describe the performance of each genotype: (i) The mean yield

Approved as JA no. 647 by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

of the genotype over all environments, (ii) The regression coefficient of the genotype, and (iii) The remainder mean square (RMS). Genotypes with a slope above one are considered to be less stable, and those with a slope less than one more stable by Finlay & Wilkinson (1963), whilst Eberhart & Russell (1966) define a stable genotype as having a slope of one and a RMS of zero. The RMS indicates how the yield data meet the linear model. Eberhart & Russell (1966) subtract a constant, which is a measure of inexplicable residual variation, from the RMS. The value thus obtained, termed S^2_d, obviously bears a simple relationship to the RMS.

Regression analyses have particular limitations when economic criteria, such as the trade off between yield and yield stability, are considered and Binswanger & Barah (1980) prefer a mean-standard deviation analysis. The method examines total variation in the context of decision theory by trading off improvement in mean yield against increased variation as measured by the standard deviation. Variation of a genotype across years measures stability over time, which is the only component of variation or risk which is relevant for a farmer contemplating adoption of a variety in any given location. Variation of a genotype across locations is a measure of adaptability. Stability and adaptability were defined in this way by Evenson et al. (1978) and have been employed in this paper when discussing the mean-standard deviation analysis. The term stability when used for the regression analysis is as defined by Eberhart & Russell (1966).

The analyses of the results of IPMAT are compared using regression and mean-standard deviation analysis. Since there are both hybrids and open-pollinated varieties in these trials, the stability of these have also been compared.

Materials and methods

The entries in the trial of pearl millet, *Pennisetum americanum* (L.) Leeke, are primarily ICRISAT products. The entries, which apart from a few controls change each year, are hybrids and open-pollinated varieties. The trial has been run by cooperators in locations in India and Pakistan (Table 1). It is planted at all locations in a randomized block design with three replications. It is grown in the rainy season with added fertilizer, and irrigation when required. Data are recorded for several characters.

Only grain-yield data were analysed. The regression analysis was done according to the method of Eberhart & Russell (1966). This employs the model:

$$Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}$$

where Y_{ij} is the variety mean of the ith variety at the jth environment, μ_i is the mean of the ith variety over all environments, β_i is the regression coefficient that measures the response of the ith variety to varying environments, δ_{ij} is the deviation from the regression of the ith variety at the jth environment, and I_j is the environmental index obtained as the mean of all varieties at the jth envi-

Table 1. Summary of trials analysed.

Year[1]	Number of entries analysed	Number of		Number of locations analysed	
		Hybrids	Varieties	High-yielding	Low-yielding
1979	19	7	12	7	7
1980	19	8	11	10	9
1981	18	7	11	8	7
1983	21	7	14	9	8
1984	23	9	14	9	9

[1] Trial not held in 1982.

ronment minus the grand mean.

The mean-standard deviation analysis of Binswanger & Barah (1980) was done using their model for a single year's multilocational testing. Their model is:

$$Y_{ijt} = \mu_i + \lambda_{ij} + \tau_{ijt}$$

where μ_i is the genotype effect (as it is genotype yield in a specific year the year effect and the genotype × year interaction is also included in this term), λ_{ij} is the sum of the location effect and the location × genotype interaction, and τ_{ijt} is the location × genotype × year interaction. In applying their model to a single year's data of genotypes across locations, the year by location interaction contributes to the variation observed for Y_{ijt}, and this interaction can be assessed given a further year's data. Because λ_{ij} and τ_{ijt} cannot be estimated separately the standard deviations (s.d.'s) which are obtained from a single year's data overestimate stability and thus exaggerate the riskiness of the genotypes (Binswanger & Barah, 1980). Therefore, these s.d.'s are termed variability-relevant s.d.'s and are neither estimates of stability nor of adaptability and are larger than either of them individually. They are simply derived by calculating the across location s.d. for each genotype. For the method of calculation of the adaptability-relevant s.d. and the stability-relevant s.d. see Barah et al., 1981.

Using data given in Barah et al. (1981) for sorghum trials in India, the correlation of s.d.'s for variability (from single-year data across locations) with s.d.'s for stability over years was determined.

Results and discussion

Regression analysis. The hybrids were on average marginally higher-yielding than the varieties. They differed little from the varieties in their regression coefficients but had consistently higher S^2_d values (Table 2).

Since the varieties had lower S^2_d values than the hybrids the importance of the deviations from the regressions needs to be considered. Therefore, the amount of genotypic variation in the experiments was related to the amount of variation between environments (Table 3). The environments had by far the largest mean square, the interactions of the genotypes with the environment the next largest, and the genotypes the smallest. For G × E interaction, the mean square due to regression coefficients was invariably much smaller than that due to deviations from the regressions. Since the varieties were superior to hybrids for the deviations from the regressions (S^2_d values) it is possible that by using varieties crop yield variation is reduced. However, a regression analysis does not provide a decision-theoretic criterion for trading off a higher mean yield against reduced stability, particularly as each genotype has two parameters for stability-slope and S^2_d.

Mean-standard deviation analysis. The analysis of Binswanger & Barah (1980) provides a way in which the benefits of reduced variability are measured against loss in yield. The variability-relevant s.d.'s (as in this case the s.d.'s are for across locations within years) are plotted against mean yields using the same scale for both axes (Fig. 1). Entries

Table 2. Mean grain yields (kg ha^{-1}), mean S^2_d values, and mean regression coefficients for the hybrids and varieties.

Year	Mean grain yield		Mean S^2_d		Mean slope	
	Hybrids	Varieties	Hybrids	Varieties	Hybrids	Varieties
1979	2300	2237	163,533	61,222	1.02	0.98
1980	2096	1974	62,791	33,085	1.01	0.99
1981	2236	2262	117,970	33,189	1.04	0.97
1983	2068	2109	157,951	50,968	0.97	1.02
1984	2028	1773	62,228	40,166	1.07	0.96

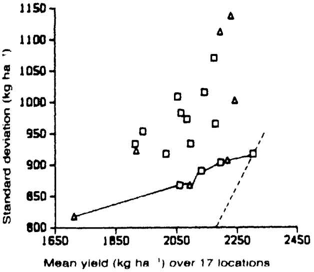

Fig. 1. Mean yield and standard deviation of the entries in 1983 (Iso-utility curve ---, risk efficient frontier ——, hybrids △, varieties □)

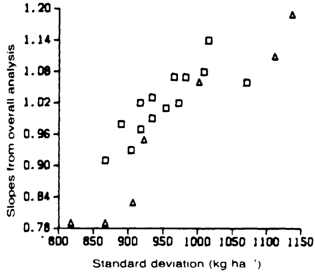

Fig. 2. Relationship between the regression coefficients and standard deviations of the entries in 1983 (Hybrids △, varieties □)

of equal utility lie on a line which is called an iso-utility curve. The slope of the line is 2.0 for representative, risk-averse farmers (Binswanger & Barah, 1980). Later studies in Southeast Asia and Central Asia, using a similar experimental approach have confirmed that the value of 2.0 for the slope of the iso-utility curve is appropriate in assessing the trade off between yield and s.d. (Binswanger & Sillers, 1983).

In each year a line is drawn connecting entries that are risk efficient because for a given level of yield performance, they have the lowest s.d. That line maps a risk efficient frontier which is also indicated in Fig. 1. The best entry must lie on this frontier, and it is chosen with regard to the iso-utility curve. In no year was an entry chosen other than the one that was highest yielding over all environments. Thus, in no case did reduced variability compensate for lower yield.

Correlation of variability, stability and adaptability s.d.'s. Barah et al. (1981) examined sorghum trials data across four years and eighteen locations. Correlations between the s.d.'s (for variability, stability and adaptability) presented in this paper were calculated (Table 4). In three of four years variability s.d.'s correlated well with adaptability and stability s.d.'s, indicating the appropriateness of single-year data for use in a choice-theoretic framework.

Table 3. Genetic variation relative to environmental variation standardised to 100 from the mean squares of a regression analysis.[1]

Source of variation	Year				
	1979	1980	1981	1983	1984
Environment (linear)	100	100	100	100	100
Genotype (= mean values)	2.2	1.5	4.1	2.8	6.6
G × E (linear = b values)	0.6	1.0	2.5	1.1	1.9
G × E (deviation = S^2_d values)	14.9	12.8	16.3	19.7	19.0

[1] Using the model of Eberhart & Russell (1966).

Table 4. Correlation coefficients[1] between s.d.'s for variability, stability and adaptability from data for sorghum presented in Barah et al. (1981).

	Stability s.d.	Adaptability s.d.
Variability s.d. 1971	0.93	0.79
Variability s.d. 1972	0.69	0.56
Variability s.d. 1973	0.73	0.74
Variability s.d. 1974	0.36	0.34
Mean variability s.d.	0.95	0.85

[1] r>.58 significant at p = .05; r>.71 significant at P = .01; r>.82 significant at P = .001. Sample size in all years = 11.

Comparison of the two analyses. The two methods of analysis were tested over a range of environments to see if a comparison of the methods is environment dependent. In all years the lowest-yielding set of environments, the highest-yielding set, and all environments were analysed separately.

Selection of the best entry was then made using the results from both methods of analysis in all years, and in the three sets of environments (Table 5). In all years the first choice was the same in the two methods (Table 5). In the regression analysis there is no choice-theoretic framework; the highest-yielding entry was invariably chosen, without regard to slope or RMS, because in no case was the highest-yielding entry exceptionally unstable for either of these parameters. In the mean-standard deviation analysis, the top-yielding entry was always chosen because it was always the highest in a preference-based ranking as well.

For each year for all three sets of environments the correlation between the slopes and S^2_d values of the regression analysis and the s.d.'s of the mean-standard deviation analysis was determined (Table

Table 5. Entries selected on the basis of the two methods of analysis over five years.

Year	Entry number of entry selected on b and $\bar{x}$			Entry number of entry selected on sd and x					
	All locations	High-yield locations	Low-yield locations	All locations		High-yield locations		Low-yield locations	
				1st choice	2nd choice	1st choice	2nd choice	1st choice	2nd choice
1979	3	3	12	3	5 or 12	3	5 or 13	12	3
1980	8	7	2	8	1	7	8	2	4 or 8
1981	6	6	6	6	2 ?*	6	5	6	2
1983	10	10	2	10	2 ?	10	?	2	15
1984	23	23	23	23	20	23	20	23	3

*? indicates several possibilities for 2nd choice.

Table 6. Correlation between various statistical parameters of the two analyses.

Year	Environments	Correlation coefficients[1] between parameters:[2]				
		sd,b	sd,S^2_d	b,S^2_d	b,$\bar{x}$	sd,$\bar{x}$
1979	all	0.73	0.40	− 0.14	0.54	0.34
	high	0.51	0.66	− 0.17	− 0.09	− 0.19
	low	0.83	0.50	0.11	0.31	0.52
1980	all	0.96	0.02	− 0.17	0.64	0.63
	high	0.81	0.22	− 0.23	0.60	0.38
	low	0.78	0.22	− 0.35	0.42	0.52
1981	all	0.80	0.37	− 0.13	0.65	0.71
	high	0.72	0.33	− 0.36	0.32	0.51
	low	0.69	0.46	0.23	0.29	0.28
1983	all	0.86	0.44	0.01	0.43	0.47
	high	0.74	0.76	0.39	− 0.13	− 0.06
	low	0.91	0.68	0.48	0.42	0.60
1984	all	0.90	0.35	− 0.08	0.63	0.49
	high	0.70	0.50	− 0.20	− 0.01	0.16
	low	0.87	0.53	0.10	0.24	0.47

[1] $r>.44$ significant at $P = .05$; $r>.56$ significant at $P = .01$; $r>.68$ significant at $P = .001$.
[2] sd = standard deviation; b = regression coefficient; S^2_d = deviations from regressions (adjusted); $\bar{x}$ = mean across locations.

6). The relationship between the slopes and the s.d.'s was highly significant (at P = 0.001) in all years and all environmental sets (except for the high-yielding set of environments in 1979). An example of this relationship is plotted for 1983 for all environments (Fig. 2). Moreover, the s.d.'s were also sometimes significantly correlated with the S^2_d values, even though the slope was not at all related to the S^2_d values.

General discussion

Limitations of an environmental index. Although regression analyses have been extensively used, they have limitations, some of which have been pointed out by Knight (1970), Witcombe & Whittington (1971), and Binswanger & Barah (1980).

One major disadvantage of the regression analysis is not found with mean-standard deviation analysis. This disadvantage is the invalidity of the RMS statistic as a measure of stability in certain circumstances. (Note that the deviations from the regressions were found to be the largest component of variation involving the genotype in the regression analysis). For example, as found by Pfeiffer & Braun (1985), a disease-resistant entry in a trial where all other entries are susceptible will deviate greatly from the performance of the other entries and hence from its regression line in environments where the disease is present. If the disease is present in only a few locations then this superiority will increase its S^2_d, and render it unstable according to the analysis, even though it is superior in stability. In contrast, in a mean-standard deviation analysis a resistant entry may tend to have a lower s.d.

The disadvantage in a regression analysis of describing the response of genotypes relative to a biologically determined environmental index (the mean of the trial entries) has been discussed by Witcombe & Whittington (1971). Essentially, regression techniques to characterise genotype responses to the environment are an oversimplification. Interactions are still occurring which are not identified because biological indices do not give information on the physical nature of the environment. Knight (1970) has also discussed the disadvantage of a biologically determined environmental index.

A choice-theoretic framework. The regression analysis cannot provide an answer to the questions as to how the superior stability of genotypes assessed by slope and S^2_d translates into real economic benefit for the farmer. In a specific location, how does a low S^2_d reduce the variation from year to year (the temporal variation), as compared to a low regression coefficient? How large a reduction in mean yield can be accepted to select an entry with a lower S^2_d or lower slope? How should the relative importance of slope and S^2_d be assessed?

Binswanger & Barah (1980) have recognized these difficulties and pointed out that regression analyses are not suitable for studies on risk measurement, instead they have used a mean-standard deviation analysis where improved stability (reduced variation across years), or improved adaptability (reduced variation across locations) can be traded off against reduced yield. The analysis of Binswanger & Barah (1980) differs fundamently from the regression approach; it does not attempt to predict how a genotype will behave across locations, nor does it attempt to separate location and year effects from G × E interactions. Instead, analyses of variance of each genotype are used. For the stability parameter location effects are removed but year effects remain, and for the adaptability parameter year effects are removed and location effects remain.

Use of single-year data to assess risk. A major disadvantage of Binswanger & Barah's analysis for its universal application is that the choice-theoretic framework relies on the stability parameter which requires across year, across location data. Such data is expensive to obtain, and in most plant breeding programmes the composition of the set of entries in a multilocational trial changes each year, as poor yielding entries are discarded and new entries are promoted to the trial. The data analysed in this paper are therefore typical – the yield data is for locations within years and not across years. The assumption has, therefore, been made that temporal and spatial variation are related, and that adaptability s.d.'s are related to stability s.d.'s.

Evenson et al. (1978) reported that for rice trials in India adaptability and stability were related. Barah et al. (1981) have also shown for sorghum in India that stability and adaptability were closely related. Moreover, a further examination of their data revealed that the variability s.d.'s within individual years correlated well in three out of four years with stability and adaptability s.d.'s (Table 4). Poor fits between these parameters in certain years are not unexpected, because estimates of s.d.'s are less precise in a single year than across years. (Estimates of slopes from the regression analysis are expected to behave in the same way with the precision of the estimates being higher across more than one year). When the variability s.d.'s were averaged over the years then these averages correlated more with stability than adaptability. These data indicate that single-year data are relevant to a choice-theoretic framework, but more data are required from other crops and other trials. The choice-theoretic framework is unaffected by the larger absolute values of the variability s.d.'s since the method is comparative. It is not the size of the variability s.d.'s but their relationship to stability s.d.'s which is important.

Whether the relationship found in the sorghum data is true for any set of trials is clearly dependent on how well the environmental variability between locations matches changes in the environment across years.

Conclusions. It is an inescapable conclusion that an analysis other than regression analysis is required to obtain an overall picture of how stability and mean yield are to be traded off. The mean-standard deviation analysis has shown for these, and other data, that in all years no entry was selected other than on the basis of its high mean yield. The results of the mean-standard deviation analysis were little changed relative to a regression analysis if the lowest- and highest-yielding environments were considered, so the comparison is not environment dependent. A lack of marked change between the environmental sets, and the invariable selection of the highest-yielding entry when a choice-theoretic framework is used, indicates that the breeder's procedure of selecting among the highest-yielding entries across environments is satisfactory.

What is even more remarkable, from the IPMAT data, is the similarity between the regression coefficients and the s.d.'s (and to a lesser extent, that between the S^2_d values and the s.d.'s). The s.d., a simple statistic, predicted well one of the major parameters of stability of a regression analysis. The s.d. is a simpler statistic that can be used in a choice-theoretic framework, and the mean-standard deviation analysis does not suffer from the disadvantages of a regression analysis which employs an environmental index. However, with respect to the mean-standard deviation analysis more data are required from more multilocational trials to determine the relationships between stability, adaptability, and estimates from single-year data.

A regression analysis requires many entries to be included in the analysis as the environmental index is more accurately assessed with a larger number of entries. No such restriction exists within the mean-standard deviation analysis. It can be carried out on only two or three entries. Thus, before a decision on release is made only the few highest-yielding entries need to be considered, and often these entries will have been retained in the trial for several years so that a true stability-relevant s.d. can be calculated.

The mean-standard deviation analysis is superior in some respects to the regression analysis, and should be used more extensively in the analysis of multilocational trials. Nevertheless, both methods complement each other, since the mean-standard deviation analysis provides a choice-theoretic framework whilst the regression analysis enables prediction of performance across environments.

Acknowledgements

The vital help of our cooperators who have grown the trials is gratefully acknowledged. The IPMAT reports of 1979 and 1980 were prepared by Dr. Anand Kumar who with Dr. K.N. Rai coordinated the trials in different years.

I am grateful to Mr. M.N.V.R. Rao for his assistance in analysing the data and to Dr. Tom Walker for his valuable suggestions in the analysis of the data.

References

Barah, B.S., H.P. Binswanger, B.S. Rana & N.G.P. Rao, 1981. The use of risk aversion in plant breeding; concept and application. Euphytica 30: 451–458.

Binswanger, H.P. & B.C. Barah, 1980. Yield risk, risk aversion, and genotype selection. Conceptual issues and approaches. ICRISAT Research Bulletin 3: 23.

Binswanger, H.P. & D.A. Sillers, 1983. Risk aversion and credit constraints in farmers' decision-making: a reinterpretation. J Development Studies 20: 5–21.

Eberhart, S.A. & W.A. Russell, 1966. Stability parameters for comparing varieties. Crop Sci 6: 36–40.

Evenson, R.E., J.C. O'Toole, R.W. Herdt, W.R. Coffman & H.E. Kauffman, 1978. Risk and uncertainty as factors in crop improvement research. IRRI paper series No. 15, Manila, Philippines.

Finlay, K.W. & G.N. Wilkinson, 1963. The analysis of adaptation in a plant breeding programme. Aust J Agric Sci 14: 742–754.

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

Perkins, Jean M. & J.L. Jinks, 1968. Environment and genotype environmental components of variability. III. Multiple lines and crosses. J Heredity 23: 339–356.

Pfeiffer, W.H. & H.J. Braun, 1985. Yield stability in bread wheat. Paper presented to the IFPRI/DSE Workshop on Sources of Increased Variability in Cereal Yields, Feldafing, FRG, November 1985.

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

Yates, F. & W.G. Cochran, 1938. The analysis of groups of experiments. J Agric Sci Camb 28: 556–580.