Methods for Determining Natural Areas for Oat Varietal Recommendations[1]

T. W. Horner and K. J. Frey[2]

SYNOPSIS

Substantial reduction of the variety-location interaction affecting genotypic progress resulted from proper subdivision of Iowa into subregions, following a procedure described herein. For example, optimum subdivision into four parts resulted in a reduction of 30%. Implications with respect to a breeding program were briefly discussed. The data involved 9 locations, 5 years and oat varieties previously selected for wide adaptability.

IN A plant breeding program for the small grains, the strains most likely to be retained are those with a satisfactory regional or area mean performance. Not very often is a strain selected because it produces well in a particular niche. Similarly, small grain varieties are most often recommended on a regional or large area basis. This is in spite of the fact that the interaction of varieties × test locations within the region may be, and in fact usually is, of considerable magnitude. These considerations raise the questions: *(a)* will subdivision of the region into more homogeneous subregions, from the standpoint of climate and soils, be useful in recommending crop varieties, and *(b)* the boundaries and number of subregions for the most information.

The answers depend in part upon first, the relative magnitudes of the variety × location interactions for the adjacent test locations as compared to those more distant, and second, the decrease in variety × location interaction for a particular subdivision compared to the interaction for the region as a whole. This study is concerned with the latter two objectives using data from oat yield plots grown in Iowa during the years 1950–1955, inclusive. Iowa is considered a region and any divisions of the state as subregions.

This problem is different somewhat from that considered by Sprague and Federer (3). They determined the optimum number of years, places and replications for maximizing genotypic gain. Apart from cost considerations, their conclusion was to increase the number of locations at the expense of the number of replications at each location. They assumed that the magnitudes of the variety component and the variety × location component were not affected by the number of locations. This is a reasonable assumption in some instances but not in others.

MATERIALS AND METHODS

The yield data used in this study were collected from oat variety tests, each consisting of 3 replications and 18 varieties, conducted at each of the 9 locations in Iowa indicated in figure 1 for the 5 years, 1950, 1952, 1953, 1954 and 1955. Each experiment was treated as a randomized block design. Within a year the varieties were the same from one location to another, but varieties differed from year to year. Only 5 varieties were included throughout the 5 years.

The varieties included in the oat yield tests are released named strains known to be reasonably well adapted to Iowa and new strains developed and released recently by experiment stations in the north central states. Some new varieties are included each year.

Since the varieties in the yield tests changed from year to year it was not possible to obtain a combined analysis of all of the data directly. For each year a variety × location interaction mean square was obtained from the combined analysis of all locations. This interaction was divided into two portions, that between and that within subregions for different subdivision patterns of the state. Pattern here refers to the way the 9 test locations are combined, e.g., locations 1, 2, and 3 would be combined, 4, 5, and 6 would be another combination, and 7, 8, and 9 would be a third.

By dividing the state into 2 sections it was possible to obtain 16 different patterns, for 3 sections 12 patterns, for 4 sections 47 patterns, and for 5 sections 46 patterns. Only the seemingly logical patterns were used. The variety × location mean squares within subregions for each number of divisions were arranged into 2-way tables according to years and patterns and analyzed by sources: patterns, years and error (patterns × years).

To test whether the pattern with the lowest variety × location mean square was significantly smaller than that for the state as a whole the statistic: $\dfrac{x_L - x_s}{\sqrt{(1/5)\ \text{error M.S.}}}$ was computed where x_L was mean square of the pattern with the lowest mean square within sections (pooled over sections and averaged over 5 years), x_s was the mean square for the whole state. Since there were many degrees of freedom for the pooled error this statistic was approximately distributed as a similar quantity with the true variance in the denominator. Significance levels for the latter were obtained using

$$1 - P(x_n) = 1 - \left\{ (1/\sqrt{2\pi}) \int_{-\infty}^{x_n} e^{-(1/2)u^2}\, du \right\}^n$$

where $P(x_n)$ was the probability integral of the extreme (x_n) in a set of n items (2, 4).

EXPERIMENTAL RESULTS

The variety × location mean squares and tests of significance for the patterns with the lowest mean squares

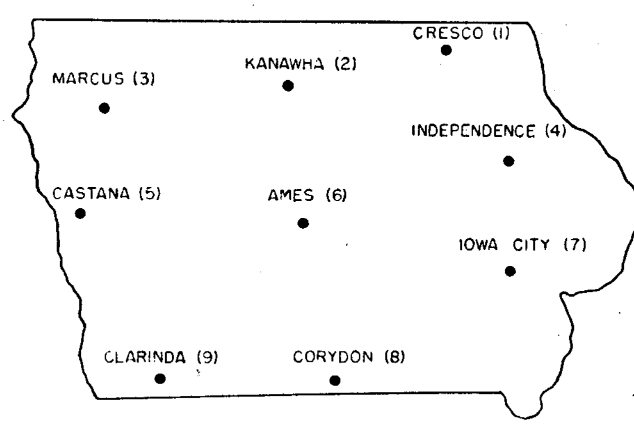

FIG. 1.—Location of oat variety trials in Iowa from 1950 to 1955 inclusive.

[1] Journal Paper No. J-3078 of the Iowa Agricultural Experiment Station, Ames, Iowa. Projects 1176 and 1285. Received December 24, 1956. In cooperation with the Field Crops Research Branch. A.R.S., U.S.D.A.

[2] Assistant Professor of Statistics and Professor of Agronomy, respectively, Iowa State College.

within areas for different numbers of subregions are presented in table 1. As was expected the F tests of the pattern-by-year 2-way tables showed nonsignificance irrespective of the number of subregions. By using Tippett's test (4), the patterns with smallest mean square for each number of subregions were found to be significantly lower than that (4435) for the state as a whole. The optimum pattern for each number of subregions can be determined from figure 1 by combining the locations according to lines 4 through 8 in table 1.

The variety × location variance components for the pattern with the lowest mean square for each number of subregions are graphed in figure 2. The estimated components seemed to follow a straight line. The average increment of reduction was 100 over the range from 1 to 5 subregions. The estimated variety × location component for the best pattern when dividing the state into 5 subregions was only 60% of that for the state as a whole.

In light of these results an attempt was made to determine whether certain test locations contributed more than others to the variety × location mean square. Variety × location mean squares were calculated for all 36 possible location pair combinations for each year. This resulted in a 36 × 5 two-way table which was analyzed into the sources: location pairs, years, and error (location pairs × years). This analysis based on mean squares gives the same F value for location pairs as the analysis of the V × L component estimates when the pooled error term over locations is used in the computation of the latter. There were highly significant differences among the average mean squares for location pairs. The ranks and averages for several of the location pairs are presented in table 2. Location pairs, 3,6; 6,9; 1,6 and 6,7 were respectively greater at the 5% level (1) than the pair ranking 3rd, 11th, 18th or 34th. Among location pairs which did not involve location Ames *(6)* there were no significant differences except Cresco *(1)*–Clarinda *(9)* was higher than the three bottom pairs.

When the location pairs sum of squares was partitioned into general and specific effects mean squares it was found that the general mean square for effects was highly significant. Estimates of the general location effects upon the variety × location variance components are presented in table 3.

Obviously, the contribution of location 6, the Ames test, is overwhelmingly larger than any other location. Corydon and Castana contributed the least to the variety × location mean squares. This was surprising because all test locations were contiguous to location 6, the center test area, while location 8 is separated from locations 1, 2 and 3 by intervening test areas. The reasons for this situation are under investigation.

Table 1.—Variety × location mean squares and significance tests of best patterns for differing number of subregions.

	Number of subregions				
	1	2	3	4	5
No. of patterns	1*	16	12	47	46
F test of pattern difference		NS	NS	NS	NS
S. E. of pattern mean		117	171	233	296
Extreme pattern [one with smallest V × L mean square]		(1235)** (46789)	(123) (456) (789)	(14) (26) (35) (789)	(12) (35) (47) (6) (89)
Mean square	4435	4104	3813	3527	3234
Significance level		<.05	<.005	<.005	<.005

*No subdivision of state.
**Test locations in subregions.

DISCUSSION

An attempt was made to determine whether the state of Iowa could be divided into logical subareas for making oat varietal recommendations. The magnitude of the variety × location mean squares among tests within subareas was used as a measure of homogeneity within the subareas. The reduction in the variance component (of the pattern with smallest mean square) due to variety × location interaction was completely linear with an increasing number of subareas. This leads to the conclusion that each of the yield trial locations should be considered as a separate area for variety recommendations. However, this would be rather impractical. In general from patterns which produced the smallest mean squares southern Iowa should be considered as one region. Locations 7, 8, and 9 were together in each of the lowest mean square patterns when the state was divided into 2, 3, or 4 subareas. These test locations tend to have similar climatic and edaphic conditions. It would appear further that the northern two-thirds of the state should be divided on a north-south basis as when the state is divided into 4 subregions.

There might be some basis for dividing the state into 5 subareas with location 6 by itself. The reduction expected in the variety × location mean squares within subregions was an extra 10% and location 6 seems to be inconsistent in its contribution to this component. However, until more is known about the causes of the oversized contribution from location 6, the separation is not logical. Environmental factors which would separate location 6 from location 2 are not apparent. Perhaps the yield tests at location 6 are not as representative of the subarea as other locations are of their respective subareas.

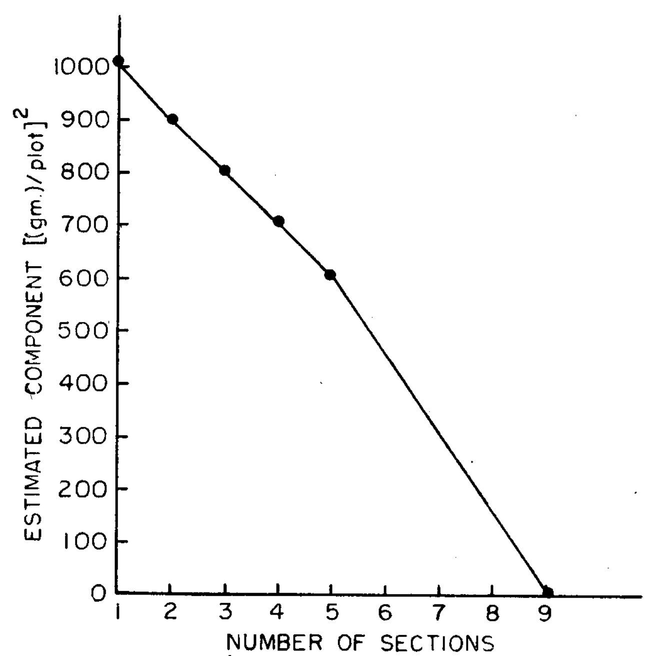

FIG. 2.—Estimated variety × location components for optimum patterns graphed against number of sections into which the state was divided.

With 4 subregions for oat variety recommendations the variety × location variance component was reduced 30% for the best pattern as compared to the state as a whole. This reduction may be even more important than is indicated by the figure itself since the variation which is no longer in the variety × location variance would presumably contribute to variety variation within the subregions. Hence, the potential gain from selection of the top producing variety in a subregion and consequent recommendation for the subregion only would be larger than the potential gain from selection of the top variety for the region as a whole. The comparison of a subregion program versus a statewide program should compare the predicted gain for the same expenditure of resources. Although potential gain from a subregion program may be larger than for a regional program, actual gain may be less for a fixed number of locations and replications since variety means for particular subregions are based on fewer items. Even if actual gains were the same, the subregion program would have additional disadvantages.

The reduction in the variety × location variance component might be even greater if the varieties tested had been selected on a subregion basis in the varietal development program. Many of the varieties tested in the oat yield tests were maintained because of their superior overall performance. With the continual selection of oat strains for subregion adaptation throughout the plant breeding process the reduction might be greater. The definition of natural subregions for oat variety recommendation can be meaningful only if the sample of varieties tested is representative. Implications of subregions on an oat breeding program will be discussed in a subsequent paper.

Table 2.—Ranks and average variety × location mean squares of location pairs.

Rank	Location pair	M.S.	Est. variance comp. (g./plot)^2
1	(3) Marcus, Ames (6)	8882	2492
2	(9) Clarinda, Ames (6)	7832	2142
3	(1) Cresco, Ames (6)	7435	2010
4	(7) Iowa City, Ames (6)	6045	1546
11	(3) Marcus, Clarinda (9)	4815	1136
18	(1) Cresco, Castana (5)	4416	1003
34	(2) Kanawha, Castana (5)	2893	496
36	(7) Iowa City, Corydon (8)	2620	405

Rank	Location pair	Rank	Location pair
1	3, 6	19	1, 3
2	6, 9	20	3, 4
3	1, 6	21	2, 5
4	6, 7	22	3, 7
5	1, 9	23	3, 8
6	2, 3	24	7, 9
7	4, 9	25	4, 5
8	5, 6	26	4, 7
9	4, 6	27	1, 2
10	2, 6	28	5, 7
11	3, 9	29	1, 4
12	1, 8	30	3, 5
13	1, 7	31	5, 8
14	5, 9	32	4, 8
15	2, 9	33	2, 8
16	6, 8	34	2, 5
17	2, 7	35	8, 9
18	1, 5	36	7, 8

Table 3.—General location contribution to variety × location variance components.

Location	Location effects (g./plot)^2
6	647
9	194
3	143
1	105
2	-127
7	-169
4	-178
5	-255
8	-387

SUMMARY

A 5 year experiment using yield data from oat varieties grown at 9 locations in Iowa was examined from the standpoint of subdividing the state into subareas which minimized variety × location interaction within the subareas. Optimum patterns of division into 2, 3, 4 and 5 subregions resulted respectively in a 11%, 21%, 30%, and 40% decrease in the average variety-location component within sections compared to the same statistic for the state as a whole. The graph of the estimated variety × location components for the pattern with lowest mean square against number of subareas resulted in a line which was almost straight. Variety × location interactions which involved location 6 were much higher than those involving other locations. The optimum pattern of subdivision of Iowa into four subareas is shown below:

Subarea	Locations
A	Clarinda, Corydon, and Iowa City (9, 8, 7)
B	Marcus and Castana (3, 5)
C	Ames, Kanawha (6, 2)
D	Cresco and Independence (1, 4)

LITERATURE CITED

1. DUNCAN, DAVID B. Multiple range and multiple F tests. Biometrics 11:1–42. 1955.
2. PEARSON, E. S., and HARTLEY, H. O. Biometrika Tables for Statisticians. University Press, Cambridge, England. 238 pages. 1954.
3. SPRAGUE, G. F., and FEDERER, W. T. A comparison of variance components in corn yield trials: II, Error, year × variety, location × variety, and variety components. Agron. Jour. 43: 535–541. 1951.
4. TIPPETT, L. H. C. Biometrika 17:117–29. 1925.