

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

PATTERN ANALYSIS OF GENOTYPE X ENVIRONMENT
INTERACTIONS AND COMPARISONS WITH
ALTERNATIVE ANALYSES

A thesis presented in partial fulfilment
of the requirements for the degree

of
Master of Science
in
Plant Science
at
Massey University

Carol Cullen

1981

81131-18

Abstract

The occurrence of genotype-environment interactions is a problem affecting the interpretation of cultivar trials. Several analyses have been used to try to resolve the inconsistencies of cultivar performances which occur when these interactions are present.

An assessment of several techniques was carried out using three sets of data. Two sets of barley data came from one season's trials covering the barley growing areas of New Zealand. Ten wheat cultivars were tested in four locations in the lower North Island in two seasons.

The analyses which were examined were Analysis of variance, linear regression, Cluster Analysis and Principal Component Analysis. The parameters of Wricke, Hanson and Eberhart and Russell were also studied.

The Analysis of Variance revealed significant location, genotype and genotype-location interaction effects for the barley data. The wheat data had significant years, years x locations, genotype and genotype-year interactions effects.

There was a strong linear relationship between the genotype means and the environmental index for the Finlay-Wilkinson regression analysis. Following refinement of the error term $\beta_{i,s}$ with significant differences from 1.0 could be seen for several Barley and Wheat genotypes. It was noted that a conflict existed between the aim of finding significant differences from 1.0 and the assumption of

independence of effects for the underlying model. It was suggested that an independent measure of environments be used.

The parameters of Wricke, Hanson and Eberhart & Russell were each related to different concepts of stability and the genotypes ranked accordingly. The three parameters gave reasonably consistent results for the rankings of the cultivars.

In the barley data the cultivars Goldmarker and Magnum had uniformly low rankings. The wheat cultivar Gamenya was generally found to rank highly. These were measures of variability over environments so a high ranking infers a low level of variability and vice versa.

A comparison of the different clustering strategies available was carried out and Wards Incremental Sums of Squares method was chosen as the major strategy. This was applied to each data set using both genotype-environment effects and means. A probabilistic cut off measure was used for truncation of the dendrogram.

The clusters formed could be related to the previous analyses and seemed to adequately summarise the different responses present.

A Principal Component Analysis was carried out and the number of components needed to account for 75% of the total variation were examined. For the barley data sets relatively large numbers of components were needed for this (five and six). This made interpretation and presentation of the genotypic performances difficult. For the wheat data two components

explained a satisfactory level of the total variation and the arrangement of the genotypes on these two axes agreed closely with the clustering results.

Varimax rotation did not aid greatly in the interpretation of the components.

It was felt that the roles of these different analyses were complementary in interpreting genotypic performances.

Acknowledgements

I would especially like to thank my supervisor, Dr Ian Gordon for his assistance and guidance throughout this study.

I would also like to thank:-

- The Department of Agriculture, University of Queensland, in particular D.E. Byth, for supplying the programme YLGR.
- C.B. Dyson, MAF, Ruakura for supplying the data from the Barley cultivar trials.
- J.M. McEwan and staff at CRD, DSIR, Palmerston North for providing the data from the Wheat Cultivar Trials.
- The staff at the Computer Centre, Massey University, for assistance and advice.

Finally I would like to thank my parents for financial and moral support, and my flatmates for tolerance.

CONTENTS

	<u>Page</u>
<u>Abstract</u>	(i)
<u>Introduction</u>	1
<u>Chapter 1 Literature Review</u>	3
1.1 Introduction to the concept of genotype x environment interactions.	3
1.2 Analysis of Variance.	5
1.3 Partitioning of the genotype x environment variance component.	11
1.4 Regression Analysis	14
(i) Basic models.	14
(ii) Modifications suggested for Regression Analysis.	18
(iii) Criticisms of Regression Analysis.	23
1.5 Pattern Analysis	26
(i) Introduction.	26
(ii) Classification	28
- introduction	
- similarity measures	
- sorting strategies	
- which strategy to use?	
- how many clusters?	
- applications to genotype x environment interactions.	
(iii) Ordination	39
- introduction	
- computation of Principal Component Analysis	
- rotation to a terminal solution	
- how many Principal Components?	
- applications to genotype x environment interactions.	
1.6 Crop Evaluation in New Zealand.	47

	<u>Page</u>
<u>Chapter 2 Materials and Methods</u>	49
2.1 Objectives.	49
2.2 Barley Data.	50
2.3 Wheat Data.	53
2.4 Data Analysis	56
(i) Programming.	56
(ii) Analysis of Variance.	58
(iii) Ecovalences.	58
(iv) Hanson's Parameter.	58
(v) Regression Analysis.	59
(vi) Cluster Analysis.	61
(vii) Principal Component Analysis.	63
(viii) Ratio to a Standard Cultivar.	64
<u>Chapter 3 Results and Associated Discussion</u>	66
3.1 Analysis of Variance.	66
(i) Barley: 13 genotypes.	66
(ii) Barley: 14 genotypes.	68
(iii) Wheat.	70
3.2 Stability Parameters.	72
(i) Barley: 13 genotypes.	72
(ii) Barley: 14 genotypes.	74
(iii) Wheat.	76
3.3 Regression - Finlay and Wilkinson	78
(i) Barley: 13 genotypes.	78
(ii) Barley: 14 genotypes.	80
(iii) Wheat.	82
3.4 Regression of Effects.	84
3.5 Cluster Analysis.	90
(i) Strategy Comparisons.	90
(ii) Barley: 13 genotypes.	100
(iii) Barley: 14 genotypes.	107
(iv) Wheat.	115
3.6 Principal Component Analysis	123
(i) Barley: 13 genotypes.	123
(ii) Barley: 14 genotypes.	134
(iii) Wheat.	143
3.7 Ratios to Standard Cultivars	152
(i) Barley: 13 genotypes.	152
(ii) Barley: 14 genotypes.	152
(iii) Wheat.	156

	<u>Page</u>
<u>Chapter 4 General Discussion</u>	157
4.1 Analysis of Variance.	157
4.2 Stability Parameters.	157
4.3 Ratios to Standard Cultivars.	158
4.4 Regression Analysis - Finlay and Wilkinson.	159
4.5 Regression Analysis - Effects.	160
4.6 Pattern Analysis.	160
4.7 Barley Data Sets.	163
4.8 Wheat Data.	166
4.9 Barley Cultivar Assessments.	168
4.10 Wheat Cultivar Assessments.	170
 <u>Conclusions</u>	 171
 <u>Appendices</u>	 173
1. Principal Component Analysis: Barley 13 genotypes in 12 environments.	173
2. Correlation Matrices for Principal Component Analysis.	176
3. Listing of Program Pansy and Associated Subroutines.	179
 <u>Bibliography</u>	 197

List of Tables

1.1	Analysis of Variance Models.	7
1.2	Expectations of Mean Squares.	9
1.3	Regression Models.	15
1.4	Regression Model.	22
2.1	Barley Cultivars.	51
2.2	Barley Trial Sites and Soil Types.	52
2.3	Wheat Cultivars.	54
2.4	Wheat Trial Sites and Soil Types.	55
2.5	Outline of Program Pansy.	57
3.11	Pooled Analysis of Variance: 13 Barley Genotypes.	67
3.12	Pooled Analysis of Variance: 14 Barley Genotypes.	69
3.13	Pooled Analysis of Variance: 10 Wheat Genotypes.	71
3.21	Stability Parameters: 13 Barley Genotypes.	73
3.22	Stability Parameters: 14 Barley Genotypes.	75
3.23	Stability Parameters for Wheat Genotypes.	77
3.31	Barley Regression Parameters: 13 Genotypes.	79
3.32	Barley Regression Parameters: 14 Genotypes.	81
3.33	Wheat Regression Parameters:	83
3.41	Regression Parameters for Effects: 13 Barley Genotypes.	85
3.42	Regression Parameters for Effects: 14 Barley Genotypes.	86

3.43	Regression Parameters for Effects: Wheat.	87
3.44	Adjustment for Covariance Amongst $\alpha\lambda_T$ and λ_T effects.	88
3.45	Adjusted Pooled Analysis of Variance.	89
3.51	Cluster Means for Each Environment: Barley 13 g Means.	101
3.52	Proportion of Variance Not Explained by Clustering (WSS/TSS) Barley 13 g Means.	102
3.53	Cluster Means for Each Environment: Barley 13 g Effects.	104
3.54	Proportion of Variance not Explained by Clustering (WSS/TSS). Barley 13 g Effects.	105
3.55	Cluster Means for Each Environment: Barley 14 g Means.	108
3.56	Proportion of Variance not Explained by Clustering (WSS/TSS) Barley 14 g Effects.	109
3.59	Cluster Means for Each Environment: Wheat Means.	116
3.510	Proportion of Variance not Explained by Clustering (WSS/TSS) Wheat Means.	117
3.511	Cluster Means for Each Environment: Wheat Effects.	120
3.512	Proportion of Variance Not Explained by Clustering (WSS/TSS) Wheat Effects.	121
3.60	Principal Component Analysis: Barley 13 g.	124
3.61	Factor Structure Matrix: Barley 13 g.	125
3.62	Factor Scores: Barley 13 g.	126
3.63	Varimax Rotated Factor Structure Matrix: Barley 13 g.	127
3.64	Factor Scores for Rotated Factors: Barley 13 g.	128

3.65	Principal Component Analysis: Barley 14 g.	135
3.66	Factor Structure Matrix: Barley 14 g.	136
3.67	Factor Scores: Barley 14 g.	137
3.68	Varimax Rotated Factor Structure Matrix: Barley 14 g.	139
3.69	Factor Scores for Rotated Factors: Barley 14 g.	140
3.610	Principal Component Analysis: Wheat.	144
3.611	Factor Structure Matrix: Wheat.	145
3.612	Factor Scores: Wheat.	146
3.613	Varimax Rotated Factor Matrix: Wheat.	148
3.614	Factor Scores for Rotated Factors: Wheat.	149
3.71	Ratio to Standard Cultivar (Zephyr): Barley 13 g.	153
3.72	Ratio to Standard Cultivar (Zephyr): Barley 14 g.	154
3.73	Ratio to Standard Cultivar (Karamu): Wheat.	156

LIST OF FIGURES

1.1	A generalised interpretation of regression coefficients plotted against variety mean yields.	15a
3.51	Dendrogram for Single Linkage Clustering of 13 Barley genotypes means.	91
3.52	Dendrogram for Centroid Clustering of 13 Barley genotypes means.	92
3.53	Dendrogram for Median Clustering of 13 Barley genotypes means.	93
3.54	Dendrogram for Group Average Clustering of 13 Barley genotypes means.	94
3.55	Dendrogram for Average within Groups Clustering of 13 Barley genotypes means.	95
3.56	Dendrogram for Complete Linkage Clustering of 13 Barley means.	96
3.57	Dendrogram for Wards Clustering of 13 Barley genotype means.	97
3.58	Dendrogram for Flexible Clustering of 13 Barley genotype means.	98
3.59	Dendrogram for Wards Clustering of 13 Barley genotype effects.	106
3.510	Dendrogram for Wards Clustering of 14 Barley genotype means.	110
3.511	Dendrogram for Wards Clustering of 14 Barley genotype effects.	114
3.512	Dendrogram for Wards Clustering of Wheat Genotype means.	118
3.513	Dendrogram for Wards Clustering of Wheat genotype effects.	122
3.61	Genotype scores for components 1 and 2 for 13 Barley genotypes - unrotated.	131
3.62	Genotype scores for Components 1 and 2 for 13 Barley genotypes - varimax rotated.	132
3.63	Genotype scores for Components 1 and 2 for 14 Barley genotypes - unrotated.	138
3.64	Genotype scores for Components 1 and 2 for 14 Barley genotypes - varimax rotated.	141
3.65	Genotype scores for Components 1 and 2 for Wheat genotypes - unrotated.	147
3.66	Genotype scores for Components 1 and 2 for Wheat genotypes - varimax rotated.	150

Introduction

In order to assess the worth of a potential cultivar a Plant Breeder will carry out a series of trials which are designed to give an estimation of its relative performance. Ideally the trials should cover the range of environments, both sites and seasons, which will be experienced by the cultivar when grown commercially.

Decisions based on the results of these trial systems become confounded when genotype-environmental interactions are present. These interactions are seen as inconsistencies of relative genotypic performances over environments and mean that comparisons of genotype means are not a true indication of the situation.

The occurrence of genotype-environment interactions has been reported in many crops over a long period of time and has resulted in several different analyses being proposed to handle them (Hill, 1975). The adaptability of a genotype is a concept concerned with its responsiveness to environmental change. This is closely linked to genotype-environment interaction problems and some measure of adaptability has been the aim of some of these analyses.

The aim of a plant breeding program may be for specific adaptation to a particular environmental range or general adaptability to a wide range of environmental conditions. Some measure of adaptation will therefore be important.

Parameters designed for this purpose have involved analysis of variance techniques, partitioning into variance components and regression analysis. A more recent approach has been to describe the pattern of response of a genotype. This has involved the application of several multivariate analyses to trial data, using each test environment as a variable in the analysis.

This study has been designed to assess some of these methods in an attempt to identify the advantages and weaknesses of each. Of particular interest are the Pattern Analysis techniques since these are relatively untried in this field.

The data which has been used comes from regional trials for the evaluation of wheat and barley cultivars in New Zealand. Interpretation of these for the various stability and adaptability concepts has been carried out and related to the New Zealand situation.