Published May, 1988

Statistical Analysis of a Yield Trial

Richard W. Zobel,* Madison J. Wright, and Hugh G. Gauch, Jr.

ABSTRACT

Yield trials frequently have both significant main effects and a significant genotype × environment (GE) interaction. Traditional statistical analyses are not always effective with this data structure: the usual analysis of variance (ANOVA), having a merely additive model, identifies the GE interaction as a source but does not analyze it; principal components analysis (PCA), on the other hand is a multiplicative model and hence contains no sources for additive genotype or environment main effects; and linear regression (LR) analysis is able to effectively analyze interaction terms only where the pattern fits a specific regression model. The consequence of fitting inappropriate statistical models to yield trial data is that the interaction may be declared nonsignificant, although a more appropriate analysis would find agronomically important and statistically significant patterns in the interaction. Therefore, agronomists and plant breeders may fail to perceive important interaction effects. This paper compares the above three traditional models with the additive main effects and multiplicative interaction (AMMI) Model, in an analysis of a soybean [*Glycine max* (L.) Merr.] yield trial. ANOVA fails to detect a significant interaction component, PCA fails to identify and separate the significant genotype and environment main effects, and LR accounts for only a small portion of the interaction sum of squares. On the other hand, AMMI analysis reveals a highly significant interaction component that has clear agronomic meaning. Since ANOVA, PCA, and LR are sub-cases of the more complete AMMI model, AMMI offers a more appropriate first statistical analysis of yield trials that may have a genotype × environment interaction. AMMI analysis can then be used to diagnose whether or not a specific sub-case provides a more appropriate analysis. AMMI has no specific experimental design requirements, except for a two-way data structure.

Additional Index Words: **Additive main effects and multiplicative interaction model, Analysis of variance, Biplot, Linear regression, *Glycine max* (L.) Merr., Principal components analysis.**

YIELD trials are conducted with many genotypes grown in a number of environments (year and site combinations), usually with replication. The total sum of squares (SS) for the yield data can be partitioned into three general sources: the genotype main effect, the environment main effect, and the genotype × environment (GE) interaction. By definition, main effects are additive, and interaction (residual from the additive model) nonadditive (Snedecor and Cochran, 1980). Commonly, all three sources are statistically significant and agronomically important (Kempton, 1984; Freeman, 1985).

The customary statistical analyses applied to yield trials, (i) analysis of variance (ANOVA), (ii) principal components analysis (PCA), and (iii) linear regression (LR), are often inadequate in effectively treating such a complex data structure (Gollob, 1968; Mandel, 1971; Bradu and Gabriel, 1978; Kempton, 1984). The following summarizes some of the more important problems: (i) The ordinary ANOVA is an additive model and therefore describes only the main effects effectively (Snedecor and Cochran, 1980). ANOVA can test the significance of the GE interaction, but this test may prove to be misleading. In any case, ANOVA provides no insight into the particular patterns of genotypes or environments that give rise to the interaction. (ii) Principal components analysis, a multiplicative model, has the opposite problem of not describing the additive main effects. Consequently, the interaction, which is by definition the residual from the additive model, is not even considered, much less analyzed effectively by this model. (iii) Linear regression models (e.g., Mandel, 1961; Finlay and Wilkinson, 1963) combine additive and multiplicative components and thus analyze main effects and the interaction. There are several deficiencies in the fitting procedure of the most commonly used model (Finlay and Wilkinson, 1963), since staged fitting of the interaction component is known not to give a least-squares fit (Gabriel, 1978). Additionally, the LR model in general confounds the interaction with the main effects (Wright, 1971), reducing its power for general significance testing.

An ANOVA test of the significance of the GE interaction may declare it nonsignificant when, in reality, the interaction is agronomically important (a more appropriate statistical model may both detect significance and describe interesting patterns in the interaction). This problem arises because the interaction contains a large number of degrees of freedom (df): given G, genotypes, and E, environments, the interaction contains $(G-1) \times (E-1)$ degrees of freedom. Even if, as is typically the case, the interaction contains 20 to 50% of the total SS (which may even exceed the SS for the genotype main effect), the interaction mean square (MS) may nearly equal the error MS and hence be declared insignificant by an F-test. (See Bancroft [1964] for appropriate tests of significance to determine model complexity.)

This paper compares the above three models with the additive main effects and multiplicative interaction (AMMI) model, also called FANOVA (Gollob, 1968), which incorporates both additive and multiplicative components into an integrated, powerful, least-squares analysis (Gollob, 1968; Gabriel, 1971, 1978; Mandel, 1971; Bradu and Gabriel, 1978; Kempton, 1984; Gauch, 1985; Freeman, 1985). Gollob (1968), Mandel (1971), and Gabriel (1978) conclude that ANOVA, PCA, and various regression analyses are in reality sub-cases of the AMMI model. These sub-cases test specific hypotheses about underlying relationships; e.g., only additive effects (ANOVA), only multiplicative effects (PCA), or multiplicative relationships between the genotype yields and the envi-

R.W. Zobel, USDA-ARS-USPSNL, Dep. of Agronomy and Dep. of Plant Breeding, Cornell Univ., Ithaca, NY 14853-1901; M.J. Wright and H.G. Gauch, Jr., Dep. of Agronomy, Cornell Univ., Ithaca, NY 14853. Contribution from the USDA-ARS, U.S. Plant Soil and Nutrition Lab., Tower Rd., Ithaca, NY 14853. Received 7 July 1986. *Corresponding author.

Published in Agron. J. 80:388-393 (1988).

ronmental means or vice versa (LR). AMMI is ordinarily the model of first choice when main effects and interaction are both important (Mandel, 1971), which is the most common case with yield trials. If, for example, only main effects (additive structure) are present in the data, then the AMMI model can be reduced to an ANOVA model; whereas, if only nonadditive structure is present, then a PCA model is indicated. AMMI results can be readily used to diagnose these and other sub-cases (Bradu and Gabriel, 1978).

The model comparisons presented here use the MATMODEL program (Gauch, 1986) to analyze yield data from a soybean trial. The results are compared to highlight the yield trial features discernable in the four approaches. This paper attempts to demonstrate that Gollob's conclusions (1968), based on psychological data; Mandel's conclusions (1971), based on an analysis of physical data; and Gabriel's conclusions (Gabriel, 1971; Bradu and Gabriel, 1978), based on physical and agricultural data, are valid for yield trial data. It is also shown that the AMMI model provides plant breeders and other plant scientists using two-way data sets, with a powerful statistical tool.

MATERIALS AND METHODS

The data used in this study were taken from Cornell University, Department of Agronomy mimeos (Wright, et al. 1978a, 1978b, 1980, 1981, 1982, 1983a, 1983b, 1985). Specific data on locations, soil types, herbicide treatments, and cultural practices are available in these mimeos and are not reported here. Test plots consisted of randomized plots with two, three, or four replications (usually four), four rows of a 7.6-m length with 0.36-m center-to-center spacings, and a plant density of approximately 37 plants per square meter. The central 4.9 m of the middle two rows was harvested and seed yield expressed in kilograms per hectare at a moisture content of 130 g kg^{-1} of seed.

For geographical reference, Chazy (C) is in the northeast in the Lake Champlain drainage; Canton (N) is in the northeast in the Saint Lawrence River drainage; Lockport (L) is in the west near Niagara River; Geneseo or Mt. Morris (G) is in the central west in the Genesee River Valley; Romulus (R) is in the center in the western drainage into Cayuga Lake; Aurora (A) is in the center in the eastern drainage into Cayuga Lake; Ithaca (I) is in the south central highlands above Cayuga Lake; Valatie (V) is in the east in the Hudson River Valley near Albany; and Riverhead (D) is on the eastern end of Long Island.

The trials contained over 70 cultivars, but not all were tested at all sites in all years. Seven of these cultivars were tested in 35 site-years (identified by the location initial and the last two digits of the year, such as A82 for Aurora in 1982) and are therefore the only ones used here. These cultivars (with maturity group in parentheses) were: Evans (0), Wilkin (0), Chippewa 64 (I), Hodgson (I) (Hodgson 78 in 1981 and later), Corsoy (II), SRF 200 (II), and Wells (II) (Wells II in 1984).

Most trials had four replicates, but some only two or three; of 980 possible plots, 912 were actually harvested. This resulted in 244 df for the model, and 667 df for error. Standard equations for unequal group sizes using unweighted means were used to calculate the error and analysis tables.

Because the experimental design was unbalanced, in principle a generalized linear model (GLM) should have been used. In our judgment, however, the required computational effort of GLM was unreasonable, especially since the unbalanced design caused some irremediable confounding of the main effects and interaction. Furthermore, the results should be very close to those obtained by the present simple calculations. The alternative strategy of discarding nearly half the observations to achieve a balanced design was unattractive. Preliminary analyses with this data set (not shown) demonstrated to us that the slight approximation involved in the adopted procedure caused only small problems relative to other inherent limitations such as sampling errors, assumptions of normality and independence required to apply F-tests, and so on. Practicality appears here to be more relevant than elegance, particularly since most model sources were either manifestly significant or manifestly not significant. Statistical analysis was performed by MATMODEL (Gauch, 1986). MATMODEL performs several analyses: ANOVA, PCA, LR (Finlay and Wilkinson, 1963), concurrence or joint regression (Tukey, 1949), and AMMI (Gollob, 1968; Mandel, 1971; Gabriel, 1978; Gauch, 1985). The appropriate statistical (structural) fixed-effect models are presented below:

The ANOVA model is

$$Y_{ge} = \mu + \alpha_g + \beta_e + \theta_{ge},$$

the PCA model is

$$Y_{ge} = \mu + \sum_{n=1}^{N} \lambda_n \zeta_{gn} \eta_{en} + \theta_{ge},$$

the LR model is

$$Y_{ge} = \mu + \alpha_g + \beta_e + K\alpha_g\beta_e + \gamma_g \beta_e + \alpha_g \delta_e + \theta_{ge},$$

and the AMMI model is

$$Y_{ge} = \mu + \alpha_g + \beta_e + \sum_{n=1}^{N} \lambda_n \zeta_{gn} \eta_{en} + \theta_{ge},$$

where Y_{ge} is the yield of genotype, g, in environment, e; μ is the grand mean; α_g is the genotype mean deviation; β_e is the environment mean deviation; λ_n is the eigenvalue of the PCA axis, n; ζ_{gn} and η_{en} are the genotype and environment PCA scores for the PCA axis, n; N is the number of PCA axes retained in the model; K is the Tukey concurrence constant; γ_e is the environment slope on the genotype means; δ_g is the genotype slope on the environment means; and θ_{ge} is the residual.

In addition, if the experiment is replicated, one may add to the above models the error term ϵ_{ger} for the deviation between the Y_{ge} treatment mean and the single observation for replicate r. Throughout, we use the error term for significance testing, in agreement with the suggestions of Bancroft (1964).

With the fore knowledge that the cultivars used in this study can be divided into three groups based on maturity grouping, it would be possible to further partition the ANOVA model to provide a comparison of the maturity grouping effects. Sites and years can also be partitioned out, yielding a complex but informative ANOVA table. This type of procedure is equally valid in the LR and AMMI models and has been used very profitably to study yield variability of a wide range of crops in the United Kingdom (Talbot, 1984). To provide a concise and clear example of the comparison of models, we have assumed no fore knowledge of the maturity grouping relationship of the cultivars in question and have retained the site-years as a single term.

A convenient scaling for the PCA scores has been selected: genotype and environment PCA scores are expressed as a unit vector times the square root of the eigenvalue; i.e., $\lambda \zeta_g \eta_e = (\lambda^{0.5} \zeta_g)(\lambda^{0.5} \eta_e)$ [genotype score = $\lambda^{0.5} \zeta_g$; environment score = $\lambda^{0.5} \eta_e$] (Gabriel, 1971; Gabriel, 1982, 1983, personal

communication; Bradu and Gabriel, 1978). Multiplication of a genotype PCA score by an environment score then gives the estimated interaction directly.

The least-squares fit to the AMMI model for balanced data (equal replication) is obtained in two steps: (i) The main effects in the additive part of the model (grand mean, genotype means, and environment means) are analyzed by the ordinary ANOVA. This leaves a nonadditive residual (namely the interaction). (ii) The interaction in the multiplicative part of the model is then analyzed by PCA. If all PCA axes were retained, the resulting full model would have as many degrees of freedom as the data and would consequently fit the data perfectly. The usual intent is, however, to summarize much of the interaction in just a few PCA axes (with $N = 1$ to 3), resulting in a reduced AMMI model that leaves a residual. Because they allow the use of F-tests to determine the significance of the PCA MS, degrees of freedom are calculated by the simple method of Gollob (1968):

$$df = G + E - 1 - 2n.$$

RESULTS AND DISCUSSION

When the data's structure agrees moderately well with the model, the analysis achieves three goals: (i) parsimony, because the model contains relatively few of the total degrees of freedom, (ii) effectiveness, because the model contains most of the total SS, leaving a residual with most of the degrees of freedom but few SS, and (iii) meaningfulness, in that the model provides agronomically meaningful insights into the data structure. We use these three criteria to compare the four models.

Table 1 presents the mean yield data from the Department of Agronomy mimeos, for the seven cultivars considered here, grown in 35 environments (site-years), plus the genotype and environment means and the first PCA axis. Yields ranged from 578 to 4620 kg ha^{-1}, with an average of 2568. The grand mean has been removed in the following statistical analyses, so the trials contain $(7 \times 35) - 1 = 244$ df.

Table 2 presents the ANOVA analysis. Note that this GE interaction exhibits the features discussed in the introduction: the interaction SS is large, nearly three times as large as the genotype SS, but after division by its enormous 204 df, the resulting small interaction MS is judged nonsignificant by an F-test. Even if this interaction had been judged significant, no detailed partitioning would normally have been provided. In short, this statistical treatment of the potentially rich and interesting structure contained in the 204 df of the interaction is superficial. ANOVA there-

Table 1. Yield for seven soybean genotypes grown in 35 environments (site-years) and values for the AMMI model, namely the means and the first PCA from analysis of the interaction.

	Genotype								
Environment	EVAN†	WILK	CHIP	HODG	S200	CORS	WELL	Mean	First PCA
	kg ha^{-1}								
A77	2757	2502	2361	2771	2878	3141	2777	2741	3.4
C77	2946	2771	1755	2192	1089	1836	1217	1972	25.9
V77	1567	1103	2266	2468	2730	2569	2616	2188	−21.7
V78	1735	1493	1506	2172	2011	2145	1553	1802	−3.7
A79	3127	2623	2488	3201	3430	2878	2791	2934	−2.3
C79	2777	2562	1728	1944	1849	1486	1264	1944	19.5
G79	2986	2367	2340	3154	2623	3040	2455	2709	0.1
R79	1843	1110	1816	2495	2105	1769	2058	1885	−9.4
V79	1083	578	1278	1500	1964	1661	1715	1397	−15.6
A80	2919	2784	2582	3208	2703	3100	2219	2788	4.3
C80	2596	3248	2280	2710	2172	2260	1432	2385	18.3
G80	3901	3194	3376	4096	3887	4250	3517	3746	−4.9
L80	3706	3820	2993	3732	3739	3161	3215	3481	7.7
D80	1937	1580	1580	2374	1997	2609	1890	1995	−5.3
R80	2199	1870	2199	2966	2461	2327	2186	2315	−4.9
V80	2334	2018	1802	1964	1601	2165	1782	1952	6.7
A81	3033	2609	2636	3013	2831	3611	2959	2956	−3.5
C81	3053	3053	1849	2522	1654	2186	1708	2289	21.8
G81	3322	2892	3208	3840	3383	4028	3235	3415	−6.3
L81	2972	2710	2636	2972	2569	2818	2777	2779	2.8
D81	2529	1997	2582	3268	2112	2529	2320	2477	−1.5
R81	2038	1385	2347	2798	2616	2966	2757	2415	−17.6
V81	3026	3127	2387	2367	2361	2461	2260	2570	13.7
A82	2186	1870	1883	2441	2441	2562	2239	2232	−5.3
L82	2663	1957	2535	2798	3241	3147	2737	2725	−11.8
G82	3652	3295	2724	3712	3901	3921	3322	3504	−2.2
V82	2273	1910	1123	1755	1184	1399	1345	1570	13.9
A83	2582	2125	2018	2313	2192	2058	1970	2180	4.9
I83	1278	1029	1688	1701	2105	1964	1856	1660	−13.0
G83	4499	4015	3329	4620	3564	4062	3625	3959	9.3
A84	3161	2717	3188	3860	3376	3423	3544	3324	−8.5
N84	2993	2603	2630	3221	3369	3201	2703	2960	−4.1
C84	3181	2690	2448	3060	2576	2784	2670	2773	4.8
I84	1950	1701	2078	2260	2145	2246	1917	2042	−4.0
G84	4015	3329	3329	3961	4277	4015	3692	3803	−4.9
Mean	2709	2361	2313	2841	2604	2737	2409	2568	
	First PCA								
	29.2	44.0	−5.2	−3.4	−24.5	−16.7	−23.5		

† Genotype and environment codes are given in Materials and Methods.

Table 2. Analysis of variance for soybean yields with seven genotypes in 35 environments.

Source	df	Sum of squares	Mean square	F
		× 10^{-3}		
Trials	244	139 991	574	5.19***
Genotype	6	8 992	1 499	13.55***
Environment	34	105 558	3 105	28.07***
GE Interaction	204	25 442	125	1.13 NS
Error	667	73 775	111	
Total	911	213 766	235	

*** = P < 0.001.

fore meets the criteria of only goal (i) because the interaction is treated as a simple residual.

Table 3 shows the three PCA axes declared significant by an *F*-test. The fourth axis, with 35 df, had an MS of only 65 × 10^3 and a nonsignificant *F* of 0.59; thus, it and the higher axes have been combined in the residual.

PCA was, from a purely statistical viewpoint, almost equal to ANOVA in summarizing the data structure. The first PCA axis with 41 df contained 75.9% of the SS, whereas the ANOVA main effects with 40 df contained 81.8%. However, in yield trials the genotype means are usually of primary interest, so practical or heuristic considerations disfavor a model that lacks these means. Furthermore, the interaction also is not specifically identified as a source. The PCA analysis was statistically efficient, but judged from a more conventional perspective, these PCA scores confound information on genotype means, environment means, and interaction in a rather undesirable manner. PCA therefore satisfies goal (ii) but is inadequate in meeting goal (i).

Table 4 shows an amended LR analysis. The original analysis (Finlay and Wilkinson, 1963) partitions the interaction into only two sources: genotype regressions and residual (containing 199 df with an MS of 127 × 10^3 and a nonsignificant *F* of 1.15). Here, an amended analysis also identifies environment regressions as a source, and when both sets of regressions are included the joint regression with 1 df must also be tested (Tukey, 1949; Wright, 1971; Gauch, 1985).

An LR analysis sometimes effectively explains much of the interaction SS, but for this particular data set even the amended analysis captures only 7.9% of the interaction SS in 19% of the degrees of freedom. The joint regression is statistically significant ($P < 0.05$) but explains only 1.9% of the interaction SS. The residual MS is significant (the residual contains 91.1% of the interaction SS), whereas the interaction itself is not. Regression *enriched* the residual MS rather than reduced it like an appropriate model would. The LR model as a whole accounts for 83% of the variance while leaving 17% in the residual. With this data set, little of the interaction variance can be explained by the hypotheses tested by the linear regression models used here. The LR model satisfies goal (i) but not goal (ii).

Table 3. Principal components analysis soybean yields with sequential sums of squares for the first three PCA axes.

Source	df	Sum of squares	Mean square	F
		× 10^{-3}		
Trials	244	139 991	574	5.19***
PCA 1	41	106 234	2 591	23.43***
PCA 2	39	18 368	471	4.26***
PCA 3	37	9 757	264	2.38***
Residual	127	5 632	44	0.40 NS
Error	667	73 775	111	
Total	911	213 766	235	

*** = P < 0.001.

The AMMI analysis is given in Table 5. The first PCA axis of the interaction captured 71.0% of the interaction SS in 19% of the interaction degrees of freedom. The MS for the first PCA axis was highly significant. The second PCA axis MS was nonsignificant (with 37 df and an SS of 2449 × 10^3, giving an MS of 66 × 10^3 and an *F* of only 0.60) and therefore was pooled with the remaining axes in the residual. Partitioning of the interaction SS by AMMI is quite effective, the MS for the first PCA axis being over ten times the MS for the residual. The complete AMMI model contained 94.7% of the SS, and the residual 5.3%. These values indicate that AMMI summarized these data very effectively. The first PCA axis accounted for the same number of degrees of freedom as the regression model while capturing ten times the interaction variance. The AMMI model satisfies both goals (i) and (ii). As discussed in the introduction, LR models are specific sub-cases of the AMMI model. The AMMI model clearly demonstrates the existence of a significant genotype × environment interaction, which is partitioned into the first PCA axis. These results and others (Gauch, 1985) lead us to concur with the suggestion by Gollob (1968), Mandel (1971), Bradu and Gabriel (1978), and Freeman (1985) that AMMI should be the first test of statistical pattern, with further tests by more restrictive models if appropriate. Bradu and Gabriel (1978) provide a method to diag-

Table 4. Finlay-Wilkinson regression analysis for soybean yields.

Source	df	Sum of squares	Mean square	F
		× 10^{-3}		
Trials	244	139 991	574	5.19***
Genotype	6	8 992	1 499	13.55***
Environment	34	105 558	3 105	28.07***
GE Interaction	204	25 442	125	1.13 NS
Joint Regr.	1	487	487	4.40*
Gen. Regrs.	5	133	27	0.24 NS
Env. Regrs.	33	1 394	42	0.38 NS
Residual	165	23 427	142	1.28*
Error	667	73 775	111	
Total	911	213 766	235	

*,*** = P < 0.05 and 0.001, respectively.

Table 5. AMMI analysis of variance for soybean yields.

Source	df	Sum of squares	Mean square	F
		× 10^{-3}		
Trials	244	139 991	574	5.19***
Genotype	6	8 992	1 499	13.55***
Environment	34	105 558	3 105	28.07***
GE Interaction	204	25 442	125	1.13 NS
PCA 1	39	18 075	463	4.19***
Residual	165	7 367	45	0.41 NS
Error	667	73 775	111	
Total	911	213 766	235	

*** = P < 0.001.

nose the appropriate model to use after testing with AMMI.

Agronomic Considerations

For reasons discussed earlier, ANOVA and PCA provide no additional insight beyond that directly available from their tables of analysis and, therefore, in this case do not satisfy goal (iii). LR can provide a family of regression lines, but the data show only a weak concurrence in this data set. When regressions of genotype means on environments or vice versa have significant MS, regression lines can demonstrate similarities between cultivars in their interaction with the environments, and similarities between environments in their effect on cultivars. In this data set, however, LR can provide very little further insight into the data, and therefore does not satisfy goal (iii).

Figure 1 presents a biplot (Gabriel, 1971; Bradu and Gabriel, 1978; Kempton, 1984) of the AMMI results. The abscissa shows the main effects (genotype and environment means), and the ordinate shows the first PCA axis. Figure 1 therefore accounts for 94.7% of the SS. It is reasonable to consider most if not all of the residual 5.3% to be merely noise without agronomic interest (examination of the residuals did not reveal any large values).

In Fig. 1, for any genotype-environment combination, the main effect equals the genotype mean plus the environment mean minus the grand mean, and the interaction is the genotype PCA score times the environment PCA score. For example, Wilkin grown in C77 has a main effect of 2361 + 1972 − 2567 = 1766 kg ha^{-1}, and an interaction effect of 44 × 26 = 1144 kg ha^{-1}. Therefore the AMMI model gives a yield estimation of 2910 kg ha^{-1} for Wilkin at Chazy in 1977. This fits well the observed yield of 2771 kg ha^{-1}, whereas by contrast the additive ANOVA yield estimation of 1766 kg ha^{-1} gives a poor fit. The ANOVA residual MS of 125 versus the AMMI residual MS of 45 demonstrates a greater accuracy for the latter model.

In a biplot presentation like Fig. 1, when a genotype and an environment have the same sign on the PCA axis, their interaction is positive; if different, their interaction is negative. If a genotype or an environment has a PCA score of nearly 0 it has small interaction effects (and hence can be fitted well by an additive model). Figure 1 displays at a glance both main effects and interaction. For example, V82 and V81 differ in main effects but not interaction (PCA score), whereas V82 differs from V79 only in interaction, since the PCA scores differ but mean yields do not. It can easily be seen that V82 differs from V77 in both main effect and interaction. The Aurora site shows much variability from year to year in main effects, but little variation in interaction. Valatie is most variable in interaction. Hence, relative rankings (but not absolute yields) of the genotypes are fairly stable in Aurora, but rankings are likely to be quite variable from year to year in Valatie, making it much more difficult to produce variety recommendations for Valatie.

We suggest that it is daylength/maturity-group relationships that are responsible for the highly significant interaction partitioned out in the first PCA axis. Wilkin and Evans at one extreme on the PCA axis are both in maturity group 0, whereas SRF 200, Wells, and Corsoy at the opposite extreme are in group II (and the group I cultivars Chippewa 64 and Hodgson are intermediate). Position along this gradient appears relatively exact since Wilkin is a slightly earlier group 0 cultivar than Evans, and Corsoy is a somewhat early group II. Concomitant with this interpretation, the northern site at Chazy favors the early group 0 cultivars (both site and cultivar have positive interaction terms), and environments with points near the bottom of Fig. 1 favor group II cultivars (both site and cultivar have negative interaction scores). The AMMI model, with a biplot display of the model, helps to visualize the overall patterns of this data set as well as specific GE interactions. It therefore satisfies all three goals: parsimony, effectiveness, and insight.

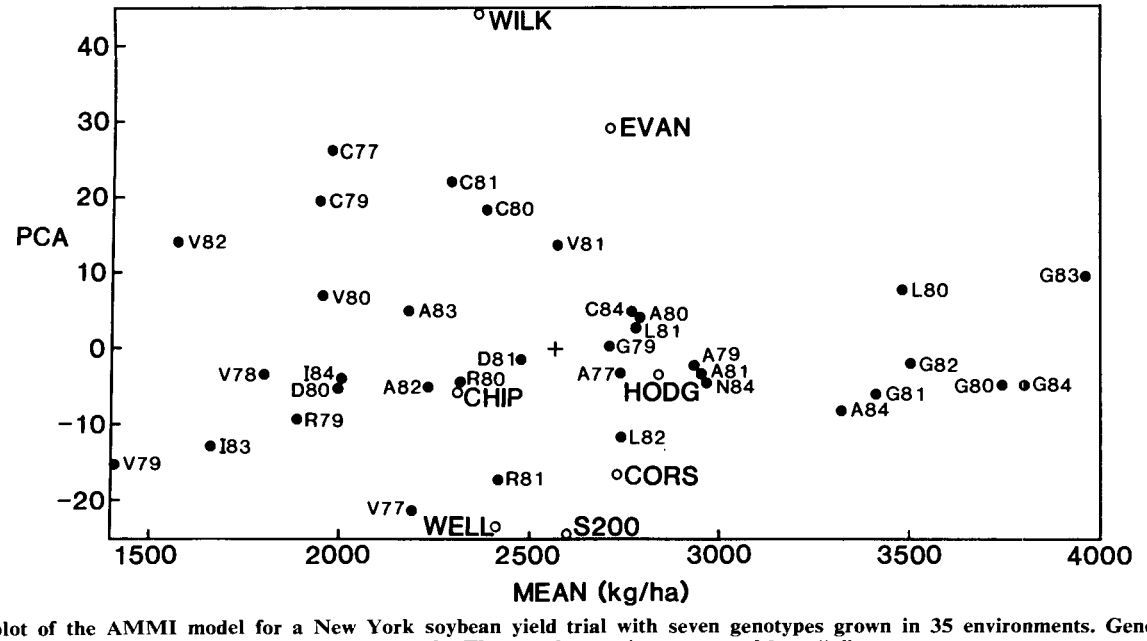

Fig. 1. Biplot of the AMMI model for a New York soybean yield trial with seven genotypes grown in 35 environments. Genotype and environment codes are given in Materials and Methods. The grand mean is represented by a "+".

The above discussion points out the additional information available through appropriate use of a multiplicative analysis in conjunction with ANOVA (AMMI) in a first-statistical test of a data set. With fore knowledge, a model encompassing further partitioning of the data prior to analysis would give even more insight into the data structure. Because AMMI and LR incorporate ANOVA, in their models, partitioning procedures appropriate for ANOVA are also valid for these other two models, though not for PCA. We suggest that this further partitioning of the data would still demonstrate a superiority for the AMMI procedure since it can further analyze the interaction data beyond that possible with ANOVA.

Clearly the four models compared had very different efficiencies in partitioning the variance into agronomically important components. Mandel (1969) describes the first three, ANOVA, PCA, and LR, as subcases of the AMMI model and concludes that they can be diagnosed from the AMMI model (a procedure for doing this using the biplot display is provided by Bradu and Gabriel [1978]). The conclusions presented by Gollob (1968), Mandel (1969, 1971), Gabriel (1971), and Bradu and Gabriel (1978) are confirmed here with our yield data set. The AMMI model should be the first model attempted, with others explored for possible improvement in efficiency, or testing of specific hypotheses about the underlying interaction patterns. The primary restriction to use of the AMMI model is its current requirement for two-way data tables (Bradu and Gabriel, 1978). Experimental design (randomized complete block, completely randomized block, Latin square, etc.) is not a significant consideration in the selection of an appropriate model(s).

REFERENCES

Bancroft, T.A. 1964. Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. Biometrics 20:427–442.

Bradu, D., and K.R. Gabriel. 1978. The biplot as a diagnostic tool for models of two-way tables. Technometrics 20:47–68.

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

Freeman, G.H. 1985. The analysis and interpretation of interaction. J. Appl. Stat. 12:3–10.

Gabriel, K.R. 1971. The biplot—graphical display of matrices with applications to principal component analysis. Biometrika 58:453–467.

———. 1978. Least squares approximation of matrices by additive and multiplicative models. J. R. Stat. Soc. Ser. B 40:186–196.

Gauch, H.G. 1985. Integrating additive and multiplicative models for analysis of yield trials with assessment of predictive success. Cornell Univ. Dep. of Agronomy Mimeo 85–7.

———. 1986. MATMODEL: A FORTRAN-77 program for AMMI analysis. Microcomputer Power, Ithaca, NY.

Gollob, H.F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33:73–115.

Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. J. Agric. Sci. 103:123–135.

Mandel, J. 1961. Non-additivity in two-way analysis of variance. Am. Stat. Assoc. J. 56:878–888.

———. 1969. The partitioning of interaction in analysis of variance. J. Res. Natl. Bur. Stand. Sect. B. 73B:309–328.

———. 1971. A new analysis of variance model for non-additive data. Technometrics 13:1–18.

Snedecor, G.W., and W.B. Cochran. 1980. Statistical methods. 7th ed. Iowa State Univ., Ames.

Talbot, M. 1984. Yield variability of crop varieties in the U.K. J. Agric. Sci. 102:315–321.

Tukey, J.W. 1949. One degree of freedom for non-additivity. Biometrics 5:232–242.

Wright, A.J. 1971. The analysis and prediction of some two factor interactions in grass breeding. J. Agric. Sci. 76:301–306.

Wright, M., R. Zobel, D. Ellis, and R. Lychalk 1978a. Soybean variety yield tests for 1977. Cornell Univ. Dep. of Agronomy Mimeo 78–11.

———, ———, ———, and ———. 1978b. Soybean variety yield tests for 1978. Cornell Univ. Dep. of Agronomy Mimeo 78–32.

———, ———, ———, and ———. 1980. Soybean variety yield tests for 1979. Cornell Univ. Dep. of Agronomy Mimeo 80–17.

———, ———, ———, and ———. 1981. Soybean variety yield tests for 1980. Cornell Univ. Dep. of Agronomy Mimeo 81–23.

———, ———, ———, and ———. 1982. Soybean variety yield tests for 1981. Cornell Univ. Dep. of Agronomy Mimeo 82–9.

———, ———, ———, and ———. 1983a. Soybean variety yield tests for 1982. Cornell Univ. Dep. of Agronomy Mimeo 83–2.

———, ———, ———, and ———. 1983b. Soybean variety yield tests for 1983. Cornell Univ. Dep. of Agronomy Mimeo 83–27.

———, ———, ———, and ———. 1985. Soybean variety yield tests for 1984. Cornell Univ. Dep. of Agronomy Mimeo 85–1.