Table of contents

List of abbreviations		**vii**
List of tables		**viii**
List of figures		**x**
1.	**Introduction**	**1**
1.1	Background	1
1.2	Genotype and environment	2
1.3	Elements of a breeding strategy	3
2.	**Adaptation and yield stability**	**5**
2.1	Analysis of adaptation	5
2.2	Definition of adaptation strategies	7
2.3	Wide vs. specific adaptation in breeding programmes	9
2.4	Targeting varieties	10
2.5	Assessment of yield stability and reliability	11
2.6	Two analytical flow-charts	12
3.	**Multi-environment yield trials**	**17**
3.1	Types of trials and requirements of the generated information	17
3.2	Additional information on environmental factors and genotypic traits	19
4.	**Analysis of variance (ANOVA) and estimation of variance components**	**21**
4.1	Models of ANOVA	21
4.2	Estimation of individual effects and comparison of means	25
4.3	Estimation of variance components	26
4.4	Data transformation	27
4.5	Computer software	29
5.	**Analysis of adaptation and identification of subregions**	**31**
5.1	Objectives of the analysis	31
5.2	Joint linear regression modelling and complementary analyses	32
5.3	AMMI modelling and complementary analyses	35
5.4	Factorial regression modelling and complementary analyses	44
5.5	Pattern analysis	49
5.6	Data transformation	51
5.7	Unbalanced data sets	55
5.8	Characterization of subregions and scaling-up of results	56
5.9	Computer software	61

6. Definition of adaptation strategy, selection environments, genetic resources and adaptive traits **67**
6.1 Adaptation strategy 67
6.2 Selection environments 71
6.3 Genetic resources and adaptive traits 73

7. Measures of yield stability and yield reliability **79**
7.1 Yield stability 79
7.2 Yield reliability 83
7.3 Computer software 86

8. Case study: durum wheat in Algeria **89**
8.1 Experiment data 89
8.2 Adaptation strategy and yield stability targets for breeding 91
8.3 Cultivar recommendation 96

References **105**

List of abbreviations

AMMI	additive main effects and multiplicative interaction
ANOVA	analysis of variance
BLUP	Best Linear Unbiased Prediction
CGIAR	Consultative Group on International Agricultural Research
CIMMYT	International Centre for Maize and Wheat Improvement
CP	main crossover point
DF	degrees of freedom
GIS	Geographic Information System
GE	genotype x environment
GL	genotype x location
GLY	genotype x location x year
GY	genotype x year
ICARDA	International Centre for Agricultural Research in the Dry Areas
IRRI	International Rice Research Institute
MS	mean square
PC	principal component
REML	Restricted Maximum Likelihood
SS	sum of squares

List of tables

2.1	Complexity of adaptation patterns as depicted by the number of significant PC axes in the analysis of GE and GL data matrices of six data sets	7
3.1	Mean yield of bread wheat varieties over two years and over four other years in trials repeated across a fixed set of locations, for two independent data sets	18
4.1	ANOVA models including the factors G = genotype and L = location or environment, and destimation of variance components, for trials in a randomized complete block design	23
4.2	ANOVA models including the factors G = genotype, L = location and Y = year, and estimation of variance components, for trials in a randomized complete block design repeated in same years (i.e. L and Y crossed factors)	23
4.3	ANOVA models including the factors G = genotype, L = location and Y = year, and estimation of variance components, for trials in a randomized complete block design repeated in different years in each location	23
4.4	Analysis of variance for 18 bread wheat varieties grown for three years in 31 Italian locations, with partitioning of GL interaction by: 1) joint regression analysis; 2) AMMI analysis; 3) definition of four subregions	24
4.5	Calculation of genotype and location main effects, and GL interaction effects, from mean values of genotypes at each location	25
4.6	Values of *t'* for calculation of Dunnett's one-tailed (or Gupta's) multiple comparisons of the top-ranking genotype with the remaining enrtries	26
4.7	Relationships of environment mean yield with experimental error, and of location mean yield with within-location phenotypic variance or standard deviation of genotype values, within-location GY interaction mean square, and average within-location phenotypic variance of annual yield values for individual genotypes, in different data sets	28
5.1	Major differences between two possible objectives for analysis of adaptation	31
5.2	Mean yield of four genotypes in four locations and across locations, mean yield of locations and within-location phenotypic standard deviation of genotype mean yields	52
5.3	Squared Euclidean distance between four locations based on genotype means in each site or genotype x location interaction effects	52
6.1	Predicted yield gain over the region of lowland northern Italy from selection of lucerne populations pecifically adapted to three or two subregions relative to selection for wide adaptation	69

6.2 Mean yield of barley lines previously selected for specific adaptation to either of two subregions or for wide adaptation to the region of northern Syria, yield ratio of specific to wide adaptation strategies, and actual yield gain relative to mean value of six control cultivars 70

7.1 Critical values of Hartley's (1950) test for comparison of several variances, applicable to comparison of yield stability measures of Type 4 83

8.1 Code, altitude, mean yield, scaled score on the first GL interaction PC axis for original and $\log_{10}$-transformed yield data, winter mean temperature and rainfall amount of test sites 90

8.2 ANOVA results and estimate of variance components, for original and transformed yield of 24 genotypes grown at 14 locations for two years 92

8.3 Comparison of analytical methods for definition of two subregions for durum wheat breeding in Algeria. Predicted yield gain over the region from a specific adaptation strategy relative to wide adaptation 94

8.4 Mean yield, environmental variance and Kataoka's index of yield reliability across environments, Type 4 stability variance slope of regression on site mean yield, scaled score on the first GL interaction PC axis, and factorial regression equation for estimating GL effects, for eight genotypes 97

List of figures

2.1 Limited and large extent of within-location variation relative to between-location variation for a major environmental factor related to the occurrence of GE interaction, and its implication on the extent of GE interaction components of variance 6

2.2 A specific adaptation strategy implying plant breeding at one national research centre, selection of novel germplasm and evaluation of genetic resources in distinct subregions, and subsequent phases of farmers' selection in each subregion 9

2.3 Flow chart of steps for definition of the adaptation strategy and yield stability targets of breeding programmes from analysis of multilocation yield trials repeated in time 13

2.4 Flow chart of steps for making variety recommendations from analysis of multilocation yield trials repeated in time 15

5.1 Genotype yields modelled by joint regression and nominal yields modelled by factorial regression on rainfall amount, for durum wheat varieties across Italian locations 34

5.2 AMMI analysis of the genotype-location data matrix in Table 4.4 36

5.3 Scores on the first two GL interaction PC axes of 18 bread wheat genotypes and four Italian subregions 37

5.4 Nominal yields of lucerne varieties modelled as a function of the score on the first GL interaction PC axis of ten Italian locations, or the first GE interaction PC axis of four artificial environments 40

5.5 Scores on the first two GL interaction PC axes of 20 Italian locations, definition of four subregions for bread wheat variety recommendation by grouping sites with the same expected winning genotype, and classification of locations into five groups based on cluster analysis of site scores on PC 1 and PC 2 41

5.6 Cluster analysis of test locations performed on site scores on the first GL interaction PC axis, using the lack of significant GL interaction within group of locations as the truncation criterion for definition of groups 43

5.7 Geographic position of test locations for lucerne in northern Italy, and provisional definition of three subregions for a specific adaptation strategy based on AMMI analysis complemented by cluster analysis and additional information on relevant environmental variables of sites 43

5.8 Yield response of two hypothetical genotypes modelled as a function of site mean yield for original, log-transformed and within-location standardized data under the assumption of within-location phenotypic standard deviation of genotype yields proportional to site mean yield 53

5.9 Nominal yields of genotypes modelled as a function of the first GL interaction PC axis of locations for original and log-transformed data, with indication of scaled PC 1 scores of sites and estimation of the main crossover point 54

5.10 Expected pair of top-ranking cultivars for yield at each combination of 10 annual rainfall by 10 mean winter temperature values of locations, according to an AMMI-1 model in which site score on PC 1 is predicted by multiple regression as a function of the two environmental variables 58

6.1 Hypothetical selection locations for three adaptation strategies compared in terms of predicted yield gains 68

7.1 Yield responses across five environments of hypothetical stable-yielding genotypes according to two concepts of stability, using non-regression or regression stability measures 80

7.2 Frequency of yield (or relative yield) values across environments of two genotypes having same mean yield and contrasting stability as measured by the variance of yield values, and estimation of a yield reliability index equal to the lowest yield that is expected in 75% of cases 84

8.1 Geographic position of test locations and definition of two subregions for a specific adaptation strategy based on pattern analysis and AMMI + cluster analysis results 89

8.2 Nominal yield and nominal yield reliability as lowest yield expected in 80% of cases, of eight cultivars modelled as a function of the scores on the first GL interaction PC axis of 14 locations 98

8.3 Expected pair of top-ranking cultivars for yield at each combination of 10 annual rainfall by 10 mean winter temperature values of locations, according to a two-covariate factorial regression model 99

8.4 Expected pair of top-ranking cultivars for yield reliability at each combination of 10 annual rainfall by 10 mean winter temperature values of locations, according to a two-covariate factorial regression model 103

Foreword

The introduction of improved varieties is one of the most powerful and cost-efficient means of enhancing crop productivity and farmers' incomes. Efficiency in varietal development itself and in the process of matching varieties to production areas implies an understanding of plant responses to diverse environments and cropping systems in a target production zone. Multilocation testing remains the main tool for understanding varietal responses to environments, but the process is both time-consuming and expensive. The efficiency of this analytical process can be enhanced using recently developed statistical methods. This publication aims to support plant breeders by examining the opportunities offered by such methods. FAO hopes that this publication will be useful to a wide variety of persons interested in efficient, sustainable use of plant genetic resources, especially those focusing on the improvement of agriculture in food-deficit developing countries.

Following introductory remarks on the impact of genotype x environment interaction on agricultural production and plant breeding (Chapter 1), adaptation and yield stability concepts are discussed in relation to breeding and the utilization of crop varieties (Chapter 2). The potential usefulness of multi-environment yield trials is also examined (Chapters 2 and 3).

Techniques relating to analysis of variance (Chapter 4) and modelling of adaptation patterns (Chapter 5) are considered for optimizing variety recommendation and for defining the adaptation strategy and yield stability targets in breeding programmes. Attention is paid to limits and opportunities for the scaling-up of results from test sites to the target region. The application of selection theory to the analysis of multi-environment data as well as additional indications that may be obtained from adaptive traits are also considered (Chapter 6). Concepts and measures of yield stability and yield reliability, and their utilization for selection and recommendation of plant varieties, are highlighted in Chapter 7.

Information on useful software for data analysis is provided throughout the book with special emphasis on IRRISTAT, a freely-available software program developed by the International Rice Research Institute. The book also presents a case study (Chapter 8), in which a large multi-environment data set is used for exemplifying the different analytical procedures, as well as the IRRISTAT commands needed for the analysis.

Eric A. Kueneman
Chief
Crop and Grassland Service
Plant Protection Division

Acknowledgements

The revision of some parts of the text by Dr Kaye E. Basford (University of Queensland), Dr Hugh G. Gauch (Cornell University) and Dr Mike Talbot (Biomathematics & Statistics Scotland), and the suggestions made by Dr Manjit S. Kang (Louisiana State University) and Dr Bruce Walsh (University of Arizona) on specific aspects of the work, are appreciated.

Dr Kamel Feliachi (Institut Technique des Grandes Cultures, Alger) and Dr Alice Perlini (Istituto Agronomico per l'Oltremare, Florence) are thanked for permitting the use of data from the Algerian-Italian cooperation project "Amélioration et renforcement du système nationale d'adaptation variétale du blé dur en Algérie" in the case study.

The International Rice Research Institute (IRRI) is thanked for permitting the use of the software IRRISTAT and its tutorial material and for allowing its distribution in the CD attached to this publication. Dr C. Graham McLaren and Ms Violeta Bartolome (IRRI) provided information related to the software.

The excellent work of Ms Ruth Duffy in the editing and formatting of the text is greatly appreciated.

Dr Helena Gómez Macpherson (Cereals Officer, Crop and Grassland Service, FAO) made possible the preparation and release of this publication.