BETWEEN AND

The statistical modelling

BEYOND

of genotype by environment

ADDITIVITY AND

interaction in plant breeding

NON-ADDITIVITY

Promotor: dr. ir. P. Stam
hoogleraar in de plantenveredeling
i.h.b. de selectiemethoden en de resistentieveredeling

Co-promotor: dr. ir. J.-B. Denis
wetenschappelijk medewerker Laboratoire de Biométrie,
INRA-Versailles

Between and beyond additivity and non-additivity; the statistical modelling of genotype by environment interaction in plant breeding

Fred A. van Eeuwijk

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 12 januari 1996
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

Cover design: Herma Klijnstra

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

van Eeuwijk, F.A.

Between and beyond additivity and non-additivity; the statistical modelling of genotype by environment interaction in plant breeding / Fred A. van Eeuwijk - [S.l.: s.n.]
Thesis Wageningen - With summary in Dutch.
ISBN 90-9009007-X

Key-words: Analysis of variance, AMMI-model, Bilinear model, Biplot, Diagnostic biplot, Disease resistance, Factorial regression, Generalized linear model, Genotype by environment interaction, Host specificity, Interaction, Maize, Multiplicative decomposition, Multiplicative model, Non-additivity, Parsimony, Perennial ryegrass, Reaction norm, Reduced rank regression, Resistance breeding, Singular value decomposition, Sugar beet, Three-way interaction, Three-way table, Two-way interaction, Two-way table, Variety testing, Wheat

Stellingen

1. De klant verwacht eerst-en-vooral absolutie van de statisticus.

2. Statistische procedures waarvan de geldigheid berust op voorwaarden die alleen door de toepasser zelf te controleren vallen, dragen niet bij aan de geloofwaardigheid van de statistiek. Denk aan het onderscheid tussen a priori en a posteriori contrasten.

3. Analyses met minder aannames leiden enerzijds wel tot meer robuuste uitspraken, maar anderzijds ook tot minder interessante conclusies.

4. De singuliere-waarde-ontbinding zou binnen het statistiekonderwijs hetzelfde belang toegekend moeten krijgen als de t-toets.

5. Statistici zijn prima in staat binnen de grenzen van het experiment variatiebronnen te onderkennen. Helaas ontberen zij vaak de grip op de meer relevante variatiebronnen tussen experimenten.

6. Al degenen die plantenveredeling als kunst wensen te betitelen, bedoelen niet meer dan dat zij geen idee hebben welke theoretische principes dienstbaar te maken aan de oplossing van hun praktische problemen.

7. Discussies over adaptiviteit en stabiliteit in relatie tot genotypische responsies dienen vervangen te worden door discussies over de wijzen waarop de verwachting en variantie van die responsies gemodelleerd kunnen worden.

8. 'Resistentie' is voor de resistentieveredeling wat 'intelligentie' is voor de psychologie.

9. Statistische methoden die genotype-bij-milieu-interactie beter interpreteerbaar maken, bieden eveneens goede vooruitzichten op een beter begrip van de fenomenen heterosis en specifieke-combinatie-geschiktheid.

10. Er bestaat geen essentieel verschil tussen beschrijvende en verklarende modellen.

11. Voortdurende fysieke bereikbaarheid leidt tot voortschrijdende psychische afwezigheid.

12. Respect voor andermans mening in de zin van positieve waardering voor een van jouw mening afwijkend standpunt is misschien wel de vreemdste gemoedsgesteldheid denkbaar.

BETWEEN AND
The statistical modelling
BEYOND
of genotype by environment
ADDITIVITY AND
interaction in plant breeding
NON-ADDITIVITY

F.A. van Eeuwijk
Wageningen, 12 januari 1996

Bibliographic abstract

In plant breeding it is a common observation to see genotypes react differently to environmental changes. This phenomenon is called genotype by environment interaction. Many statistical approaches for analysing genotype by environment interaction rely heavily on the analysis of variance model. Genotype by environment interaction is then taken to be equivalent to non-additivity. This thesis criticizes the analysis of variance approach. Modelling genotype by environment interaction by non-additivity is little parsimonious and interaction patterns remain hard to interpret. Interpretation is hindered by the multitude of parameters that require interpretation and the fact that these parameters do not refer to external genotypic and environmental information. A viable alternative is presented in the form of multiplicative models for interaction. The latter can be distinguished in two classes; factorial regression models and multilinear models. Factorial regression models describe genotype by environment interaction in direct relation to explicit external genotypic and environmental covariables. They are ordinary linear models that allow the testing of biologically interesting hypotheses about the mechanisms responsible for genotype by environment interaction. Multilinear models are based on low rank approximations to the tables of non-additivity parameters. Parameter estimates can be obtained from multiplicative decompositions of the non-additivity tables. Multilinear models guarantee a parsimonious description of the interaction. When genotypic and environmental interaction parameters are plotted simultaneously in so-called biplots, the emerging patterns often allow biologically interesting conclusions. The successful application of multiplicative models for interaction is illustrated for a number of variables in a number of crops like white cabbage, sugar beet, perennial ryegrass, lettuce, wheat, potato and maize. The data came from plant breeding, resistance breeding, variety trials, and seed technology research. Theoretical contributions include the introduction of reduced rank factorial regression models in plant breeding, the development of generalized bilinear models, and the implementation of quadrilinear models for three-way non-additivity. In addition, the use of diagnostic biplots as a model screening device for two-way tables is described and evaluated. Besides applied and theoretical papers, the thesis contains extensive reviews of the possibilities of linear and bilinear models for modelling genotype by evironment interaction. Two opinion papers provide conceptual clarifications. The thesis not only addresses plant breeders interested in modelling genotype by environment interaction, but also statisticians and researchers interested in parsimonious modelling of interactions.

Contents

Chapter I

A brief introduction

F.A. van Eeuwijk
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

A brief introduction

1 The phenomenon and a simple model

Genotype by environment interaction (GEI) is the shared name for a set of concepts that triggered into existence a multitude of scientific papers, to which those included in this thesis have been added with hopefully sound reasons. In its most elementary form, GEI involves the increase or decrease of the phenotypic difference between a pair of genotypes when going from one set of environmental conditions to another. The appreciation of this phenomenon can differ considerably. In the search for superior *conditional* genotypic performance, the GEI part of the phenotypic response is considered to represent the most valuable part. Alternatively, for *unconditional* recommendations over environments it is a nuisance.

The phenomenon of GEI in plant breeding has elicited the development of statistical/ mathematical models for analysis and description (prediction, explanation). A widely applied, basic statistical model that can account for the occurrence of GEI is the two-way analysis of variance (ANOVA) model with interaction (including sum-to-zero constraints): $\mathcal{E}(P_{ij}) = \mu + g_i + e_j + ge_{ij}$, where $\mathcal{E}(P_{ij})$ stands for the mean phenotypic response ($\mathcal{E}(.)$ is the expectation operator) of genotype i (i = 1,...,I) in environment j (j = 1,...,J), with μ for the general mean, and g_i and e_j for the genotypic and environmental main effects. The most interesting term for our purposes is ge_{ij}, the non-additivity or statistical (genotype by environment) interaction. This term provides us with the facility to incorporate the differential reaction of genotypes to the environment. A stochastic component can be added to the model by the inclusion of an error term, often tacitly assumed to be independently identically distributed with mean zero and constant variance.

2 One way of improvement

The two-way ANOVA model can be used for all kinds of data which can be framed into the form of a two-way table. Because of the, at first sight, obvious 'causal' distinction between the genetic constitution and the non-genetic environment in the realization of the phenotype, the rearrangement of genotypic evaluations in two-way genotype by environment tables seems justified. As a consequence, the two-way ANOVA model seems an appropriate instrument for analysis and description. It can be objected that the non-additivity term in the ANOVA model addresses all kinds of non-parallelism of genotypic responses simultaneously without addressing any one of them in particular. The accompanying (ANOVA) F-test for interaction may thus not be very powerful. Specifying the interaction to be of a particular type leads to more powerful tests. Modelling of non-parallelism by general ANOVA formulations is therefore best restricted to those situations for which no a priori ideas exist about the type of non-parallelism to be expected. The price to be paid for using the general ANOVA

formulation, besides loss of power, is the obligatory estimation of many parameters for 'interaction', (I-1)(J-1), leading to unparsimonious models with parameters that will be difficult to interpret.

One improvement over the ANOVA modelling of interaction would consist in a more parsimonious reformulation (approximation) of the ANOVA interaction, or better non-additivity, increasing the power of the test for interaction and the possibilities for biological interpretation. A popular choice to achieve this for the two-way case is by means of bilinear multiplicative approximations of the non-additivity. The two-way non-additivity, ge_{ij}, is approximated by a sum of bilinear terms of the type $\gamma_{mi}\delta_{mj}$ (m=1,...,M; $M \leq \min(I-1,J-1)$). Bilinear multiplicative terms can be read as differential genotypic sensitivity to hypothetic environmental variables that create maximal distinction between the genotypes. Bilinear formulations of interaction are usually far more parsimonious than the ANOVA formulations. Furthermore, interpretation is facilitated. For the latter, the close relation of multiplicative models with the graphical device of the biplot is of major importance. In the biplot genotypes and environments are jointly depicted by vectors starting at the origin and with end point coordinates determined for the genotypes by the parameters, γ_{mi}, and for the environments by the parameters, δ_{mj}. The pattern of genotypic and environmental vectors can be used for fruitful (biological) interpretations of the GEI, especially when some additional information is available on either genotypes or environments or both.

3 Another way of improvement

Though bilinear formulations for GEI certainly embodied an improvement over ANOVA formulations in terms of non-additivity regarding power of the test, parsimony and interpretability, a major shortcoming of the ANOVA formulation is only partially resolved by the the bilinear approach. Interpretation of ANOVA model parameters is hampered by the fact that they are defined in terms of exclusively phenotypic terms without reference to external genotypic and environmental information. This objection carries over to bilinear terms in so far as these cannot be interpreted on the basis of the emerging patterns in their estimated parameters from biplots. The most preferable way of modelling GEI is that which makes use of explicit genotypic and environmental information. The non-additivity, ge_{ij}, is being replaced by one or more multiplicative terms of the type $\beta_i z_j$, $x_i\tau_j$, or $\lambda x_i z_j$, where β_i represents the genotypic sensitivity (parameter) to the environmental covariable z, τ_j the environmental potentiality (parameter) for the genotypic covariable x, and λ a scaling constant (parameter) for the cross product of the genotypic covariable x with the environmental covariable z. The covariables are not limited to a certain class, they can be continuous, ordinal and nominal. Models incorporating concomitant information on the genotypes and the environments are called factorial regression models. Factorial regression provides the means for parsimonious modelling of GEI, with a high degree of interpretability, and high power for testing for specific types of interaction. Factorial regression is generally recommended for modelling

GEI, but its application can sometimes be problematic. Additional information may be absent (in the required form), be incomplete, or be overcomplete (making selection of the right information complicated). For these situations the first choice would consist in bilinear modelling. It is contended that for the majority of genotype by environment data sets the joint application of bilinear models and factorial regression models should suffice for adequate analysis and interpretation.

4 Further ways of improvement

The basic approaches towards the statistical modelling of GEI are thus defined by bilinear modelling and factorial regression. Some worthwhile extensions are the following. Bilinear modelling is a valuable tool in the analysis of two-way interaction. Bilinear models are a special case of the more general class of multilinear models. For parsimonious modelling of three-way ANOVA interaction, tri- and quadrilinear models are available. In another generalization of bilinear models the assumption of constant variance may be dropped together with the necessity of modelling the response bilinear in the parameters. In generalized bilinear models, not the *response* itself but a *function of the expectation* is bilinear in the parameters, while the variance can be chosen to be a function of the mean. Factorial regression models can be generalized along similar lines.

In the usual (generalized) bilinear and factorial regression models only the error term is assumed to be random, while the other terms are fixed. A natural generalization is to allow more terms to be random. Reasons for choosing model terms to be random can be pragmatic as, for example, the desire to recover information or to derive shrinkage estimators to correct for selection bias.

A hybrid form of bilinear models and factorial regression models is given by the class of reduced rank regression models. Reduced rank regression models are interesting because they allow greater parsimony than factorial regression models while maintaining the property of interpretation of the interaction in relation to explicit genotypic and environmental information. The latter being a factorial regression stronghold.

5 Outline of the thesis

This thesis consists of a collection of papers. Four categories can be distinguished; 1) papers in which the application of methods prevails; 2) papers in which methodology is presented; 3) review papers; and 4) opinion papers. For bibliographical information the review and opinion papers are recommended.

The utility of statistical methods is best illustrated by means of applications to real life problems and real life data. Therefore, a reasonably large number of applied papers is included in this thesis. In chapter II it is shown how for field emergence data in white cabbage, the combined use of factorial regression and bilinear models results in a powerful analysis of GEI and an interesting interpretation from the view point of seed technology.

Chapter III illustrates the same point for an analysis of yield and quality data from sugar beet trials, in which resistance to beet necrotic yellow vein virus was assessed. Chapter IV presents a repetition of this exercise for seed yield in perennial ryegrass. Chapters VI and XIII show that sometimes a bilinear model alone may satisfy the requirements for a model for GEI. In these chapters the question of the nature of the resistance of wheat against funghi of the genus *Fusarium* is addressed. Non-specificity of resistance is inferred from the pattern of the multiplicative environmental parameters in a biplot. Finally, chapter XI contains an example of how to combine the use of bilinear models, factorial regression, reduced rank regression, and mixed models in an analysis of dry matter content in maize.

The theoretical papers treat the following subjects. Chapter V deals with reduced rank regression, chapter XII with generalized bilinear models, and chapter XV with quadrilinear models for the parsimonious description of three-way ANOVA non-additivity. Chapter IX investigates the utility of diagnostic biplots for model screening for genotype by environment tables. Diagnostic biplots present an informal, graphical approach towards the difficult problem of finding parsimonious models for genotype by environment tables.

The chapters VII, X, and XIV contain review papers. Chapter VII summarizes the analyses and results of the chapters IV, V, and VI. The discussion focusses on the merits of bilinear and reduced rank regression models. These models are contrasted with the popular regression on the mean model. Chapter X presents descriptions of fixed and mixed ANOVA, bilinear models, and factorial regression models of full and reduced rank. Also the construction and interpretation rules for biplots for bilinear and reduced rank regression models are given. All the methods reviewed in chapter X are illustrated in chapter XI. The review given in chapter XIV is a rather exhaustive exposition of the possibilities for two-way factorial regression models.

Two opinion papers complete this thesis. In chapter VIII some consequences are discussed of different conceptualizations of GEI within a quantitative genetic context. Finally, chapter XVI is, primarily, an attempt to integrate all major ideas propounded in this thesis, and to elucidate the existing connections between them. Simultaneously, a general philosophy towards modelling GEI is developed.

The order of the papers is based on didactical arguments. First the bilinear and factorial regression models are introduced by means of applied papers (chapters II to IV). Then reduced rank factorial regression is introduced as a logical generalization of both bilinear models and factorial regression (chapter V). Next, a GEI problem is solved using exclusively bilinear modelling (chapter VI). Chapter VII presents a summary of the techniques introduced so far, discusses them, and defines some topics for future research. Further contemplation on concepts and models takes place in chapter VIII. A sidestep is made in chapter IX, to take a closer look on an informal method to arrive at parsimonious modelling. In chapter X the main line of the thesis is taken up again with an exposition of models, where mixed models are added to the instrumentarium consisting of bilinear models and full and reduced rank factorial

regression. The application of all these models to a real life data set follows in chapter XI. Generalized bilinear models are introduced in chapter XII, as an answer to one of the requests for future research in chapter VII. A special kind of application of the algorithm described in chapter XII is illustrated in chapter XIII. A summary of all kinds of factorial regression models follows in chapter XIV. Bilinear and reduced rank factorial regression are subsumed under factorial regression. Mixed factorial regression is briefly discussed. A quadrilinear model for three-way non-additivity forms the topic of chapter XV. Chapter XVI should tie everything together.

Chapter II

Two classes of multiplicative models for the analysis of genotype by environment interaction in seed technology experiments assessing field emergence, with an application to white cabbage

F.A. van Eeuwijk and *H.L. Kraak*
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

Two classes of multiplicative models for the analysis of genotype by environment interaction in seed technology experiments assessing field emergence, with an application to white cabbage

ABSTRACT

When assessing field emergence for a collection of seed lots in a series of trials, it is common to observe changes in the relative performance of seed lots with respect to each other. For describing this phenomenon statistical models for genotype by environment interaction can be used. A frequently applied model in seed technology is the regression on the mean model. In this model interaction consists in the heterogeneity of the coefficients obtained from the regression of individual seed lot field emergences on the mean field emergence per trial. The regression coefficients have been interpreted as reflecting vigour. Because the range of situations that can be modelled by the regression on the mean model is rather limited, there is a need for more powerful models. Two classes of alternative models are presented. The first class of bilinear models is especially useful in exploratory contexts, the second class of factorial regression models in confirmatory settings. Both classes share a multiplicative structure for the interaction, implying that the interaction can be interpreted as originating from differential genotypic sensitivity to critical environmental variables. The difference between the two classes resides in the nature of the genotypic sensitivities and environmental variables. In factorial regression models, genotypic sensitivities and environmental variables are present in the form of measured covariables on the levels of genotypes and environments, and hypotheses concerning the relevance of these covariables to the interaction can be tested. In bilinear models genotypic sensitivities and environmental variables are estimated from the data in the table themselves. Bilinear models permit an interpretation of the interaction in the absence of covariables. For a data set on field emergence in white cabbage (*Brassica oleracea* L.) it is shown how the application of multiplicative models for interaction leads to a better understanding of the pattern of interaction, and how this pattern can be related to the results of germination and vigour tests. A brief discussion is given of the usefulness of laboratory tests for field predictions.

KEY WORDS: AMMI model, Bilinear model, *Brassica oleracea* L., Factorial regression.

INTRODUCTION

For seed lots evaluated under varying environmental conditions it is not unusual to observe that the differences between the seed lots depend on the environment. Occurrence of seed lot by environment interaction seriously impedes the use of laboratory germination tests to predict field emergence, as these tests aim to predict field emergence under 'optimal' conditions only. As a response to the problem of seed lot by environment interaction, laboratory vigour tests have been proposed, as a means to quantify differential emergence in relation to the environment. A preliminary to a good assessment of the value of both germination and vigour tests is a thorough understanding of the seed lot by environment interaction.

A reference model for the analysis of two-way data, e.g. seed lot by environment data, is the familiar two-way analysis of variance (ANOVA) model: $\mathcal{E}(Y_{ij}) = \mu + \alpha_i + \beta_j + \theta_{ij}$ (with sum-to-zero identification constraints for the parameters over the indices i and j). The mean emergence of a seed lot i (i = 1...I) in an environment j (j = 1...J), $\mathcal{E}(Y_{ij})$, where the expectation operator $\mathcal{E}(.)$ indicates that we are considering the mean of the random variable between the brackets, is written as the sum of the general mean, μ, the seed lot main effect, α_i, the environmental main effect, β_j, and the statistical interaction, θ_{ij}. The statistical interaction in this model is equivalent to the non-additivity with respect to the model with only main effects, i.e. the additive model.

The dependence of seed lot differences on the environment in the ANOVA two-way model can be expressed as follows. Consider two seed lots i and i', and two environments j and j'. In environment j the difference between seed lot i and i' is $(\mathcal{E}(Y_{ij}) - \mathcal{E}(Y_{i'j}))$, and in environment j' this is $(\mathcal{E}(Y_{ij'}) - \mathcal{E}(Y_{i'j'}))$. In case of no interaction, i.e. additivity, the difference between two seed lots is the same for every one of the environments, and the so-called tetrad $\tau_{ii'jj'} = (\mathcal{E}(Y_{ij}) - \mathcal{E}(Y_{i'j})) - (\mathcal{E}(Y_{ij'}) - \mathcal{E}(Y_{i'j'})) = (\theta_{ij} - \theta_{i'j}) - (\theta_{ij'} - \theta_{i'j'})$ will always be zero. In the absence of interaction a germination test that correctly assesses the order for an arbitrary environment j suffices for predicting field emergence (order) under all circumstances. When interaction is present, this simple procedure breaks down. At least one tetrad $\tau_{ii'jj'}$ will then not be equal to zero. In the ANOVA two-way model with interaction IxJ parameters θ_{ij} are needed for describing interaction, of which (I-1)(J-1) are independent. With increasing size of the data table it becomes increasingly complicated to recognize pattern in the individual interaction parameters θ_{ij}, and to find biological explanations for these interactions.

A model which gives a more economic description of the interaction coupled to the possibility of a biological interpretation is the regression on the environmental mean model (Yates and Cochran, 1938; Mandel, 1961; Finlay and Wilkinson, 1963). For this model it is assumed that the mean emergence of all seed lots in an environment gives a good indication of the prevailing environmental circumstances. Individual seed lots are supposed to differ in sensitivity to this measure of the environment. In model form; $\mathcal{E}(Y_{ij}) = \mu + \alpha_i + \beta_j + \rho_i\beta_j$ (with sum-to-zero constraints), or more succinctly, $\mathcal{E}(Y_{ij}) = \mu + \alpha_i + \rho_i^*\beta_j$ (sum over ρ_i^*'s = I). The

model can be depicted as a set of converging, diverging, and intersecting straight lines. The heterogeneity in the slopes accounts for the interaction, and absorbs I-1 degrees of freedom. The regression on the mean model is thus far more economic than the two-way ANOVA model with interaction. The slopes may be given an interpretation in terms of vigour (Perry, 1978). Steep slopes mean high sensitivity to the environment thus implying low vigour, flat slopes represent low sensitivity and high vigour.

A necessary condition for this vigour interpretation of slopes to make sense, is that a sufficiently high proportion of the interaction is attributable to the differences in slopes. Seed lot by environment interaction, however, may frequently be too complex to permit adequate description by a regression on the mean model. In these cases more elaborate models for interaction will be necessary. To this end we introduce two alternative classes of models. Firstly, bilinear models (Denis, 1991), also called biadditive models (Denis and Gower, 1992, 1994) and Additive Main effects and Multiplicative Interaction effects (AMMI) models (Gauch, 1988). Secondly, factorial regression models (Denis, 1980, 1988; Snedecor and Cochran, 1980; van Eeuwijk, Denis and Kang, 1995).

After a description of the main features of the models, their application will be illustrated for a data set consisting of field emergence data from white cabbage (*Brassica oleracea* L.). We will show how the models can describe interaction patterns so far undescribable by regression on the mean. The results of the analyses shed new light on the kind of information that is revealed by laboratory tests, and the way these tests may be used to predict field emergence. A brief discussion of this point will end the paper.

THEORY

Bilinear models

A frequently applied model for analysing two-way genotype by environment tables with interaction is the bilinear model (Gollob, 1968; Mandel, 1969; Gauch, 1988, Denis, 1991; Denis and Gower, 1992; van Eeuwijk, 1995). The bilinear model for the mean of genotype (seed lot) i in environment j is written as

$$\mathcal{E}(Y_{ij}) = \mu + \alpha_i + \beta_j + \sum_{m=1}^{M} \sigma_m \gamma_{mi} \delta_{mj} .$$

The parameters μ, α_i, and β_j have the same meaning as in the two-way ANOVA model. The statistical interaction in the bilinear model is represented by a sum of M terms, each term consisting of the product of a multiplicative genotypic interaction parameter, γ_{mi}, with a multiplicative environmental interaction parameter, δ_{mj}. The constants σ_1 to σ_M are proportionality constraints, called singular values, and they are ordered, $\sigma_m \geq \sigma_{m+1}$. Usual identification constraints for the additive as well as the multiplicative parameters are sum-to-zero constraints over the indices i and j. In addition, the following normalization and

orthogonality constraints are often imposed;

$$\sum_{i=1}^{I}\gamma_{mi}^2=1,\ \sum_{j=1}^{J}\delta_{mj}^2=1,\ \sum_{i=1}^{I}\gamma_{mi}\gamma_{m'i}=0,\ \sum_{j=1}^{J}\delta_{mj}\delta_{m'j}=0,\ \textit{with}\ m\neq m'.$$

The model formula shows that when the row parameters, α_i and γ_{mi}, are taken fixed, the bilinear model is a linear model in the column parameters, β_j and δ_{mj}. Alternatively, when the column parameters are taken fixed, the bilinear model is a linear model in the row parameters. Hence, the name bilinear.

There are several reasons, among which notational convenience, for distributing the singular value, σ_m, over the multiplicative parameters for genotypes and environments: $\gamma^*_{mi} = \sigma_m^c\gamma_{mi}$ and $\delta^*_{mj} = \sigma_m^{1-c}\delta_{mj}$, with $0 \leq c \leq 1$. The model then becomes

$$\mathcal{E}(Y_{ij})=\mu+\alpha_i+\beta_j+\sum_{m=1}^{M}\gamma^*_{mi}\delta^*_{mj}\ .$$

It is illuminating to consider the conditional dependence of a genotypic difference on the environment as expressed by the tetrad $\tau_{ii'jj'} = (\mathcal{E}(Y_{ij}) - \mathcal{E}(Y_{i'j})) - (\mathcal{E}(Y_{ij'}) - \mathcal{E}(Y_{i'j'}))$. For the bilinear model

$$\tau_{ii'jj'}=\sum_{m=1}^{M}(\gamma^*_{mi}-\gamma^*_{mi'})(\delta^*_{mj}-\delta^*_{mj'})\ .$$

Thus, two genotypes i and i' behave additively with respect to each other over the total set of environments when $\gamma_{mi} = \gamma_{mi'}$, for all m.

The environmental interaction parameters, or scores, δ^*_{mj}, can be interpreted as the values on hypothetical environmental variables, often called axes, to which the genotypes are differentially sensitive through their genotypic sensitivities (genotypic scores), γ^*_{mi}. The hypothetical environmental variables are mutually orthogonal (uncorrelated) and have decreasing importance for describing the interaction with increasing index m. The first axis can be construed as that environmental variable that discriminates maximally between genotypes. No environmental variable can be found that would describe (linearly) more of the non-additivity (sum of squares). The second axis then describes the maximum amount of non-additivity subject to being orthogonal to the first. The bilinear model presents the upper bound of what may be achieved by modelling interaction multiplicatively. A bilinear model with M multiplicative interaction terms describes more non-additivity (in the sum of squares sense) than whichever other multiplicative interaction model jointly comprising M genotypic covariables and M environmental covariables (but, see below).

Estimated genotypic and environmental scores can be correlated with measured genotypic

and environmental covariables for interpretational purposes. A warning must be given not to adhere too strictly to an interpretation of individual axes in terms of individual environmental variables. It is above all the pattern that emerges from the axes simultaneously that should be interpreted. This is best done by using biplots (see below).

The bilinear model provides a generalization of the regression on the mean model. In the latter model the ANOVA interaction (non-additivity), θ_{ij}, was replaced (approximated) by one multiplicative term of the form $\rho_i\beta_j$. In the bilinear model this becomes, $\gamma^*_{1i}\delta^*_{1j} + \gamma^*_{2i}\delta^*_{2j} + ... + \gamma^*_{Mi}\delta^*_{Mj}$. The bilinear model is more flexible than regression on the mean, because it is not restricted to just one environmental variable and it does not restrict environmental scores to be proportional to the environmental main effects.

The number of multiplicative terms necessary for an adequate description of the interaction, M, should preferably be low. M attains its maximum value when it is equal to the minimum of (I-1) and (J-1). In that case the bilinear model is equivalent to the two-way ANOVA model with interaction, both models have the same number of degrees of freedom for interaction. However, usually M is between one and three and a substantial gain in economy (parsimony) is achieved. For assessing the number of terms to retain, various methods may be used (Gauch, 1988, 1992; Cornelius, 1993). One of the simplest methods tests mean squares for individual axes against an estimate for error (Gollob, 1968). The mean square for axis m is constructed from the ratio of the variation described by axis m, to be calculated by taking the square of σ_m, and the corresponding degrees of freedom, calculated by (I-1) + (J-1) - (2m-1).

For complete tables parameter estimation is straightforward; α_i and β_j can be estimated by the usual procedure for the main effects of a two-way table, whereas σ_m, γ_{mi} and δ_{mj} are derived from a singular value decomposition of the two-way table after correction for main effects (Gabriel, 1978; for Genstat code see van Eeuwijk, Keizer and Bakker, 1995). For incomplete tables alternating least squares algorithms are necessary (Gabriel and Zamir, 1979, Denis, 1991; van Eeuwijk, 1995).

Biplots as visual aids in presenting results from analyses with bilinear models

The results of bilinear analyses can succinctly be presented in graphical form by means of biplots (Gabriel, 1971; Kempton, 1984; Gabriel and Odoroff, 1990). Genotypes and environments are represented by vectors in a space, with starting points at the origin and end points determined by the scores. The bi- in biplot refers to the simultaneous display of both genotypes and environments, and bears no relation to the convention of choosing for two-dimensional, planar, displays. Biplots are mostly planar for two reasons. Firstly, because facilities for three-dimensional graphics are still lacking or rather primitive in many statistical packages (but see van Eeuwijk and Keizer, 1995). Secondly, because for the majority of the cases two bilinear terms capture all, or most of the relevant non-additivity. In two-dimensional biplots genotypes are represented by vectors whose starting point is at the origin, (0,0), and

whose end point is determined by the genotypic scores on the first and second axis, $(\gamma^*_{1i},\gamma^*_{2i})$. Analogously, the environmental vectors start at (0,0) and end at $(\delta^*_{1j},\delta^*_{2j})$. For the scaling constant of the scores, c, a value of 0.5 is throughout reasonable (for other choices see Digby and Kempton, 1987).

The biplot has simple rules for interpretation. The distance between two genotypic vectors (their end points) is indicative of the amount of interaction between the genotypes. The (cosine of the) angle between two genotypic vectors roughly reflects the correlation between the genotypes with regard to their interaction. Acute angles indicate positive correlation, with fully coinciding directions representing a correlation of 1. Obtuse angles represent negative correlations, with completely contrary directions pointing to a correlation of -1. Perpendicularity of directions indicates a correlation of 0.

The relative amounts of interaction for a particular genotype over the environments can be obtained from the orthogonal projections of the environmental vectors on the line determined by the direction of the corresponding genotypic vector. Environmental projections having the same direction as the genotypic vector are typical for positive interactions, projections in the opposite direction for negative interactions. The length of the projection (counting from the origin) is proportional to the size of the interaction. For the absolute interactions the lengths of the environmental projections should be multiplied by the length of the genotypic vector, multiplying again with a factor -1 in case of opposite directions for environmental projection and genotypic vector.

To enrich the biplot various procedures are used. Firstly, the values of categorical and/or continuous covariables on the genotypes and/or environments can be added near the vector end points. This allows a visual inspection of possible covariations between scores and covariables. In the same vein, directions of greatest change for covariables may be drawn in the biplot, where this direction is found from the regressions of the covariables on the appropriate scores. For example, to find the direction of greatest change for a genotypic covariable we use its regression coefficients on the genotypic scores for the first and second axis. Projections of genotypic vectors on this direction yield approximations to the genotypic values for the covariable, in the same way as interactions between genotypes and environments may be found from projections of environmental vectors on lines determined by genotypic vectors.

In the above account the roles of genotypes and environments may throughout be interchanged without problems.

Factorial regression models

The main difference between factorial regression models and bilinear models, is that in factorial regression interaction is modelled directly in relation to genotypic (seed lot) and environmental covariables. A factorial regression model for the mean of genotype i in environment j, for which the interaction includes the genotypic covariables x_1 to x_K is

$$\mathcal{E}(Y_{ij}) = \mu + \alpha_i + \beta_j + \sum_{k=1}^{K} x_{ki}\xi_{kj} \ .$$

The parameters μ, α_i, and β_j again have their by now familiar meanings. The statistical interaction in this factorial regression model consists of the products of the environmental potentialities or valuations, ξ_{1j} to ξ_{Kj}, with respect to the genotypic characterisations, x_1 to x_K ($K \leq I-1$). Convenient constraints on the parameters are sum-to-zero constraints over i for the parameters α_i, and over j for β_j and ξ_{kj}. The genotypic characterisations, x_{1i} to x_{Ki}, can be thought of as known genotypic sensitivities to still unknown environmental variables, whose values are to be estimated. Tetrads have the form

$$\tau_{ii'jj'} = \sum_{k=1}^{K} (x_{ki} - x_{ki'})(\xi_{kj} - \xi_{kj'}) \ .$$

Interaction is absent between the genotypes i and i' when for all k, $x_{ki} = x_{ki'}$.

A factorial regression model in which the interaction part contains the environmental covariables, z_1 to z_H, can be written as

$$\mathcal{E}(Y_{ij}) = \mu + \alpha_i + \beta_j + \sum_{h=1}^{H} \zeta_{hi} z_{hj} \ ,$$

with again sum-to-zero constraints. The interaction in this model consists of the genotypes having differential sensitivity, ζ_{1i} to ζ_{Hi} ($H \leq J-1$), to the environmental covariables, z_1 to z_H. The values of the environmental variables are known, but the genotypic sensitivities need to be estimated. From the tetrad

$$\tau_{ii'jj'} = \sum_{h=1}^{H} (\zeta_{hi} - \zeta_{hi'})(z_{hj} - z_{hj'}) \ ,$$

it is obvious that additivity applies to every pair of genotypes i and i' for which $\zeta_{hi} = \zeta_{hi'}$ over all h. After the description of factorial regression models with exclusively genotypic covariables and factorial regression models with exclusively environmental variables, the structure of factorial regression models including both genotypic and environmental covariables simultaneously presents no new features (Denis, 1988; van Eeuwijk, Denis and Kang, 1995).

Covariables may be quantitative and qualitative (see van Eeuwijk, Denis and Kang, 1995). Qualitative covariables often attribute group membership, i.e. nominal covariables. The incorporation of a nominal covariable in a factorial regression model for interaction, effectively means the partitioning of the non-additivity, θ_{ij}, in a between and within group

component. Consider the case of an environmental covariable that divides the J environments in R groups (index r). The non-additivity for genotype i in environment j then is re-expressed as the sum of the non-additivity of genotype i in the environmental group r, and the non-additivity of genotype i in the environment j within the environmental group r: $\theta_{ij} = \theta_{ir} + \theta_{i,j\subset r}$.

More complicated factorial regression models are possible by combining quantitative and qualitative covariables. We add to the qualitative environmental covariable already introduced, the genotypic covariables x_1 to x_K. We then replace both between and within environmental group non-additivity by regressions on genotypic covariables;

$$\theta_{ij} = \theta_{ir} + \theta_{i,j\subset r} = \sum_{k=1}^{K} x_{ki}\xi_{kr} + \sum_{k=1}^{K} x_{ki}\xi_{k,j\subset r} .$$

An example of a quantitative environmental variable is the mean temperature in the first two weeks after sowing. A qualitative environmental covariable is location. Examples of quantitative genotypic covariables are the results of laboratory germination and vigour tests. A nominal genotypic covariable is given by the company which releases a particular genotype.

The regression on the mean model can be obtained from the factorial regression model by incorporating as exclusive quantitative covariable the environmental main effect, β_j.

The use of covariables need not be restricted to the interaction. Also the main effects may be tried to be described by a regression on covariables. We might want to replace the genotypic main effect, α_i, by a regression on a genotypic covariable like x_1; $\alpha_i = \phi x_{1i} + \alpha_i^*$, where α_i^* can be interpreted as a kind of residual.

Factorial regression models provide the means to test hypotheses about the relevance of specific genotypic and environmental covariables to the values of main effects and interactions. They are suitable for a confirmatory approach, whereas bilinear models may be preferable in exploratory contexts.

The individual terms in factorial regression models contain either row parameters or column parameters, but not both simultaneously as in the bilinear model. Therefore, the factorial regression model is an ordinary linear regression model, and all the familiar estimation and testing procedures for regression are valid and can be used (see Denis, 1988, 1991; van Eeuwijk, Denis and Kang, 1995).

AN APPLICATION TO WHITE CABBAGE FIELD EMERGENCE DATA

Data description

In 1988 field experiments were performed with white cabbage (*Brassica oleracea* L.). Twelve seed lots coming from seven cultivars of white cabbage were kindly provided by various seed companies. The seeds were graded 1.75 - 2.00 mm or 2.00 - 2.25 mm, and were used for experiments either untreated or disinfected with 10 g Rovral (N-isopropyl-3-(3,5-

dichlorophenyl)-2,4-diozo-imidazolidin-1-carboxamide) and 5 g AAtiram (50% active ingredient thiram, tetramethyl-thiuramdisulphide) per kg of seed. The data to be analysed consisted of the final emergence percentages of the 12 seed lots as sown at three locations (soil types) on either three or four dates over the season, thus making ten environments (Table 1). In each environment, i.e. soil-sowing date combination, the 12 seed lots were sown in a randomized complete block design consisting of four replicates. Per replicate per seed lot 200 seeds were sown.

In addition to the field emergence data, results from a number of laboratory tests were available; standard germination tests with untreated seeds, the same test with disinfected seeds, a cold test, and conductivity tests for intact and naked seeds (Table 2). For the standard germination tests, seedlings were classified as normal or abnormal according to ISTA rules (International Seed Testing Association, 1985). Four replicates of 100 seeds each were used. In the cold test, a medium textured sand of pH 5.6 and 10.5% organic matter was used at 43%

Table 1

Table of seed lot by environment field emergence means. Seed lots are marked 1 to 12, environments a to j. For the environments additional information is given concerning the location (A to C), the trial within a location (1 to 3, or 4), and the sowing date (number of days between April 6 and the day of sowing).

Environment	a	b	c	d	e	f	g	h	i	j	
Location	A	A	A	B	B	B	B	C	C	C	
Trial	1	2	3	1	2	3	4	1	2	3	
Sowing date	6	23	78	0	5	30	61	6	43	120	
Seed Lot											Mean
1	78	63	67	53	28	61	75	34	77	64	60
2	82	62	65	64	28	62	72	55	77	74	64
3	75	40	56	53	15	46	58	56	76	73	55
4	83	61	61	55	28	63	79	46	69	64	61
5	87	69	63	72	30	72	84	46	76	76	68
6	90	76	80	79	38	73	83	65	83	80	75
7	90	84	75	80	40	74	87	68	80	77	75
8	85	66	73	73	28	63	78	49	71	68	65
9	75	61	63	54	24	59	73	44	74	67	59
10	67	47	42	47	14	50	60	40	76	65	51
11	83	63	69	69	28	69	79	47	73	71	65
12	86	75	64	70	34	68	83	63	78	72	69
Mean	82	64	65	64	28	63	76	51	76	71	64

Table 2

Additional information for the seed lots, including results of standard germination tests for untreated (Grm. U.) and disinfected seeds (Grm. D.), results of a cold test (Cld), and results of conductivity tests with naked (Cnd. N.) and intact seeds (Cnd. I.).

Seed lot	Genotype	Grm. U.	Grm. D.	Cld	Cnd. N.	Cnd. I.
1	I	90	65	77	22.4	10.0
2	II	88	81	86	25.3	13.4
3	II	91	93	86	36.2	10.8
4	III	81	74	81	21.3	11.5
5	IV	94	86	93	16.2	6.4
6	IV	93	92	97	18.6	7.5
7	IV	94	93	98	19.6	7.0
8	V	90	83	87	16.5	*
9	I	77	79	82	23.6	8.6
10	II	77	75	66	37.5	20.2
11	VI	87	80	93	14.5	6.9
12	VII	92	91	90	19.5	5.5

moisture content, which corresponded to 60% of the maximum water holding capacity. Four replicates of 100 seeds each were sown 2.5 cm deep and placed for 14 days at 5 °C in darkness. Subsequently, they were transferred to 20 °C and illuminated during eight hours per day. Normal seedlings were counted seven days after transfer to 20 °C. For the conductivity test, for each seed lot 100 intact seeds and 50 seeds from which the seed coat was removed (naked seeds) were soaked individually in 3 ml distilled water for 24 hours at 10 °C (EJJM van Eijk, *De EC-test als kwaliteitstest voor koolzaad.* Unpublished Report, 1991, Agricultural University Wageningen, The Netherlands.). The electrical conductivity of the soak water was measured with a multiprobe seed analyzer (ASAC-1000, Neogen Food Tech Corporation, Lansing, Michigan, USA) at 4 Volt. Results were corrected for seed weight. To facilitate isolation of naked seeds, the seeds were first equilibrated for three days at 10 °C and 75% RH in a drum with ventilated air. After slight drying, the seed coat was scratched with a sharp needle, and seed coat and naked seed were easily separated.

ANOVA

For creating a reference point for more complicated analyses, an ANOVA model was fitted to the white cabbage field emergence percentages. Treatment terms in that model were the main effects of seed lots and environments and their interaction. Within the individual

trials, randomized complete block designs were used. Therefore, the environmental main effect had to be be tested over another error (error 1 = blocks within environments) than the seed lot main effect and the seed lot by environment interaction (error 2 = seed lots times blocks within environments interaction). The ANOVA model fitted the data satisfactory. Treatment terms accounted for 87% of the total sum of squares (Table 3; First add up the sums of squares due to environments, seed lots, and their interaction, and divide this sum by the total sum of squares, being the sum of the treatment sums of squares plus both error sums of squares. Multiply this ratio by 100.), and the residuals behaved well (no pattern, not shown). The majority of the variation was due to differences between environmental means, 65%. Differences between seed lots contributed 15%, and the interaction another 7%. The interaction was significant. As environments were chosen to vary widely, the importance of the interaction is best assessed by comparing its sum of squares to that of the seed lot main effect. Then the interaction is not particularly small. Possibly, differences in vigour were responsible for this interaction.

Regression on the mean

A simple statistical translation of the physiological vigour concept is given in the slope for a seed lot in the regression on the mean model. Table 3 shows the partitioning of the seed lot by environment interaction corresponding to the application of the regression on the mean model. Heterogeneity of regression coefficients for the regression of seed lot field emergence on mean field emergence was not significant and explained very little of the interaction. Another model than the regression on the mean model was necessary for describing the interaction.

Table 3.
Two-way analysis of variance and regression on the mean for field emergence data.

Source	Degrees of Freedom	Sum of Squares	Mean Square	Variance ratio	F probability
Environment	9	101620.17	11291.13	70.55	< 0.001
Error 1	30	4801.64	160.05		
Lot	11	23666.67	2151.62	46.06	< 0.001
Lot.Environment	99	10593.14	107.00	2.29	< 0.001
Regression on environmental mean	11	428.02	38.91	0.83	0.607
Deviations	88	10165.12	115.51	2.47	< 0.001
Error 2	330	15413.17	46.71		

Bilinear model and biplot

Application of a bilinear model was far more successful than that of the regression on the mean model. The first two bilinear interaction terms accounted for 72% of the interaction sum of squares, using 36 degrees of freedom (Table 4). Both terms were found significant, while the deviations were non-significant.

Fig. 1a shows a biplot of the bilinear interaction. Vector representations of the seed lots are given by lines starting at the origin and ending in the points determined by the scores on first and second axis. For the environmental vectors only the end points (environmental scores on first and second axis) are given. In the biplot of Fig. 1a we first notice that of the seed lots especially the lots 1, 3, 7 and 10 behaved strongly interactive as their vector end points are far removed from the origin, which represents an imaginary additively behaving seed lot. Of the environments especially the environments h, i and j exhibited interaction (again far removed from the origin, an imaginary additive environment). To a lesser extent, environments b, d and g were interactive. As a rule, conclusions from biplots are most trustworthy for genotypes and environments with high non-additivity, as these are best represented in biplots. Therefore, an interpretation of the interaction should be mainly based on seed lots 1, 3, 7 and 10, and the environments h, i and j (b, d, g).

With respect to the *size* of the interaction between pairs of seed lots, we observe that there was considerable interaction between the seed lots 1 and 3, as their vector end points are far apart. The same holds true for the pairs of seed lots 1 and 7, 1 and 10, 3 and 7, and 7 and 10. The *pattern* of interaction was, ignoring size, quite comparable for the groups of seed lots 1 and 9; 2 and 3; 4 and 5; and 6, 7, 8 and 12. We conclude this from the almost coinciding vector directions. The opposite directions of 3 and 11, and 8 and 10 reveal that these pairs of lots had interaction patterns with almost perfect negative correlation, -1. The orthogonal angle between the vectors for 3 and 6 (also 12) points to unrelatedness of the interaction patterns.

Interactions between environments were mainly due to differences between the environments h, i, and j on the one hand, and the rest of the environments on the other hand. This separation takes place along the first, horizontal, axis. In fact, along the first axis the trials at location (soil) C (Table 1) are separated from those of the locations A and B (the locations A, B, and C have different plotting symbols in Fig. 1a, see also Table 1). Along the second, vertical, axis, trials within locations are separated from each other, but with no clear relationship to the sowing date (check the order of a, b, c/ d, e, f, g/ h, i, j along axis 2).

Now we bring to the attention that the seed lots 2, 3 and 10 were all derived from the same small seeded genotype, and that the trials h, i, and j were done on soil belonging to the breeding company that developed this genotype. Recall that when environmental projection and genotypic vector have the same direction this indicates positive interaction. We then conclude that an important part of the interaction was due to the genotype making up the seed lots 2, 3 and 10, which performed above average at the location where it was developed, and less than average at the locations of competing companies. This can be read into the biplot

Table 4.
Decomposition of the seed lot by environment interaction for model with two bilinear interaction terms (Bilinear int. 1 and 2).

Source	Degrees of Freedom	Sum of Squares	Mean Square	Variance ratio	F probability
Lot.Environment	99	10593.14	107.00	2.29	< 0.001
Bilinear int. 1	19	5230.18	275.27	4.78	< 0.001
Bilinear int. 2	17	2444.31	143.78	2.50	< 0.001
Deviations	63	2918.65	46.33	0.99	0.503
Error 2	330	15143.17	46.71		

by projecting the environmental vectors h, i and j on the genotypic vectors 2, 3 and 10, and contrast these to the projections of the other environmental vectors. As an aid to this kind of interpretation we have drawn in Fig. **1a** for seed lot 3 the projections of the environmental vectors b, g, h, i and j on the seed lot vector.

To investigate the relationships between additive and multiplicative parameters in the bilinear model and the values of seed lot and environmental covariables, we correlated seed lot main effects and scores for first and second axis with a number of seed lot covariables (Table 5). The seed lot main effect was strongest related to the results of the cold test, with a correlation of 0.880 (we ignore, of course, the auto correlation of 1). The seed lot scores of

Table 5.
Correlations between seed lot covariables and bilinear model parameters, and the proportion of variation explained in a seed lot covariable by the regression on the seed lot scores, R^2.

Concomitant Variable	Seed Lot Mean	Seed Lot Scores Set 1	Seed Lot Scores Set 2	R^2
Seed Lot Mean	1.000	0.586	-0.702	0.836
Standard Germ. Test untreated seed	0.683	0.175	-0.551	0.334
Standard Germ. Test disinfected seed	0.557	-0.252	-0.874	0.828
Cold Test	0.880	0.335	-0.778	0.718
Conductivity Test naked seeds	-0.760	-0.892	0.314	0.894
Conductivity Test intact seeds	-0.741	-0.548	0.477	0.528

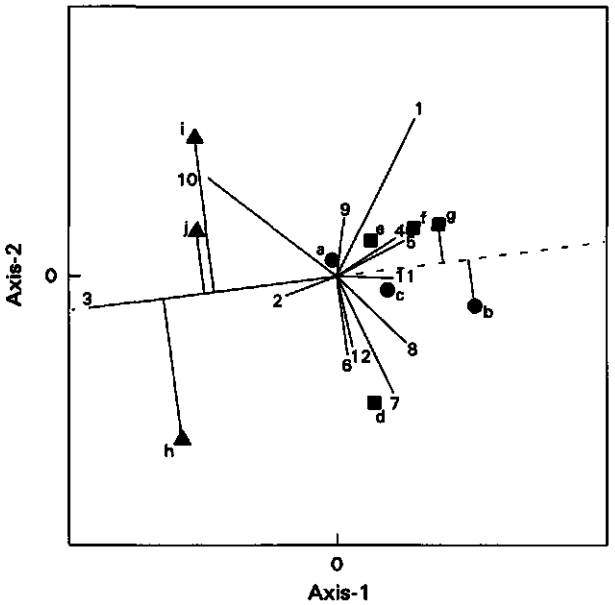


Fig. 1a. Biplot representation of the bilinear interaction, with seed lots 1 to 12 and environments a to j, scaling constant c = 0.5. Environments sharing a location have the same plotting symbol. For seed lot 3 the environmental projections are given for the environments b, g, h, i & j.

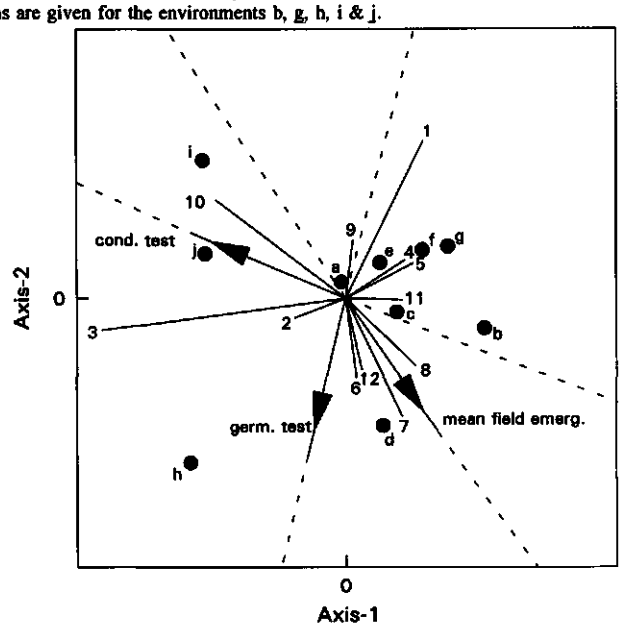


Fig. 1b. Biplot representation of the bilinear interaction, scaling constant c = 0.5, with directions of greatest change for some seed lot covariables indicated by arrow heads. The position of a covariable arrow head relative to the total length of the solid line indicates the quality of the representation of the covariable. The end of the solid line represents $R^2 = 1$.

the first axis most clearly resembled the results of the conductivity test with naked seeds, -0.892. The seed lot scores of the second axis resembled the results of the germination test with disinfected seeds, -0.874. However, precaution should be exercised towards interpreting axes individually. To help interpret the biplot pattern when considering both axes simultaneously, we added to the biplot the directions of greatest change for the genotypic covariables (Fig. 1b), as obtained from the regressions of the covariables on first and second axis. This regression may explain a smaller or larger proportion of the variation in the covariable, R^2 (Table 5). It only makes sense to indicate directions of greatest change when a substantial proportion of the variation can be explained. Therefore, only three of the five genotypic covariables appear in Fig 1b. (The positions of the end points of the solid lines in the covariable representations indicate perfect description, $R^2 = 1$, the arrowheads indicate the observed value of R^2.)

As the direction of greatest change for the conductivity test is roughly from right to left along the first axis, we conclude that the seed lots along the first axis are ordered from right to left in increasing order with regard to the values for that test. Along the second axis, seed lots are ordered from top to bottom in increasing order corresponding to the results of the standard germination test with disinfected seeds and the mean field emergence. Relative values for seed lots on the covariables are easily found from the projections of the seed lots on the covariable directions. It must be remembered that these values are only approximations, probably best for the most interactive seed lots and environments. Therefore, we restrict ourselves to the seed lots 1, 3, 7 and 10. Seed lot 1 is below average for the standard germination test with disinfected seeds. Seed lot 3 is above average for the conductivity test and mean field emergence. Seed lot 7 is above average for mean field emergence. Seed lot 10 is above average for the conductivity test. That the biplot provides good approximations to the values for the genotypic covariables for the seed lots 1,3, 7 and 10, can be checked in Table 1. However, the interest is not in the values themselves, but in the interaction pattern for the seed lots in relation to other properties of the seed lots. It is here where the strength of the biplot representation resides.

The procedure used to enlighten the relation between interaction and other characteristics for the seed lots, can also be applied to the environments. However, for the environments we had only one continuous covariable, sowing date in days after April 6. Regression on the environmental scores led to a $R^2 = 0.227$, too low to justify inclusion in the biplot. The discrete environmental covariable 'location' was already added in the form of different plotting symbols for different classes.

Factorial regression models

As an illustration of the use of factorial regression models in the analysis of seed lot by environment interaction, Table 6 shows the results of five factorial regressions, four for individual seed lot covariables (the conductivity test with intact seeds is left out, because of

a missing value), and one for the environmental covariable sowing date. The best covariable was the conductivity test with naked seeds. This was in agreement with the results of Table 5, where this test had the highest R^2. The germination test with disinfected seeds and the cold test also may be used to describe the interaction, but their explanatory power is only half of that of the conductivity test.

Following the suggestions from the biplot and the factorial regression models with one covariable, we fitted a factorial regression model in which (1) the qualitative environmental covariable location was included, i.e. the interaction between seed lots and locations (roughly; first axis of biplot) was separated from the interaction between seed lots and trials within locations (roughly; second axis of biplot), and (2) two seed lot covariables were included, namely the conductivity test with naked seeds (roughly; first axis of biplot and most important individual covariable) and the standard germination test with disinfected seeds (roughly; second axis of biplot and second most important individual covariable).

Table 7 gives the partitioning of the seed lot by environment interaction according to this model. The trends observed in the biplot re-appeared. For the between location component of the interaction, the conductivity test was far more important than the germination test. In contrast, for the within location component it was especially the germination test that was important. The joint inclusion of conductivity and germination test left non-significant deviations.

The difference between a bilinear model with two bilinear terms and the factorial regression model with location, and conductivity and germination test was non-significant. The bilinear model described 7674.49 of the non-additivity (interaction) sum of squares with 36 degrees of freedom. The factorial regression model described 6864.63 with 18 degrees of freedom. The difference of 809.86 with 18 degrees of freedom gives a mean square of 44.99, hardly different from the error of 46.71. Thus, a factorial regression model can be more

Table 6.

Amount of seed lot by environment interaction described by factorial regression models with one covariable (for codes see Table 1 & 2).

Source	Degrees of Freedom	Sum of Squares	Mean Square	Variance ratio	F probability
Lot.Environment	99	10593.14	107.00	2.29	< 0.001
(Cnd.N.).Environment	9	4461.45	495.72	10.61	< 0.001
(Grm.D.).Environment	9	2245.67	249.52	5.34	< 0.001
Cld.Environment	9	2127.42	236.38	5.06	< 0.001
(Grm.U.).Environment	9	979.92	108.86	2.33	0.013
Seed lot.Sowing date	11	876.44	79.68	1.71	0.070
Error 2	330	15143.17	46.71		

Table 7.

Factorial regression for the seed lot by environment interaction, with one qualitative environmental covariable and two quantitative seed lot covariables (for codes see Table 1 & 2).

Source	Degrees of Freedom	Sum of Squares	Mean Square	Variance ratio	F probability
Lot.Environment	99	10593.14	107.00	2.29	< 0.001
Lot.Location	22	4943.18	224.69	4.81	< 0.001
(Cnd. N.).Location	2	3893.07	1946.54	41.68	< 0.001
(Grm. D.).Location	2	482.79	241.40	5.17	0.006
Deviations	18	567.32	31.52	0.67	0.836
Lot.Trial	77	5649.96	73.78	1.57	0.004
(Cnd. N.).Trial	7	568.38	81.20	1.74	0.099
(Grm. D.).Trial	7	1920.18	274.31	5.87	< 0.001
Deviations	63	3161.19	50.18	1.07	0.339
Error 2	330	15413.17	46.71		

efficient than a bilinear model, provided one knows which covariables to include. The comparison is somewhat inflated, because we used the bilinear model for selecting interesting factorial regression models.

DISCUSSION AND CONCLUSION

Germination tests are commonly defined to be tests that predict the field emergence of seed lots under optimal environmental conditions. Vigour tests should reflect conditional environmental dependence. Both definitions seem hard to handle in practice. How to know the optimal conditions for germination tests? And, is it clear what kind of conditional dependence we have in mind when talking about vigour? With regard to germination tests, we remark that these tests become of limited value as soon as whichever type of interaction occurs. With regard to vigour, we should define which kind of statistical model we have in mind for the interaction, and how our statistical model parameters are to be translated to biological vigour parameters. When an adequate statistical model is chosen, there is, however, no need to make a distinction between laboratory tests as addressing either exclusively optimal conditions or differential conditions. A laboratory test can always be used for predicting field emergence, whether in general or conditional, without us understanding very profoundly the reason for the quality of the prediction.

We have shown how the use of bilinear and factorial regression models can lead to a better understanding of seed lot by environment interaction in seed technology. Both types of models provide far more opportunities for modelling and interpretation of interaction than the

traditional regression on the environmental mean. Laboratory tests may be useful in the description of seed lot by environment interaction for field emergence. Predictive models may be constructed on basis of laboratory tests. It is advised for the construction of these models to abstain from classification of tests as exclusively addressing optimal conditions or differential conditions.

Acknowledgements

The authors are indebted to V. van Bentum, J. van Bilsen, K. Brantjes, G. Jochems, J. Schipper, D. Sleutel, and R. Weges for providing the seed lots and for performing field experiments and some laboratory tests. They thank J.-B. Denis and R.J. Bino for comments and discussion, and L.C.P. Keizer for his help in the preparation of the manuscript.

REFERENCES

Cornelius PL. 1993. Statistical tests and retention of terms in additive main effects and multiplicative interaction model for cultivar trials. *Crop Science* **33,** 1186-93.

Denis J-B. 1980. Analyse de régression factorielle. *Biométrie et Praximétrie* **20,** 1-34.

Denis J-B. 1988. Two way analysis using covariates. *Statistics* **19,** 123-32.

Denis J-B. 1991. Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Revue Statistique Appliquée* **XXXIX,** 5-24.

Denis J-B, Gower JC. 1992. Biadditive models. *Technical report. Laboratoire de Biométrie, INRA-Versailles,* pp. 33.

Denis J-B, Gower JC. 1994. Biadditive models. Letter to the editor. *Biometrics* **50,** 310-11.

Digby PGN, Kempton RA. 1987. *Multivariate Analysis of Ecological Communities.* London: Chapman and Hall.

Finlay KW, Wilkinson GN. 1963. The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research* **14,** 742-54.

Gabriel KR. 1971. The biplot display of matrices with application to principal component analysis. *Biometrika* **58,** 453-67.

Gabriel KR. 1978 Least squares approximation of matrices by additive and multiplicative models. *Journal of the Royal Statistical Society Series B* **40,** 186-96.

Gabriel KR, Odoroff CL. 1990. Biplots in biomedical research. *Statistics in Medicine* **9,** 469-85.

Gabriel KR, Zamir S. 1979. Lower rank approximations of matrices by least squares with any choice of weights. *Technometrics* **21,** 489-98.

Gauch HG Jr. 1988. Model selection and validation for yield trials with interaction. *Biometrics* **44,** 705-15.

Gauch HG Jr. 1992. *Statistical analysis of regional yield trials.* Elsevier: Amsterdam.

Gollob HF. 1968. A statistical model which contains features of factor analytic and analysis of variance techniques. *Psychometrika* **33,** 73-115.

Kempton RA. 1984. The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science Cambridge* **103,** 123-35.

Mandel J. 1961 Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **56,** 878-88.

Mandel J. 1969. The partitioning of interaction in analysis of variance. *Journal of Research of the National Bureau Of Standards - B* **73B,** 309-28.

Perry DA. 1978. Problems of the development and application of vigour tests to vegetable seeds. *Acta Horticulturae* **83,** 141-46.

Snedecor GW, Cochran WG. 1980. *Statistical Methods 7th edn.* Ames: IOWA State University Press.

van Eeuwijk FA. 1995. Multiplicative interaction in generalized linear models. *Biometrics* **51,** 1017-32.

van Eeuwijk FA, Denis J-B, Kang MS. 1995. Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: Kang MS, Gauch HG Jr, eds. *Genotype-by-environment interaction: New perspectives,* pp. 15-49. Boca Raton, Florida: CRC Press.

van Eeuwijk FA, Keizer LCP. 1995. On the use of diagnostic biplots in model screening for genotype by environment tables. *Statistics and Computing* **5,** 141-153.

van Eeuwijk FA, Keizer LCP Keizer, Bakker JJ. 1995. Linear and bilinear models for the analysis of multi-environment trials: II. An application to data from the Dutch Maize Variety Trials. *Euphytica* **84,** 9-22.

Yates F, Cochran WG. 1938. The analysis of groups of experiments. *Journal of Agricultural Science Cambridge* **28,** 556-80.

Chapter III

Multiplicative models for cultivar by loction interaction in testing sugar beets for resistance to beet necrotic yellow vein virus

H. Paul,[1] F.A. van Eeuwijk[1] & W. Heijbroek[2]

[1] *DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)*
P.O.Box 16
6700 AA Wageningen
The Netherlands

[2] *Sugar Beet Research Institute (IRS)*
P.O.Box 32
4600 AA Bergen op Zoom
The Netherlands

Euphytica 71: 63–74, 1993.

Multiplicative models for cultivar by location interaction in testing sugar beets for resistance to beet necrotic yellow vein virus

H. Paul,[1] F.A. van Eeuwijk[1] & W. Heijbroek[2]
[1] *DLO-Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 16, 6700 AA Wageningen, The Netherlands;* [2] *Sugar Beet Research Institute (IRS), P.O. Box 32, 4600 AA Bergen op Zoom, The Netherlands*

Received 6 January 1993; accepted 2 September 1993

Key words: Additive Main effects and Multiplicative Interaction effects model (AMMI), *Beta vulgaris*, beet necrotic yellow vein virus, cultivar by location interaction, factorial regression, sugar beet, virus resistance

Summary

Sugar beet cultivars were evaluated for resistance to beet necrotic yellow vein virus (BNYVV) on various locations in two consecutive years. Resistance levels of cultivars were measured by virus assays of plants from the field and the greenhouse. Infection levels in the fields were characterised by sampling plants of a susceptible indicator cultivar. For each year, statistical analyses were performed on two-way tables of cultivar by location for yield and quality parameters. In analysis of variance (ANOVA) significant main effects and significant cultivar by location interaction were found for all parameters ($P < 0.05$). Interactions were further investigated by multiplicative models. In the Additive Main effects and Multiplicative Interaction effects (AMMI) model, interaction was written as the product of a cultivar score and a location score. Cultivar interaction scores were highly correlated to virus concentrations of the cultivars, and location interaction scores to virus concentrations of the susceptible indicator cultivar. Main effects of cultivars and locations were less clearly related to virus concentrations than interaction effects. In general, virus concentrations of plants from a greenhouse test gave higher correlations than virus concentrations of plants from the field. In the factorial regression model, virus concentrations were incorporated in the model. The model can be understood as a two-way ANOVA, with greenhouse virus concentrations and virus concentration of the indicator cultivar as concomitant variables on the cultivar and location factor. Results of analyses with both multiplicative interaction models showed that interactions of all yield and quality parameters can be described in terms of virus concentrations. Therefore, the relative performance of susceptible and partially resistant cultivars in infested fields can be estimated by means of three independent parameters, (i) the level of resistance determined in a greenhouse experiment, (ii) the yield and quality in non-infested fields, and (iii) the level of infection in the field.

Abbreviations: AMMI model – Additive Main effects and Multiplicative Interaction effects model, ANOVA – analysis of variance, BNYVV – beet necrotic yellow vein virus, ELISA – enzyme-linked immunosorbent assay, α-amino N – α-amino nitrogen, K – Kalium (potassium), Na – Natrium (sodium)

Introduction

The relationship between the performance of sugar beets (*Beta vulgaris* L.) in fields infested with beet necrotic yellow vein virus (BNYVV) and the virus concentration in the roots has been investigated in several studies (Ahrens, 1987; Bürcky & Büttner, 1989b, 1991; Giunchedi et al., 1987; Hillmann, 1984; Shimada et al., 1989). Sugar beet cultivars with various levels of resistance to BNYVV were incorporated in all studies. Negative correlations were found between virus concentrations and the parameters root yield, sugar content and sugar yield. Giunchedi et al. (1987) and Hillmann (1984) included quality parameters in their study and found a positive correlation between virus concentrations and sodium content, whereas a negative correlation was found between virus concentration and α-amino nitrogen content. Correlation with potassium content was not always clear. In these studies, the plant material for the virus assays either came from infested fields, or from greenhouse experiments, which led to similar results (Bürcky & Büttner, 1991). The results from a single field were used (Ahrens, 1987, Bürcky & Büttner, 1991; Hillmann, 1984) or the average data of several fields (Bürcky & Büttner, 1989b; Giunchedi et al., 1987; Hillmann, 1984). Shimada et al. (1989) used multiple regression analysis to study the data from various fields.

In non-infested fields, susceptible sugar beet cultivars generally perform better than partially resistant cultivars. When the level of infestation with BNYVV increases, resulting in higher levels of infection in the beets, the ranking of the cultivars changes and in severely infested fields highest yields are obtained by partially resistant cultivars (Fig. 1). Thus, an adequate statistical model for the analysis of data from cultivars with various levels of resistance in fields with varying levels of infestation has to include terms for the description of interaction.

In the present study, special attention is given to the explanation of cultivar by location interaction by means of virus concentrations. Virus concentrations are used to determine the resistance level of the cultivars as well the infection level in the trial fields. In order to describe the interaction, three statistical models are considered in this paper. Firstly, a two-way analysis of variance (ANOVA) model with interaction is used, the factors being cultivars and locations. In this model, each cell of the cultivar by location table has its own interaction parameter, so that the model uses a relatively large number of degrees of freedom for the interaction, and results of the description of the interaction are usually difficult to interpret. Secondly, a model with additive main effects and multiplicative scores for cultivars and locations is used, the so-called Additive Main effects and Multiplicative Interaction effects (AMMI) model (Gauch, 1988; Perkins, 1972; Zobel et al., 1988). Description of the interaction in this model requires a smaller number of degrees of freedom and is more accurate than in the first model. Cultivar scores represent sensitivities and location scores can be interpreted as valuations of the environmental circumstances. After the analysis, scores of cultivars and locations can be related to additional information, in the present study the virus concentrations, to improve interpretation. The third model is a factorial regression model, a two-way ANOVA model with concomitant variables on both the cultivar and location factor (Denis, 1980, 1988; Snedecor & Cochran, 1980). Virus concentrations were directly incorporated in this model.

Results of these studies were expected to provide insight in the use of statistical models for the analysis of field experiments for resistance to BNYVV. Furthermore, explanation of cultivar by location interaction by means of virus concentrations would indicate that the relative performance of susceptible and partially resistant sugar beet cultivars in BNYVV-infested fields could be estimated using information on the resistance levels of the cultivars, the performance in non-infested fields, and the infection levels in the fields (Fig. 1).

Materials and methods

Field experiments

In two consecutive years, cultivar trials for resistance to BNYVV, each carried out in a completely randomised block design, were carried out at vari-

ous locations. Data were obtained of the yield parameters root yield (ton ha^{-1}), sugar content (%), sugar yield (ton ha^{-1}), and the quality parameters sodium (Na), potassium (K) and α-amino nitrogen (α-amino N) (mmol/100 g beet) (De Nie & Van de Poel, 1989).

In 1989, the field trials were located at Colijnsplaat, Nagele, Veere and St. Maartensdijk, and included the susceptible cultivars 'Accord' and 'Univers' and the partially resistant cultivars 'Rima', 'Rizofort', 'Rizo 91'. 'M 8906', 'HM 5682', 'Donna', 'Roxane' and 'Samba'. The trial at Colijnsplaat was located on a field considered to be free of infestation with BNYVV, and was carried out in four replicates with a plot size of 7 × 3 m. The complete plots were harvested. The other trials were located on fields infested with BNYVV and were performed in five replicates. Gross size of the plots on the infested fields varied from 25 to 27 × 3 m. Net plot size, used for harvesting, varied from 18 to 20 × 3 m.

The field trials in 1990 were located at Ovezande, Wieringerwerf, Lage Zwaluwe, Nagele (two fields, coded I and II) and Arnemuiden, and included the susceptible cultivars 'Accord', 'Univers' and 'Regina' and the partially resistant cultivars 'Rima', 'Rizofort', 'Rizo 92', 'M 8917', 'Donna', 'Roxane' and 'Samba 2'. The fields at Ovezande and Wieringerwerf were considered to be disease free, whereas the other trials were located in fields infested with BNYVV. All trials were performed in four replicates with a plot size of 7 × 3 m. The complete plots were harvested.

Sampling for virus assays

At all fields in both years, a susceptible cultivar was sampled to characterise the level of infection with BNYVV. The cultivar 'Accord' was sampled in most fields, but the cultivar 'Lucy' was taken at Colijnsplaat in 1989 and the cultivar 'Univers' at Wieringerwerf in 1990. At fields that were considered to be free of infestation with BNYVV, 20 plants were sampled, whereas at the other fields 10 plants were taken per replicate. At St. Maartensdijk in 1989 and at Nagele II in 1990 the other cultivars were also sampled, taking 10 plants per replicate, to assess the

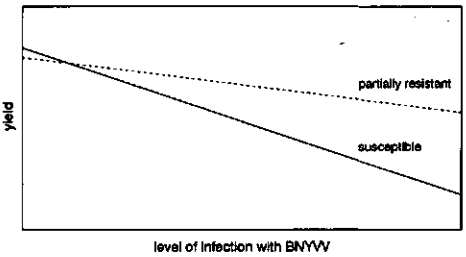


Fig. 1. Example of the occurrence of cultivar by location interaction in field trials for resistance to BNYVV with a susceptible and partially resistant sugar beet cultivar (after Richard-Molard, 1987); yield in ton ha^{-1}, and infection expressed as virus concentration (ng ml^{-1}) in samples of a susceptible indicator cultivar.

level of resistance. In 1989, samples were taken from the gross strips of the plots, or from the border rows of the field. These samples were taken in July and August. In 1990, samples were taken from special sample plots that were located next to the harvest plots, or from border rows of the field. Samples were taken in June and July. Additional samples from all cultivars were taken at Nagele II in August.

To assay the virus content, sap was extracted from the main root of the plants near the tip of the beets, using a garlic press. The sap from 10 plants of each replicate was combined into one sample. The 20 plants, taken from fields considered disease free, were analysed individually. The sap was diluted with phosphate buffered saline, containing 0.05% Tween 20 (PBS-Tween) in a ratio 1:9 (v v^{-1}).

Greenhouse experiments

In both years, the cultivars used were also tested for resistance to BNYVV in the greenhouse, using the same seed lot as was used in the field. Thirty seedlings of each cultivar were grown in a mixture of sand and soil infested with BNYVV (ratio 9:1 (v v^{-1})) for a period of five weeks at a temperature of 22/17° C (10 h/14 h). A sample of 100 mg lateral root material from each plant was crushed with PBS-Tween in a ratio 1:20 (w v^{-1}) (Paul et al., 1992).

Virus assays

Enzyme-linked immunosorbent assay (ELISA) (Clark & Adams, 1977) was used for virus detection. ELISA was performed as described by Alderlieste & Van Eeuwijk (1992) and Paul et al. (1992), with incubation times and temperatures as described by Büttner & Bürcky (1987), and using a dose response curve of purified virus, modelled by a four parameter logistic model, for the conversion of ELISA readings to virus concentrations. Virus standards were diluted in a 0.05 (v v^{-1}) or 0.1 (v v^{-1}) solution of healthy plant sap of cultivar 'Regina' with PBS-Tween. Samples with a value below 4 ng ml^{-1} were considered to be free of virus. Results of virus assays were presented as $\log_{10}$ of virus concentrations in ng ml^{-1}. Zero concentrations were given the value of 1 ng ml^{-1} before the $\log_{10}$ conversion.

Statistical analysis

For each yield and quality parameter, data were arranged in a cultivar by location table of means over replicates. Because the cultivars were not the same and different seed lots were used in the two years, results of the two years were analysed separately. Firstly, as a kind of base-line model, a two-way ANOVA model with interaction was fitted to each table:

$$\varepsilon(Y_{ij}) = \mu + G_i + E_j + GE_{ij} \qquad (1)$$

In (1) $\varepsilon(Y_{ij})$ stands for the expectation of the variable Y for the i-th cultivar at the j-th location. The term μ denotes the general mean, G_i ($i = 1 \ldots I$) the cultivar main effect, E_j ($j = 1 \ldots J$) the location main effect, and GE_{ij} the cultivar by location interaction having $(I\text{-}1)(J\text{-}1)$ degrees of freedom. The mean within block error, that is the mean of the errors obtained from the complete block analyses per location, divided by the number of replicates per location, was used to test the main effects of cultivar and location and their interaction. This error was also used for the computation of LSDs for cultivars and locations.

Secondly, AMMI-models were fitted:

$$\varepsilon(Y_{ij}) = \mu + G_i + E_j + \underbrace{\sum_{n=1}^{N} \sigma_n u_{ni} v_{nj} + \delta_{ij}}_{GE_{ij}} \qquad (2)$$

Here, μ, G_i and E_j have the same meaning as in (1). The term GE_{ij} of (1) is now split into a model part, $\sum_{n=1}^{N} \sigma_n u_{ni} v_{nj}$ and a residual, δ_{ij}. For orthogonal tables, with no cells missing, the least squares estimates for μ, G_i and E_j are the usual two-way analysis of variance estimates, whereas the multiplicative terms for cultivars, u_{ni}, and for locations, v_{nj}, also called scores, and the scaling constant σ_n are obtained from the singular value decomposition (Gabriel, 1978) of the two-way table of means corrected for the cultivar and location effect, containing the so-called residuals from additivity. The scalar constant σ_n, the singular value for the n-th set of product terms, indicates the importance of that set for the description of the interaction. Its square is equal to the sum of squares explained by the set. A corresponding mean square can be obtained by dividing this sum of squares by $(I\text{-}1) + (J\text{-}1) + 1 - 2n$, its degrees of freedom (Gollob, 1968). N indicates the number of sets necessary for an adequate description of the interaction and was assessed by testing the mean squares for successive terms against the mean within block error. After the analysis, main effects and interaction scores for cultivars and locations of the AMMI model were correlated with virus concentrations of the ten cultivars and the susceptible indicator cultivar to facilitate interpretation.

Thirdly, direct modelling of the additional information took place in factorial regression models. Virus concentrations of the cultivars that were determined in the greenhouse were used as concomitant variable on the cultivar factor, and field virus concentrations estimated as the mean virus concentration of the two samples from the susceptible indicator cultivar were used as concomitant variable on the location factor. The model used in the present study was:

$$\varepsilon(Y_{ij}) = \mu + \underbrace{\alpha_i + \xi x_i}_{G_i} + \underbrace{\beta_j + \zeta z_j}_{E_j} + \underbrace{\eta x_i z_j + \rho_i z_j + x_i \tau_j + \delta_{ij}}_{GE_{ij}} \qquad (3)$$

The term μ again stands for the general mean. For the cultivars the concomitant greenhouse virus concentrations are denoted by x_i, for the locations the concomitant field virus concentrations are denoted by z_j. Both variables were centred. G_i from (1) and (2) is replaced by a regression through the origin of the cultivar main effect on x_i, ξ represents the slope and the α_i values reflect deviations from this regression. E_j is likewise replaced by ζ and β_j values. The interaction is decomposed into three model parts and a residual part. The first model part, $\eta\, x_i z_j$, can be thought of as a regression through the origin of the residuals from additivity on an explanatory variable consisting of the product of x_i and z_j, with slope η, with one degree of freedom only. The second and third part of the interaction represent additional regressions per cultivar on z_j, giving slopes ρ_i and with I-2 degrees of freedom, and per location on x_i, giving slopes τ_j and J-2 degrees of freedom. The term δ_{ij} again denotes a residual.

Results

Field trials and virus assays

Average yield and quality data per cultivar and per location are given in Table 1 and 2, together with the results of the virus assays. In 1989, virus concentrations in plants from the field were lower at the second than at the first sampling date. In 1990, variations between sampling dates were found for the vi-

Table 1. Yield and quality data of ten sugar beet cultivars, averaged over locations, in 1989 and 1990, and virus concentrations in roots of beet plants on a heavily infested field and in the greenhouse

Cultivars	*Root yield* (ton ha^{-1})	*Sugar* content (%)	*Sugar yield* (ton ha^{-1})	*K*	*Na*	*α-amino N*	*Virus concentrations*			
				(mmol/100 g beet)			($\log_{10}$, ng ml^{-1})			
1989							St. Maartensdijk			Greenhouse
							July	August		
Accord	44.5	12.2	5.93	5.19	1.17	1.02	2.44	1.99		2.37
Univers	50.6	12.9	6.79	4.85	1.04	1.19	2.07	1.85		2.31
Rima	55.1	14.8	8.25	4.89	0.53	1.61	2.00	1.24		1.53
Rizofort	58.6	14.9	8.74	5.01	0.50	1.69	1.54	0.95		1.69
Rizo 91	52.1	15.4	8.07	4.92	0.42	1.56	1.59	1.58		1.56
M 8906	55.0	13.3	7.47	4.77	0.83	1.37	1.82	1.69		2.10
HM 5682	49.1	14.1	7.06	4.97	0.70	1.34	2.32	1.89		2.07
Donna	53.2	14.7	7.88	4.07	0.54	1.13	1.22	1.50		1.99
Roxane	53.0	14.2	7.61	5.01	0.72	1.35	2.25	1.71		2.02
Samba	54.7	14.6	8.01	4.97	0.55	1.58	2.00	1.33		1.57
LSD (95%)	2.1	0.2	0.33	0.16	0.07	0.10	0.62	0.46		0.27
1990							Nagele II			Greenhouse
							June	July	August	
Accord	65.1	15.1	9.96	4.84	0.84	1.51	1.55	1.46	1.82	2.51
Univers	69.1	14.8	10.29	4.77	0.85	1.89	1.02	1.96	2.08	2.40
Regina	66.5	14.7	9.84	4.62	0.87	1.63	0.82	1.96	1.88	2.49
Rima	69.4	16.1	11.14	4.98	0.39	2.37	0.51	0.33	1.08	1.87
Rizofort	68.6	16.0	10.95	5.03	0.41	2.24	1.08	1.13	0.95	1.91
Rizo 92	65.4	15.8	10.31	4.85	0.35	2.18	0.08	0.78	0.71	1.72
M 8917	74.5	14.9	11.07	4.73	0.79	1.82	0.41	1.44	1.78	2.31
Donna	62.0	16.3	10.06	4.14	0.42	1.61	1.23	0.91	1.47	2.17
Roxane	69.9	16.6	11.56	5.05	0.44	2.10	1.39	1.29	1.24	1.68
Samba 2	67.6	15.8	10.63	4.96	0.40	2.19	0.52	0.93	1.13	1.71
LSD (95%)	2.7	0.2	0.47	0.11	0.06	0.15	1.16	1.09	0.68	0.22

rus concentrations in field material, but in general, virus concentrations increased with time. Experimental errors were greater for the results of virus assays of plants from the field than from the greenhouse (Table 1). Fields were arranged according to increasing levels of infection with BNYVV in the susceptible cultivar (Table 2). In 1989, a few positive samples were found at Colijnsplaat, however, the average virus concentration was below the level indicating the presence of the virus. In 1990, no virus was detected at Ovezande and Wieringerwerf. Virus could be detected in all other fields, except for Lage Zwaluwe at the first sampling date. Average virus concentrations of the samples from the different fields were higher in 1989 than in 1990.

It may be noticed, that highest yields were not necessarily found on fields without infection.

ANOVA

Table 3 shows the results of the ANOVA. Main effects and interaction were significant for all yield and quality parameters in both years ($P < 0.05$).

AMMI analysis

In the Tables 4 and 5, the results of AMMI analyses are given for the parameter sugar yield in 1989 and 1990. In both years, mean squares corresponding to the first multiplicative term were highly significant when tested against the mean within block error. One set of interaction parameters ($\sigma_p u_{ip} v_{jp}$) sufficed for an adequate description, as can be seen from the high percentage of explained interaction and the non-significant residual (Table 4). In Table 6 results of the AMMI analysis with one set of interaction parameters are presented for all parameters, together with the results of the factorial regression. High percentages of explained interaction were found for all yield and quality parameters.

In Table 7, correlation coefficients of virus concentrations of plants from the field and the greenhouse with cultivar main effects and cultivar interaction scores of the different parameters are presented. Virus concentrations of the plants sampled in August gave higher correlations with both cultivar main effects and scores for interaction than virus concentrations of plants sampled in June or July. In general, virus concentrations of plants from the greenhouse gave higher correlations than virus con-

Table 2. Yield and quality data on four locations in 1989 and six locations in 1990, averaged over cultivars, and virus concentrations in the roots of a susceptible indicator cultivar

Locations	Root yield (ton ha^{-1})	Sugar content (%)	Sugar yield (ton ha^{-1})	K	Na	α-amino N	Virus concentrations
				(mmol/100 g beet)			($\log_{10}$ ng ml^{-1})
1989							July August
Colijnsplaat	66.1	15.7	10.38	3.77	0.37	1.84	0.28 0.23
Nagele	73.8	14.1	10.39	4.47	0.65	1.28	0.89 0.63
Veere	32.4	14.1	4.71	5.72	0.99	1.57	- 2.02
St. Maartensdijk	38.1	12.5	4.85	5.51	0.78	0.85	2.44 1.99
LSD (95%)	3.3	0.4	0.52	0.26	0.11	0.16	
1990							June July
Ovezande	58.2	16.1	9.41	4.36	0.34	3.30	0.02 0.00
Wieringerwerf	71.3	17.5	12.49	5.97	0.30	2.07	0.00 0.00
Lage Zwaluwe	66.7	16.2	10.77	4.15	0.55	1.63	0.12 0.62
Nagele I	82.5	14.2	11.84	4.48	0.84	1.72	0.61 0.80
Nagele II	69.1	14.5	10.02	5.03	0.83	1.56	1.55 1.46
Arnemuiden	59.1	15.1	8.96	4.78	0.58	1.45	1.94 2.25
LSD (95%)	3.5	0.3	0.60	0.14	0.08	0.19	

Table 3. Analysis of variance of yield and quality data in 1989 and 1990

Source of variation	Degrees of freedom	Root yield		Sugar content		Sugar yield		K		Na		α-amino N	
		SS[1]	MS[1]	SS	MS	SS	MS	SS	MS	SS	MS	SS	MS
1989													
Cultivar	9	540.7	60.08	37.43	4.16	23.69	2.63	3.29	0.37	2.23	0.25	1.81	0.20
Location	3	12520.0	4173.32	51.35	17.12	314.19	104.73	24.94	8.31	2.00	0.67	5.37	1.79
Cultivar × location	27	949.5	35.17	14.96	0.55	23.92	0.89	3.72	0.14	1.18	0.04	0.92	0.03
Mean within block error	135		5.50		0.07		0.14		0.03		0.006		0.01
1990													
Cultivar	9	617.5	68.61	25.90	2.88	17.97	2.00	3.87	0.43	2.77	0.31	4.91	0.55
Location	5	3996.4	799.29	78.96	15.79	95.50	19.10	21.31	4.26	2.67	0.53	23.88	4.78
Cultivar × location	45	1410.9	31.35	12.41	0.28	54.15	1.20	1.62	0.04	0.93	0.02	3.18	0.07
Mean within block error	182		9.41		0.07		0.28		0.02		0.005		0.03

[1] SS = Sum of squares; MS = Mean squares. All effects were significant at the 5% level.

centrations of plants from the field. Correlation coefficients of virus concentrations of a susceptible indicator cultivar with location main effects and location interaction scores are presented in Table 8. For the sampling dates in each year, correlations of the virus concentrations of the susceptible indicator cultivar with the location main effects and interaction scores were similar (Table 8).

For the parameters root yield, sugar content, sugar yield and α-amino nitrogen, cultivar main effects were negatively correlated to virus concentration, although results were not clear for root yield in 1990

Table 4. Analysis of variance for sugar yield in 1989 and 1990, with subdivision of sum of squares by two multiplicative models for cultivar by location interaction

Source of variation	Degrees of freedom	Sum of squares	Percentage explained	Mean square	Degrees of freedom	Sum of squares	Percentage explained	Mean square
	1989				1990			
Cultivar	9	23.69		2.63	9	17.97		2.00
[2] ξ	1	17.71	75	17.71	1	7.02	40	7.02
deviations ($\alpha_i s$)	8	5.98	25	0.75	8	10.95	60	1.37
Location	3	314.19		104.73	5	95.50		19.10
[2] ζ	1	296.34	94	296.34	1	29.71	31	29.71
deviations ($\beta_j s$)	2	17.86	6	8.93	4	65.79	69	16.45
Cultivar × location	27	23.92		0.89	45	54.15		1.20
[1] $\sigma_j u_{i1} v_{1j}$	11	21.39	89	1.94	13	48.40	89	3.72
deviations (δ_{ij})	16	2.53	11	0.16	32	5.75	11	0.18
[2] η	1	11.85	50	11.85	1	35.71	66	35.71
ρ_i	8	6.05	25	0.76	8	7.57	14	0.95
τ_j	2	3.41	14	1.71	4	5.89	11	1.47
deviations (δ_{ij})	16	2.60	11	0.16	32	4.98	9	0.16
Mean within block error	135			0.14	182			0.28

[1] Subdivision by singular value decomposition of residuals from additivity (AMMI model); [2] Subdivision by factorial regression.

Table 3. Analysis of variance of yield and quality data in 1989 and 1990

Source of variation	Degrees of freedom	Root yield		Sugar content		Sugar yield		K		Na		α-amino N	
		SS[1]	MS[1]	SS	MS	SS	MS	SS	MS	SS	MS	SS	MS
1989													
Cultivar	9	540.7	60.08	37.43	4.16	23.69	2.63	3.29	0.37	2.23	0.25	1.81	0.20
Location	3	12520.0	4173.32	51.35	17.12	314.19	104.73	24.94	8.31	2.00	0.67	5.37	1.79
Cultivar × location	27	949.5	35.17	14.96	0.55	23.92	0.89	3.72	0.14	1.18	0.04	0.92	0.03
Mean within block error	135		5.50		0.07		0.14		0.03		0.006		0.01
1990													
Cultivar	9	617.5	68.61	25.90	2.88	17.97	2.00	3.87	0.43	2.77	0.31	4.91	0.55
Location	5	3996.4	799.29	78.96	15.79	95.50	19.10	21.31	4.26	2.67	0.53	23.88	4.78
Cultivar × location	45	1410.9	31.35	12.41	0.28	54.15	1.20	1.62	0.04	0.93	0.02	3.18	0.07
Mean within block error	182		9.41		0.07		0.28		0.02		0.005		0.03

[1] SS = Sum of squares; MS = Mean squares. All effects were significant at the 5% level.

centrations of plants from the field. Correlation coefficients of virus concentrations of a susceptible indicator cultivar with location main effects and location interaction scores are presented in Table 8. For the sampling dates in each year, correlations of the virus concentrations of the susceptible indicator cultivar with the location main effects and interaction scores were similar (Table 8).

For the parameters root yield, sugar content, sugar yield and α-amino nitrogen, cultivar main effects were negatively correlated to virus concentration, although results were not clear for root yield in 1990

Table 4. Analysis of variance for sugar yield in 1989 and 1990, with subdivision of sum of squares by two multiplicative models for cultivar by location interaction

Source of variation	Degrees of freedom	Sum of squares	Percentage explained	Mean square	Degrees of freedom	Sum of squares	Percentage explained	Mean square
	1989				1990			
Cultivar	9	23.69		2.63	9	17.97		2.00
[2] ξ	1	17.71	75	17.71	1	7.02	40	7.02
deviations (α_i s)	8	5.98	25	0.75	8	10.95	60	1.37
Location	3	314.19		104.73	5	95.50		19.10
[2] ζ	1	296.34	94	296.34	1	29.71	31	29.71
deviations (β_j s)	2	17.86	6	8.93	4	65.79	69	16.45
Cultivar × location	27	23.92		0.89	45	54.15		1.20
[1] $\sigma_i u_{ij} v_{ij}$	11	21.39	89	1.94	13	48.40	89	3.72
deviations (δ_{ij})	16	2.53	11	0.16	32	5.75	11	0.18
[2] η	1	11.85	50	11.85	1	35.71	66	35.71
ρ_i	8	6.05	25	0.76	8	7.57	14	0.95
τ_j	2	3.41	14	1.71	4	5.89	11	1.47
deviations (δ_{ij})	16	2.60	11	0.16	32	4.98	9	0.16
Mean within block error	135			0.14	182			0.28

[1] Subdivision by singular value decomposition of residuals from additivity (AMMI model); [2] Subdivision by factorial regression.

Table 7. Correlation coefficients between virus concentrations of sugar beets and cultivar main effects (M) and scores for interaction (I) in 1989 and 1990

Plant material	Root yield		Sugar content		Sugar yield		K		Na		α-amino N	
	M	I	M	I	M	I	M	I	M	I	M	I
1989[1]												
Field July	−0.61	−0.52	−0.59	−0.59	−0.67	−0.54	0.73	−0.80	0.64	−0.57	−0.27	−0.47
August	−0.87	−0.70	−0.70	−0.85	−0.91	−0.78	0.12	−0.77	0.74	−0.76	−0.80	−0.64
Greenhouse	−0.67	−0.82	−0.89	−0.91	−0.87	−0.84	0.00	−0.79	0.90	−0.88	−0.89	−0.85
1990[1]												
Field June	−0.29	−0.36	0.11	−0.25	−0.11	−0.30	−0.13	−0.19	0.26	−0.22	−0.49	−0.16
July	0.15	−0.80	−0.74	−0.71	−0.38	−0.77	−0.15	−0.87	0.85	−0.78	−0.62	−0.92
August	0.13	−0.85	−0.76	−0.82	−0.42	−0.84	−0.39	−0.90	0.93	−0.89	−0.77	−0.86
Greenhouse	−0.06	−0.92	−0.81	−0.90	−0.63	−0.93	−0.48	−0.84	0.91	−0.89	−0.84	−0.79

[1] Absolute values higher than 0.63 are significant at the 5% level.

Factorial regression analysis

Results of the factorial regression analysis for the parameter sugar yield are presented in Table 4. The percentage of the main effects' sums of squares explained by the concomitant variables can be obtained by squaring the correlation coefficients of the Tables 7 and 8. For the factorial regressions, greenhouse virus concentrations were used as a concomitant variable for the cultivars and the means over both assessments of the field infection for the locations. For sugar yield, the interaction could be described for 50% in 1989 and 66% in 1990 by the regression of the residuals from additivity on the variable formed by the product of the cultivar greenhouse virus concentrations and field virus concentrations, with slope η. This regression can be interpreted as reflecting a downward correction for sugar yield for susceptible cultivars in fields with a high level of infection and an upward correction in fields without or with a low level of infection, which corresponds to the AMMI analysis. Additional regressions for each cultivar on the field virus concentrations, with slopes ρ_i, and regressions for each location on the cultivar greenhouse virus concentrations, with slopes τ_j, also accounted for a significant

Table 8. Correlation coefficients between virus concentrations of a susceptible indicator cultivar and location main effects (M) and scores for interaction (I) in 1989 and 1990

Plant material	Root yield		Sugar content		Sugar yield		K		Na		α-amino N	
	M	I	M	I	M	I	M	I	M	I	M	I
1989												
Field July[1]	−0.89	1.00	−0.97	0.84	−0.96	0.91	0.99	−0.88	0.90	−0.87	−0.95	0.24
August[2]	−0.94	0.99	−0.78	0.86	−0.98	0.92	0.99	−0.92	0.90	−0.83	−0.56	0.42
Mean[2]	−0.91	0.97	−0.84	0.83	−0.97	0.91	0.98	−0.88	0.88	−0.79	−0.64	0.41
1990												
Field June[3]	−0.17	0.93	−0.68	0.80	−0.56	0.94	0.02	0.20	0.58	−0.67	−0.60	0.67
July[3]	−0.17	0.92	−0.66	0.83	−0.55	0.93	−0.12	0.19	0.58	−0.70	−0.70	0.73
Mean[3]	−0.17	0.93	−0.67	0.82	−0.56	0.94	−0.05	0.20	0.59	−0.69	−0.66	0.71

[1,2,3]: Absolute values higher than 1.00, 0.95, and 0.81 are significant at the 5% level, respectively.

part of the interactions. Deviations from these regressions were not significant. The three regressions, together constituting the model for interaction in the factorial regression model, accounted for 89% of the interaction in sugar yield in 1989 and 91% in 1990. These percentages are about the same as those for the AMMI analysis (Table 6). For the other parameters AMMI analysis and factorial regression analysis gave similar results too. The number of degrees of freedom (n = 11 in 1989 and n = 13 in 1990) for the factorial regression description of interaction was equal to that of the AMMI model. The two models provided more or less equivalent descriptions of the interaction and, as for the AMMI analysis and correlation studies, results of the factorial regression analysis support the conclusion that measured virus concentrations of the cultivars and the susceptible indicator cultivar can be used to describe the observed interactions.

Discussion

Main effects

In the present study, a decrease in root yield, sugar content, sugar yield and α-amino nitrogen and an increase in sodium was found with increasing susceptibility of the cultivars and increasing levels of infection in the fields, while trends for potassium were not clear. Similar results for the cultivars were described earlier (Ahrens, 1987; Bürcky & Büttner, 1989b, 1991; Giunchedi et al., 1987; Hillmann, 1984; Shimada, 1989). The results also correspond with the numerous reports on cultivar trials that were performed in various countries. At a low infection level, as in 1990, correlations between virus concentrations and yield parameters were less evident than at a high infection level as in 1989. A similar finding was reported by Ahrens (1987). However, for sugar content, sodium and α-amino nitrogen high correlations with virus concentrations were found in 1990, indicating that these parameters were more sensitive to BNYVV than the other parameters. Similar observations were made by Heijbroek (1989), Pollach (1984) and Takeda et al. (1988) for susceptible cultivars, while findings in the present study on the reaction of yield and quality parameters in resistant and susceptible cultivars to infection with BNYVV confirm the results of Bürcky & Büttner (1989a), Casarini-Camangi & Canova (1987) and Rosso et al. (1988, 1989).

Interaction effects

In both years, cultivar by location interactions were significant for all yield and quality parameters ($P < 0.05$). AMMI analyses and factorial regressions demonstrated the feasibility of a description of the interaction in terms of virus concentrations, representing resistance level of the cultivars and infection level in the fields. Giunchedi et al. (1987) did not find interaction for most parameters, indicating that their fields had identical levels of infection. However, in most studies which include fields with various infection levels, interactions will occur and the use of models with terms for interaction is inevitable.

Statistical models and additional information

Virus concentration of the cultivars measured in the greenhouse was a better explanatory variable than virus concentration measured in the field. Greenhouse experiments have lower residual variance than field experiments. For the evaluation of the level of resistance, greenhouse tests provide the most accurate information. Variation in virus concentrations of plants from the field was also reported by Casarini-Camangi & Canova (1988). For the assessment of the infection level in the field, reflecting the level of infestation, a susceptible indicator cultivar was used. As an alternative, infestation levels can be determined by estimates of the number of infectious units, assessed by the most probable number technique as done by Tuitert (1990).

An AMMI model can be used even in the absence of additional information on cultivars and locations. Extra information, such as results of greenhouse tests and characterisation of the level of infection in the fields by sampling of susceptible cultivars, allows a better interpretation of AMMI

results. Alternatively, factorial regression models use extra information directly. For the experiments described here, virus measurements greatly improved the quality of the descriptions of cultivar by location interaction.

Cultivar evaluation

In breeding programmes and cultivar trials, field experiments for resistance to BNYVV are carried out on several locations with various levels of infection, including non-infested fields (Richard-Molard, 1987). The application of statistical models that include terms for cultivar by location interaction helps to interpret the results of such studies. In the present study, virus concentrations of the cultivars in a greenhouse test, and of a susceptible cultivar in the field, gave a satisfactory explanation for cultivar by location interactions in field trials for resistance. This leads to the conclusion, that the relative performance of susceptible and partially resistant cultivars, in infested fields, can be estimated by means of three independent parameters, (i) the level of resistance determined in a greenhouse experiment, (ii) the yield and quality in non-infested fields and (iii) the level of infection in the field. This approach will reduce the need for testing cultivars and breeding accessions on a large number of infested fields, and thus will increase the efficiency of sugar beet breeding and cultivar evaluation.

Acknowledgements

We thank B. Henken and L.C.P. Keizer for their assistance and Dr J. Hoogendoorn, Dr W. Lange, Dr J.B. Denis (Laboratoire de Biométrie, I.N.R.A., Versailles), Dr D. Peters (Department of Virology, Wageningen Agricultural University) and Dr J.C. Zadoks (Department of Phytopathology, Wageningen Agricultural University) for critical reading of the manuscript.

References

Ahrens, W., 1987. Auslese Rizomania-resistenter Zuckerrüben mit Hilfe des ELISA-Tests. Gesunde Pflanzen 39: 107–112.

Alderlieste, M.F.J. & F.A. Van Eeuwijk, 1992. Assessment of concentrations of beet necrotic yellow vein virus (BNYVV) by enzyme-linked immunosorbent assay. J. Virol. Meth. 37: 163–176.

Bürcky, K. & G. Büttner, 1989a. Rizomaniatolerante Zuckerrübensorten I. Streuung von Gewicht und Qualitätsmerkmalen für Einzelrüben verschiedener Sorten mit und ohne Befall durch BNYVV. Zuckerindustrie 114: 142–148.

Bürcky, K. & G. Büttner, 1989b. Rizomaniatolerante Zuckerrübensorten II. BNYVV-Gehalt junger Testpflanzen und Ertrag im Feld unter Rizomaniabefall. Zuckerindustrie 114: 225–228.

Bürcky, K. & G. Büttner, 1991. Gehalt an Beet Necrotic Yellow Vein Virus (BNYVV) in der Hauptwurzel von Zuckerrübenpflanzen verschiedener Sorten und deren Leistung unter Rizomaniabefall im Feld. J. Phytopath. 131: 1–10.

Büttner G. & K. Bürcky, 1987. Verbesserung der Nachweisempfindlichkeit von *ELISA beim Test auf BNYVV. Proc. 50th* Winter Congress IIRB, Brussels: 289–294.

Casarini-Camangi, P. & A. Canova, 1987. Infezione ed incidenza sulla produttività del virus della Rizomania in alcuni genotipi di barbabietola da zucchero. Phytopath. Medit. 26: 91–99.

Casarini-Camangi, P. & A. Canova, 1988. Aspetti epidemiologici ed infeziosità della rizomania. Phytopath. Medit. 27: 73–83.

Clark, M.F. & A.N. Adams, 1977. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. J. Gen. Virology 34: 475–483.

De Nie, L.H. & P.W. Van de Poel, 1989. Uniforme methode voor gewichtsbepaling, monstername en monsteronderzoek van suikerbieten in Nederland. Uitgave van de Nederlandse Suikerindustrie, 26 pp.

Denis, J.B., 1980. Analyse de régression factorielle. Biom. Praxim. 20: 1–34.

Denis, J.B., 1988. Two way analysis using covariates. Statistics 19: 123–132.

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

Gauch, H.G. Jr., 1988. Model selection and validation for yield trials with interaction. Biometrics 44: 705–715.

Giunchedi, L., M. De Biaggi & C. Poggi Pollini, 1987. Correlation between tolerance and beet necrotic yellow vein virus in sugar-beet genotypes. Phytopath. Medit. 26: 23–28.

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

Heijbroek, W., 1989. The development of rhizomania in two areas of the Netherlands and its effect on sugar-beet growth and quality. Neth. J. Plant Path. 95: 27–35.

Hillmann, U., 1984. Neue Erkenntnisse über die Rizomania an Zuckerrüben mit besonderer Berücksichtigung Bayerischer Anbaugebiete. Dissertation Justig Liebig Universität Giessen, 226 pp.

Paul, H., B. Henken & M.F.J. Alderlieste, 1992. A greenhouse test for screening sugar-beet (*Beta vulgaris*) for resistance to beet necrotic yellow vein virus (BNYVV). Neth. J. Plant Path. 98: 65–75.

Perkins, J.M., 1972. The principal components analysis of genotype-environmental interactions and physical measures of the environment. Heredity 29: 51–70.

Pollach, G., 1984. Versuche zur Verbesserung einer Rizomania-Diagnose auf Basis konventioneller Rübenanalysen. Zuckerindustrie 109: 849–853.

Richard-Molard, M., 1987. Variétés tolérantes à la rhizomanie – Résultats d'essais, problèmes posés par leur utilisation. Proc. 50th Winter Congress IIRB, Brussels: 111–129.

Rosso, F., M. Bimbatti & P. Meriggi, 1988. Effetti della Rizomania sulla evoluzione dei parametri tecnici della produzione. Sementi Elette 34 (1/2): 41–49.

Rosso, F., P. Meriggi, G. Vaccari & G. Mantovani, 1989. Ulteriori studi su varietà di barbabietole sensibili e tolleranti alla rizomania. Sementi Elette 35 (6): 3–14.

Shimada, I., A. Shirahata, K. Fujii, N. Tomiyama & H. Sugawara, 1989. Study on selection of rhizomania tolerant sugar beet varieties. Proc. Japan. Soc. Sugar Beet Technol. 31: 93–97.

Snedecor, G.W. & W.G. Cochran, 1980. Statistical methods. Iowa State University Press, Ames, 7th Edition, 507 pp. (Chapter 16).

Takeda, S., H. Dempo, M. Yasumura & M. Hayashida, 1988. Method to estimated the injuries by rhizomania disease on sugar beet by using the data of sugar content and harmful non-sugar components of sugar beet. Proc. Japan. Soc. Sugar Beet Technol. 30: 89–93.

Tuitert, G., 1990. Assessment of the inoculum potential of *Polymyxa betae* and beet necrotic yellow vein virus (BNYVV) in soil using the most probable number method. Neth. J. Plant Path. 96: 331–341.

Zobel, R.W., M.J. Wright & H.G. Gauch, 1988. Statistical analysis of a yield trial. Agron. J. 80: 388–393.

Chapter IV

Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass

F.A. van Eeuwijk & A. Elgersma
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

Heredity **70** (1993) 447–457

Received 23 June 1992

Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass

F. A. VAN EEUWIJK* & A. ELGERSMA†

DLO-Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 16, 6700 AA Wageningen, The Netherlands

Seed yield in perennial ryegrass was analysed for cultivar by environment interaction. Nine cultivars were evaluated in 12 trials at two locations over a 3-year period. Earlier attempts to describe the significant cultivar by environment interaction using a regression on the environmental mean or relationships with year, soil type, harvest method, or crop age, were unsuccessful. In this paper, therefore, meteorological data were introduced as explanatory variables. Three types of analysis were used. First, residuals from the cultivar by environment two-way table corrected for main effects were regressed on the explanatory variables for each cultivar separately. Secondly, the explanatory variables were used as concomitant variables for the environmental factor in a two-way analysis of variance of genotypes by environments. Finally, the matrix of residuals from additivity was subjected to a singular value decomposition, after which environmental scores were related to values of the explanatory variables using regression and a recently developed method to calculate confidence intervals for scores. All methods led to comparable conclusions about the importance of different variables in the interaction. Of equal importance were minimum temperature in the period before ear emergence, temperature sum in the period from the beginning of anthesis until peak anthesis, and mean and maximum temperature in the period from the end of anthesis until harvest. The major component of interaction was identified as a contrast between early and late cultivars. A minor component was due to cultivars that performed relatively well in the worst environment and relatively badly in the best environment. The usefulness of so-called AMMI models is discussed and compared with that of the more traditional regression on the environmental mean model.

Keywords: AMMI analysis, confidence intervals, environmental variables, factorial regression, genotype by environment interaction, perennial ryegrass.

Introduction

Perennial ryegrass is an important grass species that is propagated by seed. Seed yields are typically low. In a previous study on the seed yields of nine cultivars in 12 trials a significant interaction between cultivars and trials was found ($P < 0.001$; Elgersma, 1990a). To model this interaction a regression on the environmental mean was tried initially (Yates & Cochran, 1938; Mandel, 1961; Finlay & Wilkinson, 1963). Results from this model were unsatisfactory, as only 14 per cent of the interaction sum of squares could be explained, which was not significant when tested against the deviations from regressions.

The observation that crop development rates were similar within years in the various trials, but differed among years, indicated that meteorological information might be useful in the clarification of the cultivar by trial interaction. Examples of incorporating physical measurements of the environment into models for genotype by environment interaction can be found in Abou-El-Fittouh *et al.* (1969), Fripp (1972), Hardwick

*Correspondence.
†Present address: Department of Field Crops and Grassland Science, Agricultural University, Haarweg 333, 6709 RZ Wageningen, The Netherlands.

& Wood (1972), Perkins (1972), Wood (1976), Denis (1980), Saeed & Francis (1984), Kang & Gorman (1989), and Gorman *et al.* (1989).

For perennial ryegrass Hampton & Hebblethwaite (1983) showed that minimum temperature around anthesis accounted for 70 per cent of the variation in seed numbers for cultivar S.24 over a period of 10 years. For our data it was already established that the environmental factors year, soil type, harvest method, and crop age affected the levels of seed yield, thousand-grain weight, and seed number in the 12 trials (Elgersma, 1990a, 1990b). We were, however, unable to relate these factors to the cultivar by trial interaction. In this paper we will investigate whether the interaction can be explained by meteorological variables. The illustration of the methodology will be as important as the results obtained. We will introduce a new and simple method of relating environmental information to genotype by environment interaction. It uses an expression derived by Goodman & Haberman (1990) for the confidence limits of genotypic and environmental parameters for the interaction in so-called AMMI models (Additive Main effects and Multiplicative Interaction effects models; Perkins, 1972; Gauch, 1988). The results of this method will be compared to those of more familiar methods.

Materials and methods

Trials, cultivars and measurements

Seed yield data were obtained from 12 experimental trials (Table 1) with nine perennial ryegrass cultivars sown at two experimental sites, one having sand and the other clay, in Wageningen, The Netherlands, and harvested in 1986, 1987 and 1988. Each trial consisted of a randomized blocks design with four replications. Trials were distinguished from each other by soil, year, production year of the crop (crop age), and harvest method (see also Elgersma, 1990a). The cultivars chosen were all late flowering, though significant differences occurred for maturity dates. In all trials the cultivar Perma (Pe) was the earliest, followed by Semperweide (Se), Wendy (We) and Parcour (Pa). Compas (Co) and Trani (Tr) were intermediate, Vigor (Vi) was rather late, and Barenza (Ba) and Lamora (La) were the latest. The difference between Perma and Barenza varied from 3 to 10 days in the various trials (Elgersma, 1990a, 1990b).

Table 1 Specification of trials

Number	Soil	Harvest year	Production year	Harvest method
1	Sand	1986	First	1
2	Clay	1986	First	1
3	Sand	1986	First	2
4	Clay	1986	First	2
5	Sand	1987	First	2
6	Sand	1987	Second	2
7	Clay	1987	Second	2
8	Sand	1987	Second	2
9	Clay	1987	Second	2
10	Sand	1988	Second	2
11	Sand	1988	First	2
12	Clay	1988	First	2

Dates of ear emergence, first anthesis, peak anthesis, end of anthesis, and harvest ripeness were recorded on each plot. Subsequently five developmental periods were defined: (1) 10 days preceding ear emergence; (2) ear emergence till first anthesis; (3) first anthesis till peak anthesis; (4) peak anthesis until end of anthesis; and (5) end of anthesis until harvest. Meteorological data were recorded within 6 km of the experimental plots (Haarweg observation station, unpublished data). During each developmental period minimum, mean and maximum temperature, rainfall, relative humidity and wind velocity were calculated for each plot (the choice of these variables was based on information in Hampton & Hebblethwaite, 1983). Environmental characterization of each site was derived from these data by averaging over all plots at a particular site within a particular year. Ranges of the meteorological variables over the trials are given in Table 2 for each developmental period. As can be deduced from Table 1 the following trials had identical environmental characterizations: 1 and 3; 2 and 4; 5, 6 and 8; 7 and 9; 10 and 11; whereas 12 was the only trial with a unique characterization. Additional variables included temperature sum (defined as the length of a developmental stage multiplied by the average temperature for that stage) and period length.

Statistical analyses

All analyses had as a starting point the matrix of interaction residuals, i.e. the cultivar by trial table corrected for main effects. First, interaction residuals were regressed for each cultivar separately on the weather variables including developmental period length. Orthogonalized squares of the weather variables were included as well. Secondly, a simultaneous regression of the interaction residuals on the weather variables was performed by introducing these variables as concomitant variables for the environmental factor in the two-way analysis of variance for the cultivar by trials table (Snedecor & Cochran, 1980, Chpt. 16). The

Table 2 Ranges of environmental variables during five developmental stages. For definition of the periods see text

	Period									
	1		2		3		4		5	
	min.	max.	min.	max.	min.	max.	min.	max.	min.	max.
Length period (days)	10	10	11	23	2	9	7	16	14	25
Mean temp. (°C)	12.3	15.2	13.4	19.3	15.1	22.7	14.9	18.7	14.6	17.7
Min. temp. (°C)	7.7	9.4	9.3	12.3	9.2	14.2	10.7	13.0	10.9	12.6
Max. temp. (°C)	16.5	19.6	17.3	24.7	19.2	32.1	19.0	23.7	18.3	27.6
Temp. sum (°C days)	123.0	152.0	174.2	324.3	45.4	154.7	104.3	280.3	247.8	402.5
Rel. hum. (%)	74.3	82.3	66.1	89.1	56.7	93.5	66.1	85.4	78.7	86.6
Rainfall (mm)	11.9	57.8	3.9	37.0	0.0	5.5	6.9	50.6	32.3	105.6
Wind velocity (km h^{-1})	2.5	3.4	2.3	3.0	1.3	2.4	1.9	2.6	2.4	3.1

interaction is partitioned into a part due to regression and a part due to deviations from regression. The method was probably introduced in plant breeding by Abou-El-Fittouh *et al.* (1969) and has been refined and extended by Denis (1980, 1988) under the name of factorial regression. Both these regression methods directly relate environmental information to interaction residuals.

Alternatively, one could try to first separate out pattern from noise in the interaction residuals by means of a singular value decomposition and subsequently relate the environmental scores thus obtained to measured environmental variables. Effectively, an additive main effects and multiplicative interaction effects (AMMI) model is used (Perkins, 1972; Gauch, 1988; Zobel *et al.*, 1988), which has the form

$$Y_{ijk} = \mu + G_i + E_j + \sum_{n=1}^{N} \lambda_n a_{ni} b_{nj} + I_{ij} + \varepsilon_{ijk}.$$

Y_{ijk} is the yield for the k-th replication of the i-thp cultivar in the j-th trial, μ the general mean, G_i and E_j are the cultivar and trial effect, respectively, λ_n is the n-th singular value from the singular value decomposition (Gabriel, 1978) of the matrix of interaction residuals, a_{ni} and b_{nj} are the corresponding cultivar and trial scores, N is the number of multiplicative terms (axes) needed for an adequate description of the interaction, I_{ij} is a residual arising from the two-way table after correction for the main effects and the extraction of the multiplicative interaction effects, and ε_{ijk} represents a normally distributed intra-block error.

Estimated environmental scores were regressed on the environmental variables (the same procedure can be used to relate cultivar scores to explanatory variables). As measurements for the environmental variables tended to cluster in two groups the approach by regression was questionable, because no real check on the linearity of the relation was possible. Therefore another method was used, which does not rely on the assumption of linearity and is based on the calculation of confidence intervals for cultivar and trial scores. The general expression for an interval for a multiplicative parameter (score), ξ, in an AMMI model was derived by Goodman & Haberman (1990) as $[\hat{\xi} - \mathrm{Tsq}(\hat{\xi}), \hat{\xi} + \mathrm{Tsq}(\hat{\xi})]$, in which T denotes the upper $\alpha/2$ point for a t-distribution, s is the square root of the variance estimate, and $q(\hat{\xi})$ is a function of the observations. (Those in favour of multiple comparison procedures can replace the t-distribution with their preferred distribution.) For the multiplicative parameter for the i-th row (cultivar) corresponding to the m-th singular value, a_{mi}, $q(\hat{\xi})$ is the square root of

$$\left[\left(1 - 1\Big/I - \sum_{n=1}^{N} \hat{a}_{ni}^2\right)\Big/\hat{\lambda}_m^2\right] + \sum_{n \neq m,\, n \leq N} [(\hat{\lambda}_m^2 + \hat{\lambda}_n^2)/(\hat{\lambda}_m^2 - \hat{\lambda}_n^2)^2]\, \hat{a}_{ni}^2,$$

where I is the number of rows of the two-way table, the number of cultivars. For the column parameters, b_{mj}, the same formula is valid with I replaced by J, the number of colomns (trials), and the a_{mi}s replaced by the b_{mj}s.

Before testing a hypothesis on the relation between the observed values of an environmental variable and the environmental scores for a particular axis, the values of the environmental variable must be scaled in

the same way as the environmental scores, e.g. with mean zero and squared length unity. Signs of observed values and scores must be aligned as much as possible, at least for the largest values. Testing involves comparing observed values with the confidence intervals for environmental scores.

Results and discussion

Two-way data table and ANOVA

Table 3 shows the cultivar by trial table of means in which each entry is a mean over four blocks. The corresponding ANOVA is given in Table 4. The sum of squares for cultivar × trial interaction looks small in comparison to the total sum of squares, but is substantial in comparison to the sum of squares for cultivars. The mean square is highly significant when tested against the mean intra-block error ($P < 0.001$). The interaction was not due to non-normality: the estimated value for the Box–Cox parameter, λ (see Atkinson, 1985), was close to 1. Nor could outliers be the source of the interaction as evidenced by a non-significant maximum normed residual of 0.2538 (Stefansky, 1972).

Regressing interaction residuals on environmental variables for each cultivar separately

Table 5 shows the results of the regressions of interaction residuals on weather variables for each cultivar. First, linear terms were tried. Because there were only six independent values for each explanatory variable, correlations had to exceed (+/−) 0.811 (4 d.f.) to achieve significance at $\alpha = 0.05$. This value was surpassed only in Parcour for mean temp. 5, max. temp. 5, and min. temp. 1. The two highest correlated variables per cultivar are given, together with the correlations with four variables which were selected using the factorial regression and the singular value decomposition (see below). For the more unstable cultivars (those responsible for more than 10 per cent of the interaction) min. temp. 1 was most frequently found among the highest correlated variables, followed by temp. sum 3 and mean temp. 5.

Factorial regression

The first attempt to explain the interaction included only linear terms. The best were min. temp. 1, mean temp. 5, max. temp. 5 and temp. sum 3 (see Table 6). The regression mean squares were tested over the mean intra-block error of the original experiments. Subsequently, the contributions of a second linear term and quadratic terms were investigated. The best significant second terms are also given in Table 6. It is obvious that the factorial regression represents something of an average over the individual regressions from the previous section. The best explanatory variable in the factorial regression was min. temp. 1.

AMMI analysis

Singular value decomposition of the matrix of interaction residuals resulted in a decomposition consisting

Table 4 Two-way analysis of variance results. Error is the mean intra-block error over the 12 trials

Source	Df	SS	MS
Cultivars	8	1,769,903	221,238
Trials	11	20,643,894	1,876,718
C×T	88	1,055,090	11,990
Error	288		5,781

Table 3 Mean seed yields plus marginal means in kg ha^{-1}. For abbreviations and codes see Materials and methods and Table 1

	1	2	3	4	5	6	7	8	9	10	11	12	
Se	1030	1844	1231	2059	1476	951	1250	956	1538	459	825	860	1207
We	1194	1865	1431	2242	1410	937	1342	1079	1722	455	782	904	1280
Co	992	1838	1406	2096	1217	813	1137	804	1556	502	595	635	1133
La	1073	1496	1151	1558	927	539	921	555	1186	387	377	581	896
Pe	829	1691	1317	2084	1284	805	1353	1027	1677	403	771	901	1178
Ba	1102	1604	1210	2113	1011	567	1060	587	1147	383	498	756	1003
Pa	965	1609	1221	1899	1258	908	1533	1066	1599	561	781	797	1183
Tr	1284	1868	1580	2293	1425	917	1404	1111	1693	666	829	1066	1345
Vi	1199	1897	1390	2138	1340	847	1217	855	1270	575	636	910	1189
	1074	1746	1326	2054	1261	809	1246	893	1488	488	677	823	1157

Table 5 Cultivar, percentage of total interaction sum of squares due to cultivar, first and second highest correlated variable to interaction residuals of a cultivar, correlations of interaction residuals with min. temp. 1, temp. sum 3, mean temp. 5, and max. temp. 5

			r			
Cultivar % SS int.	First and second highest corr. variable	*r*	min. temp. 1	temp. sum 3	mean temp. 5	max. temp. 5
Semperweide	Rel. hum. 5	0.74	0.44	−0.50	−0.31	−0.36
8%	Min. temp. 3	−0.69				
Wendy	Max. temp. 4	0.70	0.19	−0.48	−0.07	−0.00
5%	Mean temp. 4	0.69				
Compas	Rainfall 3	−0.60	−0.35	0.12	0.35	0.42
8%	Length per. 5	−0.54				
Lamora	Length per. 4	−0.71	−0.32	0.49	0.14	0.16
17%	Temp. sum 4	−0.69				
Perma	Min. temp. 1	0.69	0.69	−0.61	−0.60	−0.63
14%	Temp. sum 3	−0.61				
Barenza	Temp. sum 3	0.78	−0.77	0.78	0.75	0.72
16%	Min. temp. 1	−0.77				
Parcour	Mean temp. 5	−0.91	0.86	−0.68	−0.91	−0.89
20%	Min. temp. 1	0.86				
Trani	Rel. hum. 5	−0.45	−0.29	0.32	0.25	0.25
2%	Min. temp. 3	0.35				
Vigor	Mean temp. 5	0.74	−0.69	0.57	0.74	0.69
11%	Min. temp. 1	−0.69				

Table 6 Results of factorial regressions; best explaining single variables, best explaining pairs given best explaining singles, and some selected combinations

Variable(s)	% SS int. explained
Min. temp. 1	40
Min. temp. 1 + length per. 4	52
Min. temp. 1 + min. temp. 2	50
Mean temp. 5	38
Mean temp. 5 + length per. 4	53
Max. temp. 5	38
Max. temp. 5 + length per. 4	53
Temp. sum 3	35
Temp. sum 3 + mean temp. 5	50
Temp. sum 3 + temp. sum 4	48
Length per. 3 + length^2 per. 3	46

of eight terms explaining respectively 51.1, 19.4, 10.8, 10.1, 4.5, 2.4, 1.4 and 0.3 per cent of the interaction sum of squares. Mean squares were derived from the Eigenvalues by dividing the Eigenvalues by an approximation of the appropriate number of degrees of freedom; $I+J-1-2n$ (Gollob, 1968; see section *Statistical analyses* for the meanings of symbols). When tested against the mean intra-block error only the first two Eigenvalues were significant. Calculating an estimate for the error from the non-significant Eigenvalues gave 5,764, remarkably close to the intra block estimate of 5,781.

An interaction that can be described by two multiplicative components can be represented concisely in a biplot (Fig. 1). Cultivars are represented by lines, trials by points. Both the end-points of the lines for the cultivars and the points for the environments must be interpreted as the end-points of vectors starting at the origin. The length of a cultivar line roughly reflects the amount of interaction for that cultivar; thus according to Fig. 1 most interaction is due to Parcour, Lamora, Barenza and Perma (for confirmation see Table 5). The (cosine of the) angle between cultivar lines corresponds to the correlation between the interraction residuals. Hence, Vigor and Barenza are quite alike, Barenza and Perma are strongly negatively correlated and Parcour and Lamora are very dissimilar with a correlation close to zero. Scaling is such that distance between cultivars is indicative of the amount of interaction between them (Kempton, 1984); e.g. between Parcour and Barenza much interaction is present. An interaction residual for a particular combination of cultivar and trial can be recovered by calculating the inner product between their respective vectors. This is equivalent to the length

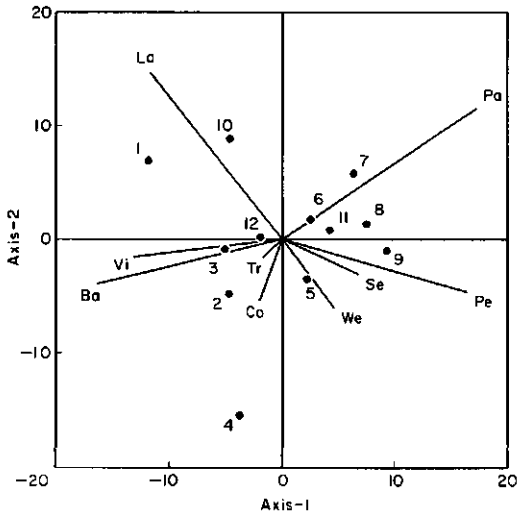


Fig. 1 Biplot constructed from cultivar (——) and trial (●) scores on the first two axes of the AMMI analysis. For abbreviations and codes see *Materials and methods* and Table 1.

Table 7 Weather variables with strongest correlations to environmental scores of AMMI-axis 1, 2, and environmental main effects, plus their mutual correlations

		Variable	r	Mutual correlations		
AMMI-axis 1	1	min. temp. 1	0.87			
	2	temp. sum 3	−0.82	−0.87		
	3	mean temp. 5	−0.81	−0.96	0.72	
	4	max. temp. 5	−0.81	−0.96	0.76	0.99
				1	2	3
AMMI-axis 2	1	length per. 5	0.66			
	2	min. temp. 2	−0.63	−0.91		
	3	temp. sum 5	0.61	0.95	−0.76	
	4	length per. 4	−0.58	−0.65	0.39	−0.82
				1	2	3
Env. main effect	1	length per. 5	−0.79			
	2	temp. sum 5	−0.78	0.95		
	3	rel. hum. 2	−0.77	0.84	0.72	
	4	min. temp. 2	−0.77	−0.91	−0.76	−0.91
				1	2	3

of the projection of either vectors onto the other multiplied by the length of the vector on which projection takes place. The sign of this product is positive if the angle between the vectors is acute, and negative for obtuse angles. Parcour can be seen to have high positive interaction residuals in trials 7, 8, and 9, which are of equal value, and a high negative residual in trial 4; Perma also has high residuals for 7, 8, and 9, but increasing in magnitude in this order; Lamora has high positive residuals in 1 and 10, and a high negative residual in 4.

AMMI analysis first axis

Figure 1 shows that the first axis forms a contrast between, on the left side, the late cultivars Lamora, Vigor, and Barenza, and, on the right side, the early cultivars Parcour and Perma. Environmental scores were regressed on the weather variables. Four regressions were found to be significant (Table 7 and Fig. 2). Addition of other linear or quadratic terms did not lead to significant increases in explanation. The selected variables were the same as those found by factorial regression. The separation of pattern and noise by means of a singular value decomposition with subsequent interpretation of scores through regression on environmental variables leads to the same conclusion as direct regression of interaction residuals on explanatory variables. This agreement of methods may serve to support the claim that the most important environmental variables are included in the set of selected variables with great certainty. However, collinearity makes it difficult to decide which variables have a primary causal effect and which have an associated effect. Cultivar reactions to the environmental circumstances represented by axis 1 could just as well be reflections of a reaction to minimum temperature in the 1st period as to temperature sum in the 3rd period, or mean or maximum temperature in the 5th period. An answer can only be obtained by additional experimentation.

Another problem was the clumping of the data points for certain variables such as temp. sum 3 and max. temp. 5 (Fig. 2). For temp. sum 3 the data were more or less divided into two clusters, thus precluding a check on linearity, and making the regression a contrast between the environments of 1987 on the one hand, and 1986 and 1988 on the other hand.

To avoid the formulation of an explicit relationship between environmental scores and measurements, appopriately scaled variable values were compared with confidence intervals for scores (interval matching, Fig. 3). For min. temp. 1 only the observed value of trial 11 was found just outside the confidence interval for the score. For temp. sum 3 trials 1 and 5 were just outside, and 11 was clearly outside the interval. For mean and max. temp. 5 trials 1, 2, 8, and 10 were outside the intervals. The agreement between scores and observed values for the different variables as assessed by interval

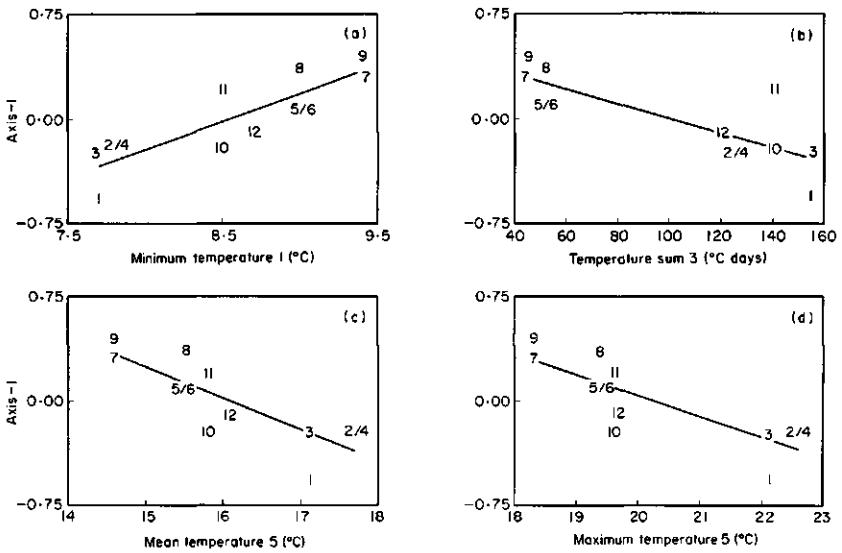


Fig. 2 Fitted (——) and observed (1...12) values for the regressions of axis 1 environmental scores on: (a) min. temp. 1; (b) temp. sum. 3; (c) mean temp. 5; (d) max. temp. 5. For codes see Table 1.

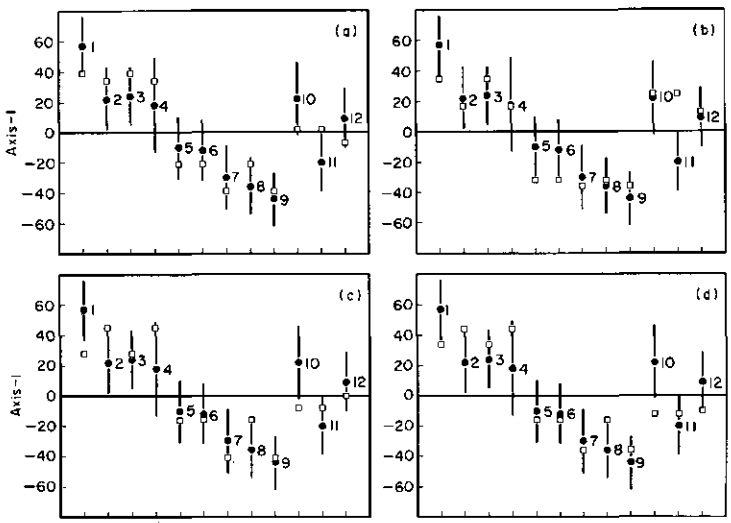


Fig. 3 Trial scores (●) on AMMI axis 1 with 95 per cent confidence bounds (——), and scaled values (□) for the environmental variables: (a) min. temp. 1; (b) temp. sum. 3; (c) mean temp. 5; (d) max. temp. 5. For codes see Table 1.

matching was thus comparable with that found by regression.

Interval matching was also used to test the hypothesis that axis 1 cultivar scores represent a contrast between early and late cultivars (Fig. 4a). The contrast involves the early cultivars Perma and Parcour with score $-3/\sqrt{[(2*-3^2)+(3*2^2)]}$ on the one hand and the late cultivars Lamora, Vigor and Barenza with score $2/\sqrt{[(2*-3^2)+(3*2^2)]}$ on the other hand. The other cultivars had a zero score. The denominator here is a normalizing factor which gives the vector of scores a squared length of unity. In Fig. 4a none of the contrast

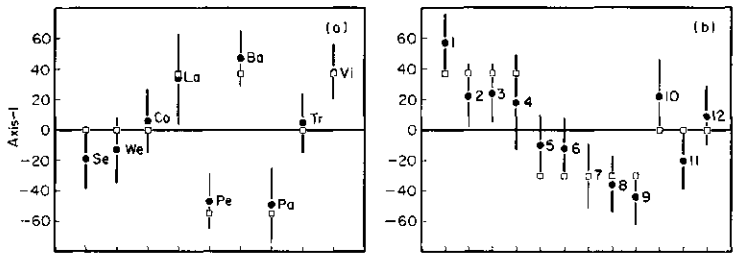


Fig. 4 (a) Cultivar scores (●) on AMMI axis 1 with 95 per cent confidence bounds (——), and scores (□) for the contrast of cultivars Perma and Parcour versus Lamora, Barenza, and Vigor. For abbreviations see Materials and methods. (b) Trial scores (●) on AMMI axis 1 with 95 per cent confidence bounds (——), and scores (□) for the contrast between the trials of 1986 vs. those of 1987. For codes see Table 1.

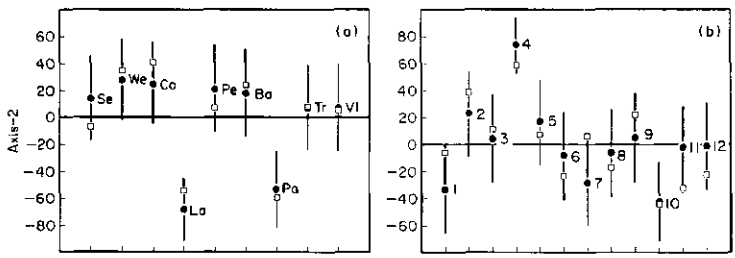


Fig. 5 (a) Cultivar scores (●) on AMMI axis 2 with 95 per cent confidence bounds (——), and scaled values (□) for the regression on the environmental mean stabilities. For abbreviations see Materials and methods. (b) Trial scores on AMMI axis 2 (●) with 95 per cent confidence bounds (——), and scaled values (□) for trial main effects. For codes see Table 1.

values fall outside the confidence limits, so that axis 1 could represent a contrast between early and late cultivars.

For the environmental scores the contrasts between sand and clay, 1986–1987, 1986–1988, 1987–1988, and first- and second-year crop were investigated. None of these gave promising results except the contrast 1986–1987. This contrast was reconcilable with the axis 1 trial scores (Fig. 4b). Not surprisingly this contrast was easily discernible in the values of min. temp. 1, lower in 1986 than in 1987, and the values of temp. sum 3, mean temp. 5, and max. temp. 5, higher in 1986 than in 1987.

AMMI analysis second axis

Figure 1 shows that the second axis is dominated by the cultivars Lamora and Parcour, which had relatively low yields in the highest yielding trial, 4, and relatively high yields in the lowest yielding trial, 10 (Table 3). Axis 2 scores for Lamora and Parcour combined with the scores for trials 4 and 10 account for 54 per cent of the sum of squares for this axis, or 10 per cent of the total interaction sum of squares. Lamora and Parcour seem to respond less to changes in the environmental factors than the other cultivars. Cultivar scores for axis 2 were correlated with the coefficients for the regressions of individual cultivar means on the average of all cultivars, $r = -0.94$. Environmental scores were correlated with environmental main effects, $r = -0.78$. Axis 2 thus seems to represent a regression on the environmental mean (Yates & Cochran, 1938; Finlay & Wilkinson, 1963). The percentage of the interaction sum of squares explained by this axis, 19 per cent was also close to that for the regression on the environmental mean, 14 per cent. Interval matching of cultivar and trial scores (Fig. 5) further subscribed to this view. In Fig. 5a zero values indicate average stability and negative values belong to the more stable cultivars, Lamora and Parcour.

The results of regressions of environmental scores on the weather variables are given in Table 7. Even the best explanatory linear term, length per. 5, could not be shown to be significantly related to the scores. A second linear, or a quadratic term did not add anything. Regressions of the environmental main effect on the weather variables also showed that axis 2 scores and environmental main effects were very similar (Table 7).

If axis 2 trial scores are indeed reflections of an underlying environmental variable it becomes difficult to explain why Lamora and Parcour respond very clearly to the environmental circumstances in trials 4 and 10, but almost not at all to the very similar circumstances in 2 (= 4), and 11 (= 10) (Fig. 1). Besides, it is hard to maintain that axis 2 is best interpreted in terms of stabilities, because as Fig. 5a reveals, Lamora and Parcour are the only cultivars without an average stability, while trials 4 and 10 are the only ones with

scores deviating clearly from 0. Probably axis 2 is better interpreted as modelling a multiple outlier for which no explanation in environmental terms is available.

Regression on the environmental mean versus AMMI

The first application of the method of regression onto the environmental mean to genotype by environment problems in grasses was probably that of Breese (1969) in his study of forage yield in cocksfoot. The method elicited enthusiasm as it seemed to provide an easily interpreted solution. Fifty-two per cent of the interaction was explained, but as Knight (1970) remarked with respect to the same data, genotype by environment interaction will remain intractable unless combinations of environmental factors are taken into consideration. Regressing individual responses on the mean of all genotypes is valuable, but only for broad studies of a collection of varieties.

Successful applications of the technique to forage yield in perennial ryegrass can be found in Troughton (1970), Samuel *et al.* (1970) and Hill & Samuel (1971). The percentages of explained interaction in these papers vary from 31 to 55. The method did not always accomplish such an adequate description of the genotype by environment interaction in grasses, as can be seen in a study of Nguyen *et al.* (1980) on tall fescue synthetics. In this case only 19 per cent of the interaction for the variable total herbage yield was described by heterogeneity of regression lines. In Gray (1982) total growth can be calculated at 15 per cent in an experiment with cocksfoot. Even worse is the figure for annual yield in smooth bromegrass (Tan *et al.*, 1979) which is 12 per cent. For forage yield in reed canary grass Barker *et al.* (1981) concluded that stability parameters, like regression coefficients, were not consistent, and that mean yield *per se* appeared to be the most reliable measure to evaluate forage yield performance.

Two frequently expressed criticisms towards the regression on the environmental mean are that the amount of interaction explained is low, and that the regressions are determined by only a few points (Westcott, 1986). Both criticisms apply to the regression solution to our genotype by environment problem. It is evident that the AMMI model possesses greater versatility in modelling interaction than the regression on the mean model, because it allows modelling in more than one dimension. Therefore, the AMMI solution to a genotype by environment problem is less prone to lead to a low percentage of explained interaction. As to the percentage of explained interaction, the AMMI model outperformed the regression on the mean model for our data because it identified a major component of interaction undetected by the regression on the mean, while it simultaneously contained the regression on the mean on another axis. Regression on the environmental mean becomes part of the AMMI solution if environmental scores mimic the environmental main effect. Perkins (1972) and Freeman & Dowker (1973) also identified an AMMI axis, the first in their cases, to be equivalent to a regression on the environmental mean.

When the majority of the interaction has a structure which deviates from the structure embodied in a regression on the environmental mean, the classical regression test on interaction, i.e. heterogeneity of slopes against deviations from regression, will often fail to detect any interaction at all. In that case the deviations from regression will provide an inappropriate measure for testing the heterogeneity of slopes. All interaction structure has to be removed from a genotype by environment table before a reasonable estimate for the error can be extracted. For our data, the estimate for the error taken from the cultivar by trial table after removing the first two multiplicative terms was almost equal to the intra block estimate. This strongly supported the conclusion that the two extracted AMMI axes represented structure, with the second axis being equivalent to regression on the mean, thereby proving the classical test on heterogeneity of slopes to be incorrect for our data.

Interval matching and power

It can be remarked that the way in which hypothesis testing proceeds with the interval matching possesses an inverse character. As the intervals become larger hypotheses become more difficult to refute. We do not think this forms a major problem as long as two conditions are fulfilled. In the expression for the confidence intervals it can be seen that the sizes of the intervals for the parameters of a particular axis are dependent on the estimate for the error, and the estimates for the scores and eigenvalues of other axes. A reliable estimate for the error, together with retainment of the appropriate number of axes, will safeguard the utility of the interval matching method. Reliability of the error estimate may be checked by comparison of the coefficient of variation for the experiment with published data or known standards. The coefficient of variation for our data was 13.1 per cent on a per plot basis. This seems reasonable when compared with the 11.7 per cent for forage yield of perennial ryegrass (Hill & Samuel, 1971), which is expected to be less variable than seed yield. Forage yield ranging from 22 to 65 per

cent was also reported in perennial ryegrass by Troughton (1970).

Causal modelling

The different methods used in this paper to investigate relations between interaction and environmental factors all identified the same variables as important. Despite the low number of independent measurements for the explanatory variables, which must have induced low power for most of the regression-related methods, this agreement of methods guarantees that the selected variables were at least statistically related to the interaction. For further insight into causal relations, developmental processes should be monitored through time, and not be characterized solely by an end product such as seed yield. This necessitates the use of statistical models that can accommodate changing relationships between numbers of variables through time, such as linear structural relationships models (Bollen, 1989). An alternative might be the use of simulation models for crop growth and development to bridge the gap between physiological studies and plant breeding practice (Hammer & Vanderlip, 1989).

Acknowledgements

We want to thank L. C. P. Keizer for his help with the figures, and J. B. Denis, C. Hackett, H. Jansen, A. P. M. den Nijs and P. Stam for comments on earlier versions of the paper.

References

ABOU-EL-FITTOUH, H. A., RAWLINGS, J. O. AND MILLER, P. A. 1969. Genotype by environment interactions in cotton – their nature and related environmental variables. *Crop. Sci.*, **9**, 377–381.

ATKINSON, A. C. 1985. *Plots Transformations, and Regression.* Clarendon Press, Oxford.

BARKER, R. E., HOVIN, W. A., CARLSON, I. T., DROLSOM, P. N., SLEPER, D. A. AND ROSS, J. G. 1981. Genotype-environment interactions for forage yield of reed canarygrass clones. *Crop. Sci.*, **21**, 567–571.

BOLLEN, K. A. 1989. *Structural Equations Modelling with Latent Variables.* Wiley, New York.

BREESE, E. L. 1969. The measurement and significance of genotype-environment interactions in grasses. *Heredity*, **24**, 27–44.

DENIS, J. B. 1980. Analyse de regression factorielle. *Biom. Praxim.*, **20**, 1–34.

DENIS, J. B. 1988. Two way analysis using covariates. *Statistics*, **19**, 123–132.

ELGERSMA, A. 1990a. Genetic variation for seed yield in perennial ryegrass (*Lolium perenne* L.). *Plant Breeding*, **105**, 117–125.

ELGERSMA, A. 1990b. Seed yield related to crop development and to yield components in nine cultivars of perennial ryegrass (*Lolium perenne* L.). *Euphytica*, **49**, 141–154.

FINLAY, K. W. AND WILKINSON, G. N. 1963. The analysis of adaptation in a plant-breeding programme. *Aust. J. Agric. Res.*, **14**, 742–754.

FREEMAN, G. H. AND DOWKER, B. D. 1973. The analysis of variation between and within genotypes and environments. *Heredity*, **30**, 97–109.

FRIPP, Y. J. 1972. Genotype-environmental interactions in *Schizophyllum commune*. II Assessing the environment. *Heredity*, **28**, 223–238.

GABRIEL, K. R. 1978. Least squares approximation of matrices by additive and multiplicative models. *J. R. Stat. Soc. Ser. B*, **40**, 186–196.

GAUCH, H. G. Jr 1988. Model selection and validation for yield trials with interaction. *Biometrics*, **44**, 705–715.

GOLLOB, H. F. 1968. A statistical model which contains features of factor analytic and analysis of variance techniques. *Psychometrika*, **33**, 73–115.

GOODMAN, L. A. AND HABERMAN, S. J. 1990. The analysis of nonadditivity in two-way analysis of variance. *J. Am. Stat. Ass.*, **85**, 139–145.

GORMAN, D. P., KANG, M. S. AND MILAM, M. R. 1989. Contribution of weather variables to genotype × environment interaction in grain sorghum. *Plant Breeding*, **103**, 299–303.

GRAY, E. 1982. Genotype × environment interactions and stability analysis for forage yield of orchardgrass clones. *Crop. Sci.*, **22**, 19–23.

HAMMER, G. L. AND VANDERLIP, R. L. 1989. Genotype-by-environment interaction in grain sorghum. III. Modeling the impact in field environments. *Crop. Sci.*, **29**, 385–391.

HAMPTON, J. G. AND HEBBLETHWAITE, P. D. 1983. The effects of the environment at anthesis on the seed yield and yield components of perennial ryegrass (*Lolium perenne* L.) cv. S.24. *J. Applied Seed Prod.*, **1**, 21–22.

HARDWICK, R. C. AND WOOD, J. T. 1972. Regression methods for studying genotype-environment interactions. *Heredity*, **28**, 209–222.

HILL, J. AND SAMUEL, C. J. A. 1971. Measurement and inheritance of environmental response amongst selected material of *Lolium perenne*. *Heredity*, **27**, 265–276.

KANG, M. S. AND GORMAN, D. P. 1989. Genotype × environment interaction in maize. *Agron. J.*, **81**, 662–664.

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

KNIGHT, R. 1970. The measurement and interpretation of genotype-environment interactions. *Euphytica*, **19**, 225–235.

MANDEL, J. 1961. Non-additivity in two-way analysis of variance. *J. Am. Stat. Ass.*, **56**, 878–888.

NGUYEN, H. T., SLEPER, D. A. AND HUNT, K. L. 1980. Genotype × environment interactions and stability for herbage yield of tall fescue synthetics. *Crop. Sci.*, **20**, 221–224.

PERKINS, J. M. 1972. The principal components analysis of genotype-environmental interactions and physical measures of the environment. *Heredity*, **29**, 51–70.

SAEED, M. AND FRANCIS, C. A. 1984. Association of weather

variables with genotype × environment interactions in grain sorghum. *Crop. Sci.*, **24**, 13–16.

SAMUEL, C. J. A., HILL, J., BREESE, E. L. AND HUNT, K. L. 1970. Assessing and predicting environmental response in *Lolium perenne*. *J. Agric. Sci., Camb.*, **75**, 1–9.

SNEDECOR, G. W. AND COCHRAN, W. G. 1980. *Statistical Methods*, 7th edn. Iowa State University Press, Ames.

STEFANSKY, W. 1972. Rejecting outliers in factorial designs. *Technometrics*, **14**, 469–479.

TAN, W. K., TAN, G. Y. AND WALTON, P. D. 1979. Regression analysis of genotype–environment interaction in smooth bromegrass. *Crop. Sci.*, **19**, 393–396.

TROUGHTON, A. 1970. Intra-varietal variation of yield in two varieties of *Lolium perenne* L. *Euphytica*, **19**, 382–389.

WESTCOTT, B. 1986. Some methods of analysing genotype–environment interaction. *Heredity*, **56**, 243–253.

WOOD, J. T. 1976. The use of environmental variables in the interpretation of genotype–environment interaction. *Heredity*, **37**, 1–7.

YATES, F. AND COCHRAN, W. G. 1988. The analysis of groups of experiments. *J. Agric. Sci., Camb.*, **28**, 556–580.

ZOBEL, R. W., WRIGHT, M. J. AND GAUCH, H. G. 1988. Statistical analysis of a yield trial. *Agron. J.*, **80**, 388–393.

Chapter V

Interpreting genotype-by-environment interaction using redundancy analysis

F.A. van Eeuwijk
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

be taken to an extension of the form

$$Y_{ij} = \mu + G_j + E_i + \sum_{l=1}^{L} U_{lj} V_{li} + \varepsilon_{ij}. \quad (2)$$

Model (2) is partially additive, partially multiplicative, and was first introduced by Gollob (1968) and Mandel (1969, 1971). In the multiplicative part, the U_{lj}s denote genotypical scores (sensitivities, stabilities) and the V_{li}s environmental scores (characterizations, indices). L indicates the number of multiplicative terms required for an adequate description of the interaction. Least-squares fitting of this model can be done in two stages. First, the additive terms are fitted in the usual way, then the remaining residual matrix is decomposed using the singular-value decomposition (Gabriel 1978). This model was already used in the context of plant breeding in the early 70s (Perkins 1972; Freeman and Dowker 1973). Recently it received renewed interest through Gauch and Zobel, who also introduced the term AMMI model, a shorthand for Additive Main effects and Multiplicative Interaction effects model (Gauch 1988; Gauch and Zobel 1988, 1989, 1990; Zobel et al. 1988). Model (2) certainly provides more modelling opportunities than model (1), but still defines the environment by quantities derived from the phenotypical observations themselves. These environmental characterizations may afterwards be related to explicitly measured environmental variables, that is *indirectly*, e.g., by regression or correlation.

Easy ways of *directly* relating genotype-by-environment interaction to environmental variables are: (a) regressing residuals from additivity on environmental variables for each genotype separately, or (b) using concomitant information on the environmental factor in the two-way ANOVA of genotypes by environments (Snedecor and Cochran 1980). The second case is a form of simultaneous regression, and will be referred to as factorial regression (Denis 1988). It amounts to regressing ANOVA interaction parameters on environmental variables. For an early example see Abou-El-Fittouh et al. (1969). An elaboration of factorial regression, but originally arrived at via a generalization of the AMMI approach, was obtained by Rao (1964). The method was dubbed principal components of instrumental variables. It can be understood as an AMMI model with a restriction on the environmental scores. These have to be linear combinations of measured environmental variables. Subsequently, the connection with multiple regression was established, e.g., by Hardwick and Wood (1972), Izenman (1975), Lefkovitch (1986), and Denis (1988). Hardwick and Wood probably were the first to note the applicability of the technique in a plant breeding context. So far Wood (1976) seems to be the only accessible application, though in rudimentary form. Finally, Van den Wollenberg (1977) developed the same method starting from canonical correlation analysis under the name of redundancy analysis.

Despite its apparent potential the technique has remained practically unknown in plant breeding. The present paper intends to stimulate interest in the method by describing the key features of the model together with an application to a real data set consisting of nitrate concentrations in lettuce.

Theory

Multivariate multiple regression

In order to describe the relationship between a set of genotypical responses and a number of environmental variables one could carry out multiple regressions for each of the genotypes on the set of explanatory environmental variables. Multiple regression aims at maximizing the multiple correlation coefficient; a measure of the association between a dependent variable and a set of independent variables. It can be shown that the multiple correlation coefficient is the maximum correlation between the dependent and a linear functions of the independents. The multiple regressions for a number of genotypical responses on the same set of environmental variables can be written in the form of a multivariate multiple regression model as follows

$$\mathbf{Y} = \mathbf{1}\mathbf{C}' + \mathbf{X}\mathbf{M} + \mathbf{E} \quad (3)$$

in which the columns of the matrix $\mathbf{Y}_{n \times m}$ represent the genotypical responses, the columns of the matrix $\mathbf{X}_{n \times q}$ the environmental variables; $\mathbf{1}_{n \times 1}$ stands for a vector of ones, $\mathbf{C}_{m \times 1}$ for the m intercepts, $\mathbf{M}_{q \times m}$ for the matrix of regression coefficients, while $\mathbf{E}_{n \times m}$ stands for a matrix of independently distributed normal errors with zero expectation and variance σ^2. Inclusion of a term for the row main effect changes (3) into a factorial regression model (Denis 1988), which is more appropriate in the context of genotype-by-environment interaction. However, for ease of exposition below, (3) will be used as a reference model, generalizations to factorial regression being obvious. Model (3) will be called the full-rank regression model for reasons to be explained shortly. In the full-rank model each genotype possesses unique sensitivities to every one of the environmental variables, no inter-relatedness between genotypical responses exists.

The environmental information as collected by the researcher will generally not have the form that is most relevant to the plants. Environmental variables of importance to the plants can be approximated by linear combinations of measured variables (possibly transformed, and including squares and cross products). In addition, it seems reasonable to assume that different genotypes react to *similar* environmental factors, though with varying sensitivity. A model that describes genotype-by-environment interaction in terms of heterogeneity in genotypical sensitivity to *common* linear combinations of environmental variables is given by the redundancy analysis model (Rao 1964; Hardwick and Wood 1972; Izenman 1975; Van den Wollenberg 1977; Davies and Tso 1982). The supposition of common linear combinations of environmental variables as the basis of genotype-by-environment interaction marks the distinction between the redundancy analysis model and the multivariate multiple regression model. The common linear combinations are found by rotation of the axes in the space spanned by the fitted values of the full-rank regressions for the genotypes. The rotation step may be followed by a reduction step in which only the most explanatory linear combinations are retained.

Redundancy analysis

Instead of maximizing the correlations between the individual dependent variables and the set of independent variables, as in

multiple regression, in redundancy analysis linear combinations of independent variables are formed that account for successively maximal proportions of the total sum of squares over the set of dependent variables. The quantity of central importance is the index of redundancy, introduced by Stewart and Love (1968).

Let $\mathbf{Y}^t = (Y_1, \ldots, Y_m)$ and $\mathbf{X}^t = (X_1, \ldots, X_q)$ be two sets of centered variables, and $\mathrm{SSSP}(\mathbf{Y}) = \mathbf{S}_{11}$, $\mathrm{SSSP}(\mathbf{Y}, \mathbf{X}) = \mathbf{S}_{12}$, $\mathrm{SSSP}(\mathbf{X}, \mathbf{Y}) = \mathbf{S}_{21}$, and $\mathrm{SSSP}(\mathbf{X}) = \mathbf{S}_{22}$, with SSSP a sums of squares and sums of products matrix. The index of redundancy is defined as

$$R^2(\mathbf{Y}:\mathbf{X}) = \frac{\mathrm{trace}(\mathbf{S}_{12}\,\mathbf{S}_{22}^{-1}\,\mathbf{S}_{21})}{\mathrm{trace}(\mathbf{S}_{11})}, \tag{4}$$

being the proportion of the total sum of squares in the Y-set which is accounted for by the linear prediction of Y by X. The analogy with the squared multiple correlation coefficient from multiple regression is obvious.

The coefficient vector $\mathbf{b}$ for the linear combination of independent variables $\mathbf{b}^t\mathbf{X}$ that describes the maximum proportion of the total sum of squares in the set of dependent variables Y can be found by maximizing the following function of $\mathbf{b}$

$$\xi(\mathbf{b}) = \mathbf{b}^t\mathbf{S}_{21}\mathbf{S}_{12}\mathbf{b} - \lambda(\mathbf{b}^t\mathbf{S}_{22}\mathbf{b} - 1) \tag{5}$$

(Van den Wollenberg 1977). For understanding (5) one should note that the sums of products between the dependent variables Y and the linear combination of independent variables $\mathbf{b}^t\mathbf{X}$, are given by $\mathrm{SSSP}(\mathbf{b}^t\mathbf{X}, \mathbf{Y}) = \mathbf{b}^t\mathbf{S}_{21}$, and the sum of the squares of these sums of products is simply $\mathbf{b}^t\mathbf{S}_{21}\mathbf{S}_{12}\mathbf{b}$. For convenience, and without loss of generality, the linear combinations are scaled to unit sum of squares, explaining the second term on the right in (5).

Differentiating (5) with respect to $\mathbf{b}$ and setting the result equal to zero leads, after some reshuffling, to the generalized eigenvalue problem

$$(\mathbf{S}_{21}\mathbf{S}_{12} - \lambda\mathbf{S}_{22})\,\mathbf{b} = 0. \tag{6}$$

The first eigenvector, $\mathbf{b}$, contains the weights for the X-variables, which are called canonical coefficients. The first eigenvalue, λ, represents the amount of the total sum of squares in Y explained by the linear combination $\mathbf{b}^t\mathbf{X}$. This linear combination represents the first redundancy variate. Subsequent redundancy variates, uncorrelated with preceding ones, can be obtained from subsequent eigenvectors.

Inspection of (6) also reveals the inter-connectedness of redundancy analysis and principal components analysis. When the Y- and X-set are the same $\mathbf{S}_{12} = \mathbf{S}_{21} = \mathbf{S}_{22}$ and (6) reduces to the equation for the principal components problem.

In the terminology of the genotype-by-environment problem, theoretical environmental variables are formed that minimize the total residual sum of squares of the regressions of the genotypical responses on these linear combinations of environmental variables. Genotypes, now, can be characterized by their covariances with the newly formed theoretical environmental variables.

Reduced rank regression

An alternative derivation of the method of redundancy analysis, which displays more clearly its least squares properties, is given by Davies and Tso (1982). They subsumed redundancy analysis under the wider class of reduced-rank regression models. The basic assumption underlying these models is that the matrix of regression coefficients is a matrix of low rank, in any case of lower rank than the full-rank multivariate multiple regression coefficients matrix. Reduced rank regression models arise naturally in situations where a number of Y-variables are known to be inter-related, as for genotypical responses.

The reduced-rank equivalent of the full-rank regression model (3), assuming the number of environments n to be greater than the number of measured environmental variables q, is written as

$$\mathbf{Y} = \mathbf{1}\,\mathbf{C}^t + \mathbf{Z}\,\mathbf{A} + \mathbf{E}, \tag{7}$$

in which $\mathbf{Z}_{n\times s}$ contains $s \leq q$ redundancy variates, linear combinations of the original environmental variables, that is, $\mathbf{Z} = \mathbf{X}\,\mathbf{B}$, with $\mathbf{B}_{q\times s}$ a matrix whose columns contain the weights for the environmental variables in X, the canonical coefficients. The columns of $\mathbf{A}_{s\times m}$ are made up of the covariances of the m responses in Y with the redundancy variates in Z, they are comparable with the regression coefficients in the Finlay-Wilkinson model.

Effectively, the reduced-rank argument is carried through by a factorization $\mathbf{M} = \mathbf{B}\,\mathbf{A}$ in (3). When M has rank $s = q$, model (7) represents the full-rank model (3), whereas for $s < q$ (7) denotes a reduced-rank model. The factorization can be found following a least squares argument (Davies and Tso 1982).

Methods

Assessing rank; maximum likelihood

A major issue arising in the application of redundancy analysis concerns the determination of the maximum rank s. It is appealing to base this decision on the residual sum of squares from the rank s fit

$$\mathrm{SS}_{\mathrm{res}(s)} = \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 + \sum_{i=s+1}^{\mathrm{rank}(\hat{\mathbf{Y}})} \lambda_i, \tag{8}$$

with $\|\mathbf{D}\|^2 = \sum_{ij} d_{ij}^2$ for a matrix $\mathbf{D}$ with elements d_{ij}, $\hat{\mathbf{Y}}$ the matrix of fitted values from the full-rank regression, and λ_i the i-th eigenvalue from (6) (which is equivalent to the i-th eigenvalue of $\hat{\mathbf{Y}}^t\hat{\mathbf{Y}}$ or $\hat{\mathbf{Y}}\hat{\mathbf{Y}}^t$). $\mathrm{SS}_{\mathrm{res}(s)}$ consists of the ordinary residual sum of squares from the full-rank fit plus a contribution of the least significant eigenvalues of (6).

Assuming the errors making up the matrix $\mathbf{E}$ in (7) to be distributed independently normal, with zero mean and variance σ^2, the loglikelihood can be written as

$$\mathrm{loglik} = -\tfrac{1}{2}\,\mathrm{n\,m}\,[\log_e(2\pi\sigma^2) + 1], \tag{9}$$

with $\log_e$ denoting the natural logarithm. From (9) the maximum likelihood estimator for σ^2 is obtained as $\hat{\sigma}^2 = \mathrm{tr}(\hat{\mathbf{E}}^t\hat{\mathbf{E}})/\mathrm{nm}$, with $\hat{\mathbf{E}}$ containing the residuals from the rank s fit (Van der Leeden 1990). The loglikelihood ratio test for the hypothesis of rank t against $t-1$ is most conveniently written as

$$\mathrm{lr} = \mathrm{nm}\log_e\left[\frac{\mathrm{SS}_{\mathrm{res}(t-1)}}{\mathrm{SS}_{\mathrm{res}(t)}}\right], \tag{10}$$

with $\mathrm{SS}_{\mathrm{res}(t-1)}$ and $\mathrm{SS}_{\mathrm{res}(t)}$ the residual sums of squares fom the rank $s-1$ and rank s fit. Asymptotically lr in (10) has a χ^2 distribution with a number of degrees of freedom equal to the difference between the degrees of freedom for the rank t model and the rank $t-1$ model. Assume that the data, Y, are corrected for the genotypical and environmental main effect, then the number of degrees of freedom for redundancy variates is equal to $q + (m-1) - (2t-1)$ for the t-th redundancy variate, where q stands for the rank of the X matrix ($n - 1 > q$), and $(m-1)$ for the rank of the corrected $\mathbf{Y}_{n\times m}$ matrix ($n > m$).

Assessing rank; randomization test

As an alternative for the loglikelihood ratio test a randomization test can be used. A possible approach is based on permutation of the rows of the **X** matrix (Ter Braak 1988). Calculate the first eigenvalue, then permute the rows of **X** and recalculate the first eigenvalue, repeat this v times. The significance level for the first eigenvalue is $(u+1)/(v+1)$, where u is the number of eigenvalues of the permuted set greater than the eigenvalue for the unpermuted **X**. For testing the second eigenvalue, correct **Y** for the first axis, etc. When the errors are uncorrelated, the columns of **X**, the environmental variables, may be permuted independently.

Variable selection and model building

Selection of variables in redundancy analysis can be performed either by techniques akin to those in discriminant analysis or, alternatively, by techniques used for multiple regression problems. One possibility is a stepwise procedure within a factorial regression set-up, in which the usual ANOVA tests for contrasts can be used (Snedecor and Cochran 1980). Subsequently, the rank of the matrix of regression coefficients, and simultaneously the number of axes to retain, can be assessed by means of the test given in (10). This procedure was recommended for modelling a matrix in terms of concomitant variables for rows (and/or columns) by Gabriel and Odoroff (1985).

An alternative, using backward elimination, is inspired by an idea of Jolliffe (1986, p. 108) in the context of principal components. Discard variables with high absolute coefficients in redundancy variates which express exact or nearly exact linear relationships between the explanatory variables, i.e., with zero or near-zero eigenvalues. This can be done iteratively. Fit the full-rank model, test whether or not the last redundancy variate contains significant information, e.g., by (10), and if yes discard the variable with highest absolute coefficient. Repeat this for the now reduced set of explanatory variables until the last redundancy variate appears no longer non-significant. Note that we will end up with a full-rank model, but a reduced set of explanatory variates. Nevertheless, the rotation in the space spanned by the fitted values of the individual genotypes can add to the interpretation of the interaction.

A word of caution should be expressed with respect to too heavy reliance on statistical variable selection procedures. Especially for genotype-by-environment problems, a reasonable choice of variables expected to be most influential should be possible beforehand, thereby reducing the need for elaborate statistical selection procedures.

Goodness of fit for individual genotypical responses

Evaluation of individual fits to responses can be done by considering the reduced-rank regression as a method to derive best linear predictors, the redundancy variates, for the set of responses. The regressions of the responses on the s redundancy variates then can be treated in a univariate fashion, making use of univariate evaluation procedures. Mean square errors of fit can be compared with known levels of precision for the type of response. In addition, prediction error on an independent set can be a useful evaluation criterion.

Precision of estimates of canonical coefficients

In order to say something about the precision with which canonical coefficients are estimated, a result of Tyler (1982) can be used. This shows that the canonical coefficients corresponding to the i-th redundancy variate, $\mathbf{b}_i$, can be interpreted, if scaled appropriately, as the vector of regression coefficients for the regression of $\mathbf{a}_i'\mathbf{Y}$, a linear combination of the responses Y weighted by their covariances with the i-th redundancy variate, on the X-set. Using the normalization $\mathbf{a}_i'\mathbf{a}_i = 1$ the regression of $\mathbf{a}_i'\mathbf{Y}$ on **X** gives as regression coefficients $\sigma_i\,\mathbf{b}_i$. Standard errors and t-values from the regression may be used for exploratory purposes.

Visualization of results

An important aid in the interpretation of the results of eigenvalue techniques is the biplot (Gabriel 1971). For an exposition on the use of biplots in genotype-by-environment problems see Kempton (1984).

In case of the AMMI model it is customary to depict scores for genotypes and environments on the first two axes in two-dimensional biplots. A rank-two approximation of the matrix of interaction residuals can be found from the biplot using the inner-product definition. Imagine the scores for the genotypes and the environments to determine vectors in two-dimensional space. Then, the interaction effect of a certain genotype in a certain environment is approximated by the inner-product between their respective vectors. The inner-product between two vectors is simply the length of the orthogonal projection from one vector onto the other, multiplied by the length of the other. A factor -1 or 1 is used as a multiplication factor depending on the angle between the two vectors; -1 for obtuse angles, 1 for acute angles. Ranking of interaction effects for all the genotypes in a particular environment can easily be done by just considering the ordering of the orthogonal projections of the genotypical vectors on that environmental vector.

For redundancy analysis the story is about the same as for the AMMI analysis. The major difference is that for redundancy analysis it is not the matrix of interaction residuals, but the matrix of *fitted* interaction residuals, which forms the raw material. Biplots for redundancy analysis have as an additional feature the possibility of representing measured environmental variables. For details on this and related aspects see Ter Braak (1990).

Computation

The calculations for a redundancy analysis can be done by any package that includes facilities for the singular value decomposition of matrices (in which case the matrix of full-rank fitted values must be the input) or for solving generalized eigenvalue problems such as (6). The calculations for the Application section were programmed in Genstat (1987). The package CANOCO (Ter Braak 1988) includes redundancy analysis among a number of other multivariate techniques, all furnished with facilities for forward selection of variables and permutation tests.

Application: nitrate concentration in lettuce

Data

In the period between March 1987 and June 1988 eight lettuce (*Lactuca sativa* L.) genotypes (Table 1) were evaluated at 18 harvesting times (Table 2) with respect to their nitrate concentrations (Reinink 1991). Each evaluation consisted of an experiment in eight blocks. The 18 evaluations in time were treated as environments in which genotypical performances were assessed. The average nitrate concentrations (g/l) of the eight genotypes observed in the 18 environments are given in Table 3. After a pre-

liminary selection eight environmental variables thought to exert influence on nitrate concentration (Tables 4, 5) were chosen for a characterization of the circumstances. Their usefulness to describe the genotype-by-environment interaction was investigated.

Preliminaries

Before searching for an explanation in terms of environmental variables, the existence of interaction has first to be proven. This involves testing for interaction [see Krishnaiah and Yochmowitz (1980) for a review] and, when present, determining whether the interaction is not due to a few outliers or removable by transformation. Then various methods should be tried to relate environmental variables to the interaction. In what follows the results of the following methods will be used: (a) stepwise regression of residuals from additivity on the set of environmental variables for each genotype separately; (b) factorial regression on the environmental variables; (c) AMMI analyis; (d) redundancy analysis. Different methods will elucidate different aspects of the data. At the same time, however, certain main features should become evident, as if looked upon from different angles.

Testing interaction in the two-way analysis of variance set-up (Table 6), using the mean intra-block error as an estimate for the error gave a highly significant result, $P \ll 0.001$. Another estimate for the error can be obtained via principal components analysis of the matrix of interaction residuals, which is part of the AMMI analysis, using the non-significant eigenvalues. The eigenvalues ex-

Table 1. Lettuce genotypes (*Lactuca sativa* L.) and their abbreviations

Pa	Panvit
DM	Deci-Minor
Pi	Pinto
GT	Große Brune Tête
RW	Reichenauer Winter
Wi	Winterbutterkop
Tr	Trocadero
Ls	*Lactuca sativa capitata*

Table 2. Trial numbers and harvesting times (day-month-year) of the trials corresponding to the environments 1 to 18

Trial	Harv. time	Trial	Harv. time	Trial	Harv. time
1	08-04-1987	7	25-11-1987	13	10-05-1988
2	06-05-1987	8	06-01-1988	14	18-05-1988
3	03-07-1987	9	19-02-1988	15	03-06-1988
4	10-09-1987	10	08-03-1988	16	14-06-1988
5	07-10-1987	11	30-03-1988	17	20-06-1988
6	05-11-1987	12	26-04-1988	18	30-06-1988

Table 4. Measured environmental variables in the environments of Table 2

Number	Variable
1	Electrical conductivity of the medium
2	Summed global radiation in Joule/cm^2/day on eighth last day before harvest
3	As 2 on fourth last day before harvest
4	As 2 on second last day before harvest
5	As 2 on last day before harvest
6	Daylength on sowing day in hours
7	As 6 on introduction NFT system
8	As 6 on harvesting day

Table 3. Mean nitrate concentrations (g/l) over the eight replicates of a randomized blocks design for the genotypes from Table 1 in the environments of Table 2

Environment	Genotype							
	Pa	DM	Pi	GT	RW	Wi	Tr	Ls
1	3.113	2.835	2.629	1.988	2.199	2.414	1.248	2.380
2	3.379	3.222	2.848	2.823	3.002	2.950	2.176	3.196
3	3.067	2.326	2.511	2.120	2.692	2.598	1.032	2.355
4	3.202	2.663	2.230	1.638	2.187	2.171	1.062	1.599
5	3.921	3.365	3.028	2.653	2.935	2.931	2.007	2.942
6	4.153	3.970	3.444	2.813	2.865	3.232	2.341	3.289
7	4.851	4.512	4.010	3.504	3.135	3.624	3.080	3.612
8	4.547	4.203	3.429	2.944	2.616	3.052	2.817	3.070
9	3.721	3.505	3.337	2.425	2.177	2.525	1.917	2.830
10	3.581	3.298	3.287	2.389	2.159	2.681	1.744	2.726
11	3.312	3.130	2.959	2.280	1.797	2.152	1.365	2.178
12	3.439	3.329	3.254	2.561	2.843	3.035	1.927	3.058
13	3.195	3.047	2.948	2.696	2.610	2.902	1.914	3.138
14	2.890	2.297	2.295	2.237	1.930	2.414	1.462	2.274
15	2.700	2.430	2.172	2.004	2.194	2.392	1.374	2.144
16	3.143	2.710	2.429	2.260	2.406	2.438	1.536	2.464
17	2.746	2.470	2.226	2.126	2.332	2.185	1.287	2.621
18	3.273	2.384	2.555	2.167	2.545	2.386	1.616	2.813

Table 5. Values of the environmental variables of Table 4 in the environments of Table 2

Environment	Environmental variable							
	1	2	3	4	5	6	7	8
1	2.1	1,136	993	911	881	14.75	11.23	13.15
2	2.1	1,345	1,277	1,250	1,815	11.78	13.28	14.93
3	2.1	1,700	2,191	2,586	2,556	16.14	16.48	16.37
4	2.1	1,076	1,090	1,323	1,065	14.81	13.61	12.74
5	2.2	960	779	539	457	13.61	12.46	10.89
6	2.0	316	482	421	556	12.81	11.23	9.04
7	2.0	145	117	102	42	11.91	10.29	8.05
8	1.9	109	93	127	42	10.96	8.71	7.73
9	2.2	555	504	415	383	9.64	8.50	9.90
10	2.0	641	663	596	780	8.45	7.98	11.78
11	2.0	676	666	541	546	7.70	9.04	12.60
12	1.6	1,951	2,427	2,413	2,286	9.22	12.05	14.39
13	1.5	1,651	1,789	1,276	1,518	11.78	13.01	15.21
14	1.5	2,281	2,359	2,376	2,514	13.15	14.45	15.62
15	1.5	1,244	1,456	1,604	1,398	14.07	15.37	16.23
16	2.3	1,398	1,852	2,719	2,975	14.51	15.96	16.45
17	1.5	2,041	1,515	1,350	988	14.93	16.23	16.48
18	1.5	1,326	1,416	1,779	1,580	15.26	16.39	16.41

Table 6. Two-way analysis of variance on the genotype-by-environment matrix of Table 3. The error is the mean intra block error over the 18 trials

Source	Df	SS	MS
Genotypes	7	31.333	4.476
Environments	17	29.325	1.725
Interaction	119	6.772	0.057
Error	882		0.009

Table 7. Sum of squares for interaction per genotype, selected explanatory set from a stepwise regression, percentage sum of squares explained, R^2, by the selected set, and by the first and second redundancy variate (linear combinations of variables 7 ad 8), and residual mean square from regression on first and second redundancy variate

Genotype	SS int.	MR-set	R^2			RMS-RA
			MR	RA1	RA2	
Pa	0.970	3 4 5 8	80	44	17	0.023
DM	1.297	6 8	81	84	0	0.013
Pi	0.674	4 7	61	35	30	0.015
GT	0.314	6 7	39	31	10	0.012
RW	1.744	1 3 5 7	85	75	3	0.024
Wi	0.440	2 3 4 6	75	49	0	0.014
Tr	0.495	1 2 3 5 8	67	9	5	0.027
Ls	0.840	2	31	27	4	0.036

plained respectively 61, 16, 11, 5, 4, 2, and 1% of the interaction sum of squares. Eigenvalues below 0.7 times the average percentage (i.e., $0.7 \times 100/7 = 10\%$) can be interpreted as noise (Jolliffe 1986). So the first three eigenvalues represent structure, the rest noise. Approximate degrees of freedom can be attributed using Mandel's (1971) simulation studies. Summing the last four eigenvalues and dividing by the appropriate degrees of freedom, 34.8, led to an error estimate of 0.023, again leading to a highly significant interaction. A reason for the difference between both estimates of error might be the extra contributions of environment-by-block, and genotype-by-environment-by-block interactions, to the estimate derived from the non-significant eigenvalues. As a corollary it can be remarked that the dimensionality of three for the interaction implied inappropriateness of the regression-on-the-mean model.

A check on outliers revealed no severe anomalies in the data. The maximum normed residual, the maximum absolute interaction residual divided by the square root of the interaction sum of squares (Stefansky 1972), amounted to only 0.21, which was far from significant. The estimate for the Box-Cox parameter for a power transformation (see Atkinson 1982) included the value 1 in its 95% confidence interval, so that there was no reason for a transformation either.

Multiple regression

The environmental variables from Table 5 were used as the explanatory set in stepwise regressions for the interaction residuals of the individual genotypes. The cut-off values were chosen as $F_{in} = F_{out} = 4$ (Montgomery and Peck 1982). The results are given in Table 7. Substantial parts of the interaction sums of squares can be described by the environmental variables. The problem, however, is that no pair of genotypes has the same set of explanatory variables. In fact all environmental variables end up three times in the eventual explanatory set, except variable 1

Table 8. Selected sets of variables for factorial regression, with order of variables within sets reflecting stepwise selection. Further columns; distribution of explained sums of squares over redundancy variates, and total sums of squares explained. All subset regressions and redundancy axes are significant at at least 5%, unless non-significancy (ns) is indicated

Variable (s)	RA-1	RA-2	RA-3	Total
7, 8	3.645	0.514		4.158
8, 6	3.593	0.522		4.117
7, 2	3.647	0.393		4.040
7, 3	3.586	0.378		3.965
2, 6	3.444	0.499		3.943
5, 6, 1	3.290	0.537	0.112^{ns}	3.938
3, 6	3.404	0.441		3.845
7, 6	3.430	0.413		3.844
4, 6, 1	3.139	0.495	0.159^{ns}	3.794
5, 6	3.062	0.348		3.441
7	3.361			3.361
8	3.360			3.360
4, 6	3.012	0.326^{ns}		3.338
2	3.019			3.019
3	2.846			2.846
6, 1	2.287	0.469		2.756
4	2.535			2.535
5	2.403			2.403
6	1.952			1.952
1	0.903			0.903

Table 9. Correlations between environmental variables and environmental scores from AMMI- and redundancy analysis (axes are linear combinations of variables 7 and 8)

Variable	AMMI-1	AMMI-2	AMMI-3	RA-1	RA-2
1	−0.40	−0.45	−0.11	−0.42	0.13
2	0.85	0.10	−0.12	0.84	−0.24
3	0.82	0.02	−0.17	0.83	−0.23
4	0.78	−0.13	−0.07	0.85	−0.12
5	0.76	−0.03	−0.07	0.80	−0.19
6	0.63	−0.45	0.36	0.66	0.59
7	0.89	−0.19	0.26	0.95	0.30
8	0.90	0.05	−0.12	0.95	−0.30
RA-1	0.94	−0.08	0.07	1.00	0.00
RA-2	−0.01	−0.40	0.63	0.00	1.00

(only two times) and variable 3 (four times). The multiple regression approach thus leads to a highly idiosyncratic description of the interaction.

Factorial regression

More parsimonious descriptions of the interaction residuals are possible with factorial regression. Just as in the case of separate multiple regressions, the interaction is related directly to environmental variables, but this is done simultaneously for all genotypes. Testing of the contributions of one or several variables can be done by means of usual F-tests. With the inclusion or exclusion of a variable, 7 degrees of freedom from the interaction are involved. Contributions were tested against the remainder of the interaction at 5%. The remainder might be tested against an independent estimate of the error, e.g., 0.023.

Table 8 gives the results of an all-subsets procedure. For pairs and trios, variables are given in order of inclusion following a stepwise procedure: for pairs starting from every one of the individual variables, which were all found significant at 5%, for trios starting from each of the pairs remaining after the elimination part of the preceding step. The pair consisting of variables 7 and 8 (daylengths at the introduction of the NFT system and at harvesting time) performs best with respect to the amount of the interaction sum of squares explained. However, the pair 8 and 6 (daylength at sowing date) does only slightly worse.

AMMI analysis

Part of the AMMI analysis was already presented above under Preliminaries (testing for interaction). To gain some insight into the meaning of the axes, the correlation of the environmental scores for the axes 1 to 3 with the environmental variables was calculated (Table 9). Only axis 1 shows a relationship with the environmental variables, especially with variables 7 and 8. For an easier understanding of the meaning of this result one can look at the biplots of axis 2 against 1, 3 against 1, and 3 against 2 (Fig. 1 a, b, c). The scaling is such that the score vectors for the environments have squared lengths equal to the eigenvalues, whereas the genotypes have squared lengths of 1. With this scaling in the biplot of axis 2 against 1 the squared distance between the environmental approximates to twice the amount of interaction between them (Kempton 1984).

AMMI axis 1, AMMI-1, can be seen in Fig. 1 to represent roughly a contrast between summer (environments 2, 3, 13, 15, 17, and 18, having high positive scores) and and winter (environments 7, 8, 9, and 11, having high negative scores). This conclusion is in accordance with the high positive correlations of AMMI-1 with daylength at introduction NFT, variable 7, and harvesting date, variable 8. Daylength is greater in summer than in winter. AMMI-2 is dominated by the environments 4 (highly positive) and 13 (highly negative). To say AMMI-2 represents a contrast between spring and autumn would be overinterpreting. Just as AMMI-2 is not very easily relatable to environmental circumstances, neither is AMMI-3.

Redundancy analysis

One way of starting the redundancy analysis is by investigating the possibilities for rank reduction of the matrices of regression coefficients of the factorial regressions. In Table 8 the distribution of the interaction sum of squares over the redundancy variates is given for each

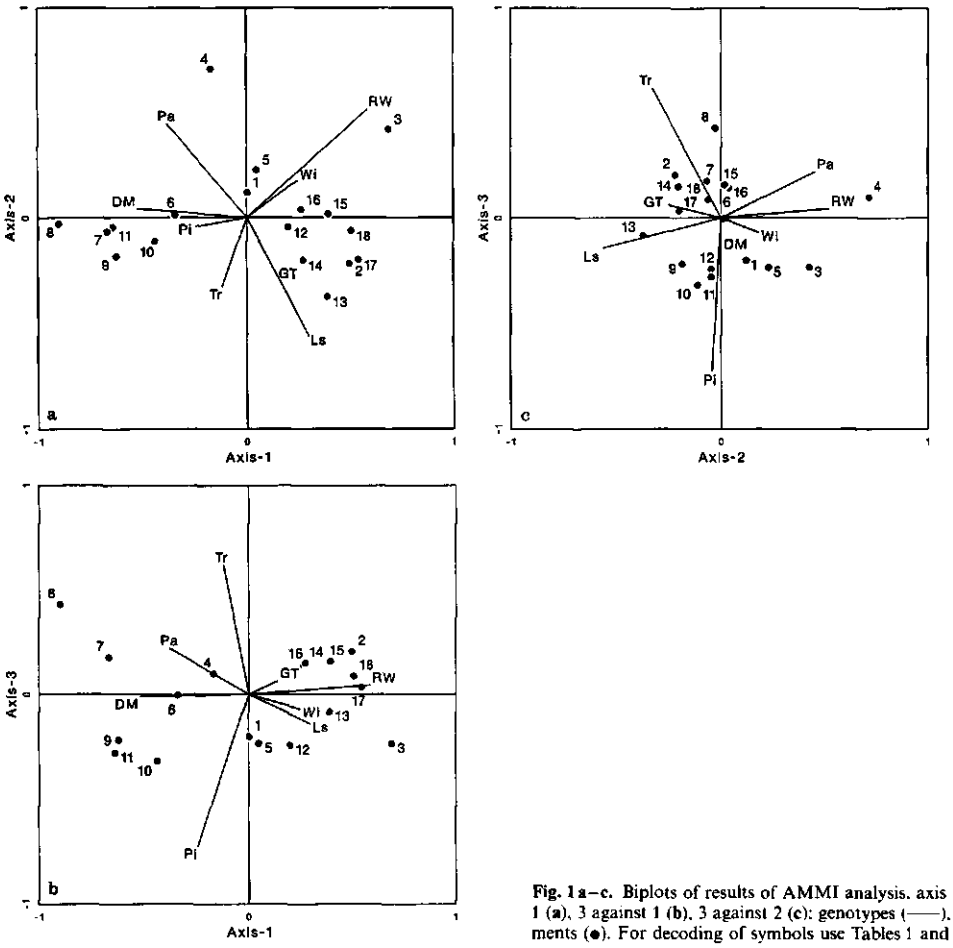


Fig. 1a–c. Biplots of results of AMMI analysis, axis 2 against 1 (**a**), 3 against 1 (**b**), 3 against 2 (**c**); genotypes (——), environments (●). For decoding of symbols use Tables 1 and 2

factorial regression. It seems natural to take the best set, pair 7 and 8. The test for a rank reduction using (10) reads: $\text{lr} = 18 \times 8 \times \log_e \left[\frac{6.772 - 3.645}{6.772 - (3.645 + 0.514)} \right] = 25.859$. The quantity lr is, under the null-hypothesis, of no second dimension, asymptotically distributed as a χ^2 with $q + (m-1) - (2t-1) = 6$ degrees of freedom, $q = 2$, $m = 8$, $t = 2$ (see Methods section). This means that no rank reduction is possible, as the 5% point for $\chi^2_6 = 12.592$. The necessity for the full-rank model was confirmed by a permutation test for the second dimension, conditional on the first dimension, $P \leq 0.05$ (see Methods section). The coefficients for the (standardized) variables in the first redundancy variate were 0.13 for 7, and 0.13 also for 8, approximate t-values (see Methods section) were 3.36 and 3.37. Corresponding values of the second redundancy variate were 0.40 and −0.40, with t-values of 4.85 and −4.85. The coefficients were scaled in such a way that the sum of squares for the environmental scores was 1.

The first redundancy variate is the sum of the daylenghts at harvesting time and a month earlier, so high values will be found in summer and low values in winter (recall that X-variables were centered to mean zero), while intermediate values will be found in spring and autumn. The second redundancy variate is the difference between both daylength variables. During summer and winter daylength will not change very much, resulting in almost zero values for this redundancy variate. However, in spring and autumn daylength changes, and the second redundancy variate will become positive in autumn and

negative in spring. The two redundancy variates together thus describe a reaction of nitrate concentration to day-length throughout the year.

The biplot for the nitrate data (Fig. 2) immediately reveals that the genotype-by-environment interaction is a season-dependent phenomenon; the environments are arranged in a closed curve running counter-clockwise from summer at the right via autumn at the top, winter at the left, and spring at the bottom, to summer again at the right. Scaling is just as for the AMMI biplots; that is, environmental scores have sum of squares equal to the eigenvalue of the corresponding axis. The distance between environments is proportional to the amount of interaction between them. Most interaction can be identified between the extreme winter environments 7, 8 and 9 on the left, and the extreme summer environments 3, 15, 16 and 17 on the right.

The data set offers the opportunity for an internal check of the adequacy of the model because, for some dates, data are available from 1987 as well as 1988. To be more specific; environment 1 (8-4-87) may be expected to be located between 11 (30-3-88) and 12 (26-4-88), 2 (6-5-87) has to be in the neighbourhood of 13 (10-5-88) and 14 (8-5-88), and 3 (3-7-87) has to be near 18 (30-6-88). Inspection of Fig. 2 corroborates these expectations, thereby vindicating the chosen model.

Further evidence for the correctness of the redundancy solution is given by the position of the genotype RW in the biplot. This genotype was selected for its extremely low nitrate concentrations under low light conditions (Reinink et al. 1987). The genotype RW has above average nitrate concentrations in summer, so that highly positive inner-products results from the projection of summer points (3, 15, 16, 17, 18) on the RW vector, whereas RW has below average nitrate concentration in winter, and highly negative inner-products result from the projection of winter points (8, 9, 10) on the RW vector.

The cosine of the angle between the genotypical vectors may be interpreted as an estimate of the correlation between genotypical responses over environments. Genotypes RW and DM seem to behave as antipodes.

Information about the fits for the individual genotypical responses (in fact individual genotypical deviations from additivity) to the redundancy variates is given in the last three columns of Table 7. There it can be seen that the genotypes with the greater amounts of non-additivity, DM and RW, seem especially to determine the first redundancy component; that ist, their explained sums of squares are the highest. For the second component, genotypes Pa and Pi seem to be the most important. The proportion of variance explained by the regressions on both redundancy variates is a measure for the quality of the representation of the individual genotypes in the biplot. Genotypes DM and RW are well represented, genotypes Tr and Ls are poorly represented.

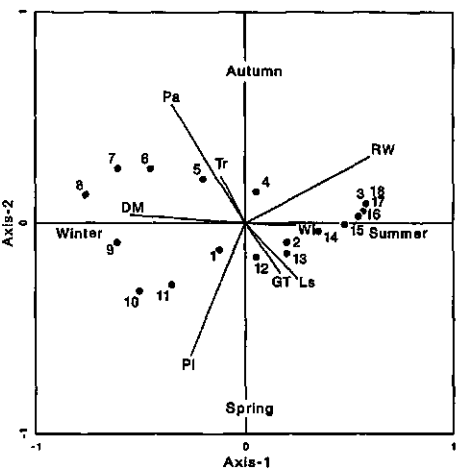


Fig. 2. Biplot of results of redundancy analysis; axis 2 against 1; genotypes (——), environments (●). For decoding of symbols use Tables 1 and 2

Table 10. Backward elimination of variables by discarding variable with highest coefficient on non-significant last redundancy variate

Variable	Coefficients						
1	0.02	0.02	−0.13	0.03	0.07	0.26	–
2	−0.44	−0.70	–	–	–	–	–
3	0.58	0.40	−0.02	−0.73	−0.45	–	–
4	−0.32	0.38	−0.19	0.77	–	–	–
5	0.60	–	–	–	–	–	–
6	−0.16	−0.30	0.46	0.18	0.08	−0.13	0.26
7	0.14	0.44	−0.60	–	–	–	–
8	−0.43	−0.33	0.44	−0.18	0.36	0.21	−0.19
Last eigenvalue	0.000	0.003	0.012	0.041	0.061	0.194	0.522
Sum eigenvalues	5.512	5.370	4.998	4.775	4.449	4.365	4.117

The residual mean squares for all genotypes except Ls are quite comparable, supporting the view that the redundancy analysis has taken up almost all structure from the data. The exception, Ls, has a higher residual mean square, probably due to interaction caused by factors other than the amount of light. The mean of the residual mean squares over the genotypes is 0.021, which is close to the 0.023 that was derived from the AMMI analysis.

An alternative to the above procedure is to start off from a full-rank model incorporating all environmental variables, and then test the significance of the last redundancy variate. Upon non-significance the variable with the highest coefficient is discarded (see Methods section). This process is repeated until the last redundancy variate

turns out to be significant. The results are given in Table 10. The test for the fourth redundancy variate for the model with the variables 1, 3, 6, and 8 reads $lr = 18 \times 8 \times \log_e \left[\frac{6.772 - (4.459 - 0.061)}{6.772 - 4.459} \right] = 3.748$. Compared to the 5% value of χ^2_4, 9.488, this means non-significance. As variable 3 had the highest coefficient it was discarded. The test for the third redundancy variate for the model with the variables, 1, 6, and 8 reads $lr = 18 \times 8 \times \log_e \left[\frac{6.772 - (4.365 - 0.194)}{6.772 - 4.365} \right] = 11.162$. The 5% value for χ^2_5 is 11.070. On this criterion the final set would be 1, 6, and 8. However the loglikelihood ratio test is slightly over-sensitive (see Discussion) and, therefore, it is better to continue until clearer significance for the last redundancy variate is found. After removing variable 1, a final set, 6 and 8 (daylength at sowing and at harvest time) is found for which both redundancy variates are clearly significant; lr for the second redundancy variate is 25.847 ($P < 0.001$). The interpretation is equivalent to the one arrived at earlier. The first redundancy variate is again a sum of both environmental variables, with most extreme values in summer and winter, and the second their difference, being extreme in spring and autumn. This is not surprising; in Table 8 it could already be seen that the pairs 7 and 8, and 8 and 6, explain the interaction almost equally well. Variables 6 and 7 have a correlation of 0.82, and should be exchangeable in combination with 8. In fact all pairs of variables selected in Table 8, except those including variable 1, would have led to the interpretation given above.

Discussion

Comparison of analyses

Various methods can lead to a very similar interpretation of the interaction. This important conclusion follows from the analyses in the Application section. In analysing genotype-by-environment tables one should use different approaches and, upon agreement, interpretation is straightforward, whereas upon disagreement closer inspection is necessary thereby acknowledging the differences between the method and the kind of structure they are supposed to detect.

For the nitrate data, AMMI and redundancy analysis gave comparable results, though the first extracts environmental scores as linear combinations of residuals from additivity, whereas the second forms environmental scores from linear combinations of measured environmental variables. The first AMMI axis paralleled the first redundancy axis, while the second and third AMMI axis more or less collapsed into the second redundancy axis (Table 9). The resemblance of AMMI and redundancy solutions means that for the redundancy analysis all relevant variables were selected (Ter Braak 1987). This is a useful diagnostic for the interpretation of interaction.

The individual regressions per genotype were mainly given as a reference point for the other analyses. Individual regressions have the advantage of high specificity, but the disadvantage of low parsimony. Moreover, it seems more likely that genotypes react to *common* environmental factors as can be uncovered by redundancy analysis. The dimension reduction property of the redundancy analysis was eventually not used for the redundancy analysis departing from factorial regression. Nevertheless, the interpretation of the interaction was certainly facilitated by the rotation, and the axes do bear on the physiology of the plants, as witnessed by the repeatability of the environmental scores in time. Besides, the fact that the positions of the genotypes in the redundancy biplot (Fig. 2) were scattered over all four quadrants means that the axes transcend a purely statistical interpretation, because in the latter case genotypes would be more likely to be situated near the lines $y = x$ and $y = -x$, since genotypes would bear no particular relationship to the extracted axes.

The dimension-reducing faculty of redundancy analysis proved very beneficial in the backward elimination procedure in the search for a good subset. However, strictly speaking, after final selection of variables 6 and 8, further rank reduction was not allowed. Real rank reduction can be seen to occur in Table 8 for the sets 5, 6, 1; 4, 6, 1; and 4, 6. For these sets the last redundancy variate turned out to be non-significant. Though one would not base an interpretation on these sets, since better ones are available, the estimation of the regression coefficients for these sets should be more accurate using the lower rank approximation of the matrix of regression coefficients due to the separation of structure in the retained dimension(s), and noise in the discarded dimension(s) (Gauch 1982).

In the Application section a slight over-sensitiveness of the loglikelihood ratio test was mentioned. This phenomenon is best illustrated by situations for which F-tests, as well as loglikelihood ratio tests, can be calculated. Consider the inclusion of variable 1 in the model after having fitted main effects. The loglikelihood ratio test is $lr = 144 \times \log_e [6.772/(6.772 - 0.903)] = 20.608$ (see Table 8), to be compared with a χ^2_7 distribution, so $P = 0.004$. The F-test is $f = [0.903/7]/[6.772 - 0.903)/112] = 2.46$, to be compared with an $F_{[7;112]}$ distribution, giving $P = 0.022$. Somewhat less obvious is the following example. Take the pair 8, 6, and the trio 8, 6, 1. From Table 8 we know that 6 and 8 together explained a sums of squares of 4.117. Adding 1 raises this amount to 4.365 (see Table 10). An F-test for inclusion of 1 has the form $f_{[7;98]} = [(4.365 - 4.117)/7]/[(6.772 - 4.365)/98] = 1.44$, $P = 0.198$, so inclusion of variable 1 seems not to be supported by this F-test. On the other hand having found that both

redundancy variates are significant for the pair 8 and 6 (Table 8), a possible test for the need of the inclusion of 1 is to test the third redundancy variate for the trio 8, 6, 1. The loglikelihood ratio test here gives $P = 0.048$ (see Application section). A reason for the liberality of the loglikelihood ratio test could be that, though it is based on the comparison of two estimates for the residual variance, it does not take into account the different degrees of freedom on which the estimates are based. However, in general, F-test and loglikelihood ratio test do not deviate much, and it seems recommendable anyway, not to adhere too strictly to the results of significance testing. They are best used as rough guides.

Extensions and other applications of the redundancy analysis model

An appealing extension of redundancy analysis is the so called partial redundancy analysis, in which not only environmental variables, but also one or more covariables, are present (Davies and Tso 1982). To obtain the partial redundancy analysis solution the environmental variables are first regressed on the covariables, after which the residuals of these regressions replace the environmental variables in the subsequent redundancy analysis. In this way the contribution of particular environmental variables conditional on the contribution of other environmental variables is testable.

In the same vein, AMMI analysis and redundancy analysis can be combined. First, extract the significant redundancy variates; next, search for structure in the residuals by performing a singular value decomposition on them to see whether there is any structure left. Of course, covariables or conditioning can again be incorporated in this analysis.

Instead of interpreting the genotypical responses as variables and the environments as sample points, one could analyse the reversed situation of the genotypes within environments constituting variables and the genotypes over environments being sample points. Explanatory variables can then express either group structure in the genotypes or contrasts between them. A straightforward generalization of the redundancy analysis model even makes it possible to investigate both types of dependence simultaneously (Denis 1988; Velu 1991).

Another interesting application of redundancy analysis lies in the search for informative genotypes with respect to environmental circumstances, say indicator genotypes. Consider the model consisting of the genotypical main effect and the first dimension of the singular-value decomposition of the data corrected for the genotypical main effect. This model is almost equivalent to a regression on the mean model; the genotypical scores, in a reparameterized form, are estimates for the regression coefficients, and the environmental scores are estimates for the environmental main effects. Rewrite this model as a redundancy model by choosing as explanatory variables the (centered) genotypical responses themselves. When subsequently a subset selection procedure is applied to the explanatory genotypical responses a maximally adequate subset of informative genotypes will be retained. A similar approach is possible with respect to the environments.

Acknowledgements. I am indebted to Kees Reinink for allowing me to use the data set and for helpful discussions. I also thank Jean-Baptiste Denis, Aad van Eijnsbergen, John Gower, Hans Jansen, Piet Stam, Cajo ter Braak, Eeke van der Burg, and two anonymous referees for their comments on drafts of the manuscript. Paul Keizer was responsible for the graphics.

References

Abou-El-Fittouh HA, Rawlings JO, Miller PA (1969) Genotype by environment interactions in cotton – Their nature and related environmental variables. Crop Sci 9:377–381

Atkinson AC (1982) Regression diagnostics, transformations and constructed variables. J R Stat Soc Ser B 44:1–36

Davies PT, Tso M K-S (1982) Procedures for reduced-rank regression. Appl Stat 31:244–255

Denis JB (1988) Two way analysis using covariates. Statistics 19:123–132

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agric Res 14:742–754

Freeman GH, Dowker BD (1973) The analysis of variation between and within genotypes and environments. Heredity, 30:97–109

Gabriel KR (1971) Biplot display of multivariate matrices with application to principal component analysis. Biometrika 58:453–467

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

Gabriel KR, Odoroff CL (1985) Some reflections on strategies of modelling: How, when and whether to use principal components. In: Sen PK (ed) Statistics in biomedical, public health and environmental sciences, Elsevier science publisher B.V., North Holland, pp 315–331

Gauch HG Jr (1982) Noise reduction by eigenvector ordinations. Ecology 63:1643–1649

Gauch HG Jr (1988) Model selection and validation for yield trials with interaction. Biometrics 44:705–715

Gauch HG Jr, Zobel RW (1988) Predictive and postdictive success of statistical analyses of yield trials. Theor Appl Genet 76:1–10

Gauch HG Jr, Zobel RW (1989) Accuracy and selection success in yield trial and analyses. Theor Appl Genet 77:473–481

Gauch HG Jr, Zobel RW (1990) Imputing missing yield trial data. Theor Appl Genet 79:753–761

Genstat 5 Committee (1987) Genstat 5 reference manual. Clarendon Press, Oxford

Gollob HF (1968) A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33:73–115

Hardwick RC, Wood JT (1972) Regression methods for studying genotype-environment interactions. Heredity 28:209–222

Izenman AJ (1975) Reduced-rank regression models for the multivariate linear model. J Mult Anal 5:248–264

Jolliffe IT (1986) Principal components analysis. Springer-Verlag, New York

Kempton RA (1984) The use of biplots in interpreting variety by environment interactions. J Agric Sci 103:123–135

Krishnaiah PR, Yochmowitz MG (1980) Inference on the structure of interaction in two-way classification model. In: Krishnaiah PR (ed) Handbook of statistics, vol 1. North-Holland Publishing Company, Amsterdam, North Holland, pp 973–994

Lefkovitch LP (1986) Linear predictivity: An alternative for MANOVA and Multivariate multiple regression. Biom J 28:771–781

Mandel J (1969) The partitioning of interaction in analysis of variance. J Res Nat Bureau of Standards B Math Sci 73B:309–328

Mandel J (1971) A new analysis of variance model fo non-additive data. Technometrics 13:1–18

Montgomery DC, Peck EA (1982) Introduction to linear regression analysis. Wiley, New York

Perkins JM (1972) The principal components analysis of genotype-environmental interactions and physical measures of the environment. Heredity 29:51–70

Rao CR (1964) The use and interpretation of principal component analysis in applied research. Sankhya A 26:329–358

Reinink K (1991) Genotype x environment interaction for nitrate concentration in lettuce. Plant Breed 107:39–49

Reinink K, Groenwold R, Bootsma A (1987) Genotypical differences in nitrate content in Lactuca sativa L. and related species and correlation with dry matter content. Euphytica 36:11–18

Snedecor GW, Cochran WG (1980) Statistical methods, 7th edn. Iowa State University Press, Ames.

Stefansky W (1972) Rejecting outliers in factorial designs. Technometrics 14:469–479

Stewart D, Love WA (1968) A general canonical correlation index. Psych Bull 70:160–163

Ter Braak CJF (1987) Ordination. In: Jongman RHG, Ter Braak CJF, Van Tongeren OFR (eds) Data analysis in community and landscape ecology. Pudoc, Wageningen, The Netherlands, pp 91–169

Ter Braak CJF (1988) CANOCO – a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). Agricultural Mathematics Group, Wageningen, The Netherlands

Ter Braak CJF (1990) Interpreting canonical correlation analysis through biplots of structure correlations and weights. Psychometrika 55:519–531

Tyler DE (1982) On the optimality of the simultaneous redundancy transformations. Psychometrika 47:77–86

Van den Wollenberg AL (1977) Redundancy analysis an alternative canonical correlation analysis. Psychometrika 42:207–219

Van der Leeden R (1990) Reduced rank regression with structured residuals. DSWO Press, Leiden.

Velu RP (1991) Reduced rank models with two sets of regressors. Appl Statist 40:159–170

Wood JT (1976) The use of environmental variables in the interpretation of genotype-environment interaction. Heredity 37:1–7

Yates F, Cochran WG (1938) The analysis of groups of experiments. J Agric Sci 28:556–580

Zobel RW, Wright MJ, Gauch HG (1988) Statistical analysis of a yield trial. Agron J 80:388–393

Chapter VI

Genotype x strain interactions for resistance to *Fusarium* head blight caused by *Fusarium culmorum* in winter wheat

C.H.A. Snijders & F.A. van Eeuwijk
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

Theor Appl Genet (1991) 81:239–244

Genotype × strain interactions for resistance to *Fusarium* head blight caused by *Fusarium culmorum* in winter wheat

C. H. A. Snijders and F. A. Van Eeuwijk

Centre for Plant Breeding Research (CPO), P.O. Box 16, NL-6700 AA Wageningen, The Netherlands

Received March 23, 1990; Accepted July 13, 1990
Communicated by A. R. Hallauer

Summary. In 3 consecutive years, a set of 17 winter wheat genotypes, representing a wide range of *Fusarium* head blight resistance, was inoculated with four strains of *Fusarium culmorum*. *Fusarium* head blight ratings were analyzed. The interaction between genotypes, strains, and years was described using a Finlay-Wilkinson model and an Additive Main effects and Multiplicative Interaction effects (AMMI) model. The interaction consisted primarily of a divergence of genotypical responses with increasing disease pressure, modified by genotype-specific reactions in certain years. The divergence was mainly caused by one very pathogenic strain. The *Fusarium* head blight resistance in this study can be described as horizontal resistance in terms of Vanderplank, with the exception of three genotypes selected from one particular cross that showed a 'strain-year combination' dependent resistance which was ineffective in 1 year.

Key words: *Fusarium* head blight resistance – Plant breeding – *Fusarium culmorum* – Genotype × environment interaction – Wheat – AMMI model

Introduction

In The Netherlands, *Fusarium* head blight in wheat (*Triticum aestivum* L.) is predominantly caused by *Fusarium culmorum* (W. G. Smith) Sacc. and *Fusarium graminearum* Schwabe. Both species of *Fusarium* have a worldwide distribution as soil inhabitants and cause, in addition to head blight root, foot and stem rot. Both fungi are generalists infecting cereals and a large number of other hosts, including corn, peas, and alfalfa (Booth 1971). Both *Fusarium* spp. are nonobligate parasites and facultative saprophytes. Variation for *Fusarium* head blight resistance in wheat exists (Atanasoff 1924; Parry et al. 1984; Schroeder and Christensen 1963; Snijders 1990). The resistance found until now is of a moderate form. Complete resistance has not been demonstrated. It is not clear whether or not *Fusarium* head blight resistance can be described as horizontal resistance in terms of Vanderplank (1984), i.e., whether or not the variation in resistance in the population of the host is independent of the variation in the population of the pathogen.

In a 3-year study of *Fusarium* head blight resistance, Mesterhazy (1984) found significant genotype × isolate interaction each year between 11 isolates of *F. graminearum* and two wheat genotypes. Using two isolates of *F. graminearum* and two isolates of *F. culmorum* for artificial inoculation of 21 genotypes, Mesterhazy (1988) found significant interactions for genotype × *Fusarium* species and genotype × *Fusarium* isolate. However, interaction patterns were not stable over experiments, and genotype ranking was only slightly influenced by the isolates. No evidence has been found for the occurrences of races of *Fusarium culmorum* or *F. graminearum* adapted to different wheat genotypes. Also, in studies with *F. graminearum* in corn ear rot tests, significant but inconsistent isolate × genotype interaction patterns were found (Atlin et al. 1983; Mesterhazy 1982; Mesterhazy and Kovacs 1986). However, large genotype rank reversals did not occur. This phenomenon is not restricted to *Fusarium* of wheat and corn. Environmental lability of interactions between wheat cultivars and isolates were also reported for *Cercosporella herpotrichoides* (Scott and Hollins 1977).

In an initial study of *Fusarium culmorum* in wheat, a significant host genotype × pathogen strain interaction was observed (Snijders 1987). The experiments were continued to investigate the consistency of the interaction patterns, i.e., whether or not strain-specific resistance to *Fusarium culmorum* head blight in wheat exists.

Materials and methods

Host and pathogen

A set of 17 winter wheat cultivars and SVP[1]-lines was composed, representing the whole range of *Fusarium* head blight susceptibility based on data available in 1985. In field trials in 3 consecutive years this set was tested for resistance to *Fusarium* head blight. Ten strains taken from monospore cultures of isolates of *Fusarium culmorum*, collected in The Netherlands, were prescreened for pathogenicity in the glasshouse. Two nonpathogenic strains were discarded, and from the remaining strains four were drawn: IPO 39-01, IPO 329-01, IPO 348-01, and IPO 436-01, originating from isolations from a grain of seed, culm, head, and leaf sheath, respectively. The lyophilized strains are deposited at the Research Institute for Plant Protection (IPO), Wageningen. Each year, conidiospores for inoculation were produced in 1-l Erlenmeyer flasks containing 250 ml sterilized cereal seeds: the 1st year, wheat seeds (cultivar Arminda), the next 2 years, a wheat (Arminda) and oat (bulk) seed mixture (3:1). A lyophilized strain was used as starting inoculum. The cultures were incubated in darkness at 25 °C for 2 weeks, followed by 3-week incubation at 5 °C. To prepare spore suspensions, conidia were washed from the kernels with water. Since wheat is most susceptible to *Fusarium* head blight at anthesis (Schroeder and Christensen 1963), experimental inoculations were made at that time. The spore suspensions were applied at 1 l/10 m^2. To ensure a high relative humidity during the nights after inoculation, the field was sprinkled in the evening for 1 h each day over a period of 2 weeks. Head blight ratings were determined as the product of the percentage of heads infected and the proportion of infected spikelets per infected head (Snijders and Perkowski 1990). In all experiments, interplot interference was prevented.

Field trial 1986

On November 22, 1985, seeds were sown in sandy soil in Wageningen at a standard density of 330 seeds/m^2 in rows 0.25 m apart. A split-plot design was established, with two blocks. Each main plot, consisting of one genotype, was divided into subplots of 0.90 × 0.75 m, over which the strains of *F. culmorum* were randomized so that the experimental subplots were separated from each other by border subplots of the same size. Further details are described in Snijders and Perkowski (1990). *Fusarium* head blight was assessed 26 days after first inoculation.

Field trial 1987

On November 4, 1986, seeds were sown in Flevoland in clay soil at a standard seed density of 330 seeds/m^2 in rows 0.25 m apart. A split-plot design was established, with three blocks. The four strains of *Fusarium culmorum* were randomized over the main plots. A distance of at least 4 m between the main plots prevented interplot interference. The main plots were divided into subplots of 2.00 × 0.75 m, over which the wheat genotypes were randomized. On June 25, when 30% of the wheat genotypes flowered, all genotypes were inoculated. The spore concentrations varied from 25,000 to 250,000 spores per milliliter. At the time when 100% of the genotypes flowered, July 2, a second inoculation was done. For the inoculation a spraying machine was used, which sprayed from 0.3 m above the crop. Spore concentrations varied from 25,000 to 250,000 spores per milliliter. On July 21, 26 days after the first inoculation, head blight was assessed. Observations were made on culm length.

Field trial 1988

On November 10, 1987, seeds were sown in Flevoland in clay soil. The same design was used as in field trial 1987. Each subplot consisted of a hill plot (Ø 0.25 m) seeded with 3 g seeds, at 0.5 m apart. Experimental inoculation was done on June 2, when 30% of the wheat genotypes flowered, and repeated on June 9 and June 16, by which time 100% of the genotypes flowered. For the inoculation a spraying machine was used. The spore suspensions had a concentration of 250,000 spores per milliliter. On June 30, 28 days after the first inoculation, head blight was assessed. Observations were made on time of anthesis and culm length.

Statistical analysis

For the analysis of variance of *Fusarium* head blight ratings in the three consecutive experiments, the split-plot model with fixed effects was used (Steel and Torrie 1981). For a description of the interactions, a Finlay-Wilkinson regression model (Finlay and Wilkinson 1963) and an Additive Main effects and Multiplicative Interaction effects (AMMI) model (Bradu and Gabriel 1978; Gauch 1988; Kempton 1984; Zobel et al. 1988) were used.

Results and discussion

As no head blight was observed in control and border plots, interplot interference was assumed to be absent. Inoculum concentration for individual inoculations and total amount of inoculum had no influence on the *Fusarium* head blight ratings. No significant correlations were found between *Fusarium* head blight and time of anthesis, and head blight and culm length. From preliminary analyses (data not shown), it was concluded that within the experiments of 1986 and 1987, there was a statistically significant interaction between wheat genotypes and *Fusarium* strains, which could not be removed by transformation of the data to an angular or logistic scale. In 1988 there was no significant interaction between genotypes and strains, nor was there a significant strain effect.

The means over the replicates of the genotypical assessments per strain within each of the 3 years are presented in Table 1. This table shows the high pathogenicity of strain IPO 39-01. The nonadditivity of the head blight ratings is striking. The head blight data of Table 1 were subjected to a Finlay-Wilkinson analysis, for which each strain-year combination was treated as a separate environment. The model may be written

$$Y_{ijk} = \mu + G_i + \beta_i E_j + I^*_{ij} + e_{ijk},$$

where μ is the mean value over all genotypes and environments, G_i is the effect of the i^{th} genotype, the regression coefficient β_i is a measure of the stability of the i^{th} genotype, E_j is the effect of the j^{th} environment, I^*_{ij} is the residual interaction after allowing for differences in stability between the genotypes, and e_{ijk} is the error for the k^{th} individual within the ij^{th} genotype-environment.

1 The Foundation for Agricultural Plant Breeding (SVP) is now part of the Centre for Plant Breeding Research (CPO)

Table 1. *Fusarium* head blight incidence[a, b] of 17 wheat genotypes for four *F. culmorum* strains and 3 years. Genotypes are presented in ascending order of incidence averaged over strains and years

Wheat genotype	1986				1987				1988			
	IPO 39-01	329-01	348-01	436-01	39-01	329-01	348-01	436-01	39-01	329-01	348-01	436-01
SVP 72017-17-5-10[c]	2.0	1.5	3.0	1.5	7.3	0.8	0.5	2.1	5.3	2.7	2.0	2.7
SVP 77076-4	9.0	1.0	3.0	1.5	13.5	0.3	0.1	0.2	7.0	1.7	3.3	3.7
Arina	8.0	2.5	5.0	4.0	12.0	0.3	0.1	2.8	6.3	1.0	2.0	3.3
SVP 77076-38	18.0	1.0	4.5	7.0	8.9	0.2	0.1	0.7	2.0	2.3	1.3	1.7
SVP 77076-1	6.0	3.0	1.0	1.5	11.1	0.1	0.4	2.7	7.0	4.7	5.0	8.3
Saiga	4.5	7.5	4.5	9.0	15.5	1.1	0.3	1.7	9.3	9.7	6.0	11.7
SVP 77078-30	9.0	13.0	1.0	2.5	17.8	0.6	0.4	8.4	15.7	5.3	3.3	4.7
SVP 72003-4-2-4	23.5	4.0	9.5	3.5	16.3	0.7	0.4	1.4	5.0	6.3	7.3	13.7
SVP 77079-15	27.5	4.5	2.5	8.5	35.2	1.9	0.5	5.8	3.3	6.3	4.0	8.3
SVP 75059-28	11.0	3.5	3.0	1.5	54.0	1.2	1.2	19.3	4.0	2.7	1.3	6.3
SVP 73030-8-1-1	60.0	7.0	7.5	9.0	36.3	3.0	5.0	7.5	13.0	4.7	5.0	5.7
SVP 73016-2-4	47.0	18.0	14.5	22.5	44.1	5.4	3.5	11.6	6.7	12.5	3.3	4.0
SVP 73012-1-2-3	67.5	16.0	17.5	17.0	34.2	7.0	2.5	9.3	11.0	9.0	6.0	25.7
SVP 75059-46	25.5	5.0	5.0	10.5	69.3	5.0	1.7	13.2	30.0	22.3	20.0	20.3
Nautica	62.5	20.5	20.0	30.5	32.2	1.3	0.8	4.8	40.0	25.0	18.0	20.3
SVP 75059-32	32.5	5.0	9.0	42.5	57.3	5.2	4.3	30.5	37.7	14.7	38.3	31.0
SVP 72005-20-3-1	62.5	16.5	27.5	23.0	58.5	3.7	2.7	21.7	36.7	26.3	13.3	20.3
Mean	28.0	7.6	8.1	11.5	30.8	2.2	1.4	8.5	14.1	9.2	7.6	11.3

[a] Head blight ratings were determined as the product of the percentage of heads infected and the proportion of infected spikelets per infected head

[b] Values presented are means over blocks

[c] SVP-line code: the first two digits indicate the year of crossing, followed by three digits representing the crossing number. The number after each hyphen is a selection number

Table 2. Summary of the results from the Finlay-Wilkinson analysis

Term	*df*	SS	MS
Genotype	16	12,368	773**
Environment	11	15,201	1,382**
Genotype × environment	176	13,867 (100%)	79**
Regressions	16	5,662 (41%)	354**
Concurrence	1	4,586	4,586**
Deviations from concurrence	15	1,076	72*
Deviations from regressions	160	8,205 (59%)	52*
Error	>51		25

* Significant at $P \leq 0.01$
** Significant at $P \leq 0.001$

To fit the Finlay-Wilkinson model, first the genotypical and environmental main effects are estimated in the customary way for ANOVA. Subsequently, the individual genotypical responses are repressed on the estimated environmental main effects to find estimates for the parameters β_i. The heterogeneity between regression lines has to account for the genotype × environment interaction. This approach is quite usual for yield data, but may seem somewhat unorthodox for disease incidences. Problems with respect to inference may be expected from failure of the assumptions for analysis of variance, such as homogeneity of variance and normality. However, in this study the Finlay-Wilkinson model only served as a starting point for a more appropriate model, and no ultimate conclusions are derived from the model itself. The results of the Finlay-Wilkinson analysis are shown in Table 2. The heterogeneity between lines accounted for 41% of the total interaction and the description seems to be acceptable. The plot of the fitted regression lines, Fig. 1, approaches a special case of the Finlay-Wilkinson model, namely, the situation where all regression lines intersect at the same point. This model is equivalent to the concurrence model (Mandel 1969)

$$Y_{ijk} = \mu + G_i + E_j + cG_iE_j + I^*_{ij} + e_{ijk},$$

where μ, G_i, E_j, I^*_{ij}, and e_{ijk} have the same interpretation as in the Finlay-Wilkinson model, while c is the only extra parameter needed for a description of the interaction. With this model 81% of the interaction that was explained by the heterogeneity of the Finlay-Wilkinson fitted lines can be covered (Table 2). This means that the genotype × environment interaction as described by the Finlay-Wilkinson model consists mainly of a divergence of (centered) genotypical responses. This interpretation gains even more credibility from the strong associations existing between the evaluations of the genotypes over the set of environments as measured by Spearman rank

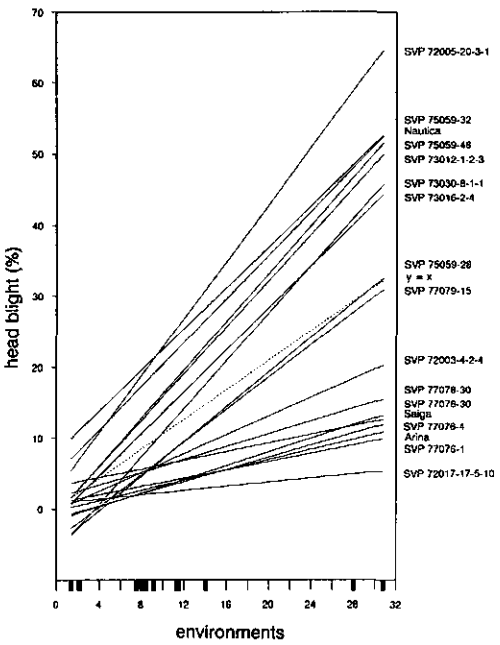


Fig. 1. Regression lines of individual head blight ratings on mean head blight ratings per environment (formed by strain-year combinations) for 17 individual genotypes. The hanging symbols on the abscissa represent the environments, which are presented in Table 1

correlations (data not shown). These were all positive; 58 out of 66 were significant at $P \leq 0.05$.

At first glance, this would seem an adequate explanation of the interaction. However, a first problem arises in the context of the deviations I_{ij}^* from the regressions. When tested against an error estimate of 25 with at least 51 *df* (Table 2), being the geometric mean of the error estimates for genotype-environment means over the 3 years, the deviations appear to be significant. This implies that in addition to the divergence of the regression lines, other factors are involved in the interaction. Obviously, the Finlay-Wilkinson model does not remove all pattern from the data. Furthermore, a plot of the residuals against the fitted values exhibited an increase of the variance with the mean.

A second problem is that a considerable part of the environmental range is not represented by actual measurements, invalidating an interpretation of the regression coefficients as stability measures. The regressions mainly express a contrast between the high disease incidences in the environments formed by IPO 39-01 in 1986 and 1987, and the rest of the strain-year combinations. To a major extent the slopes were determined by the two high incidence environments (Fig. 1; Table 1). The influence of these two environments was investigated more closely by performing an analysis without them. The overall treatment sum of squares decreased dramatically from 41,436 to 13,579. However, the proportion of genotype × environment interaction remained more or less the same, 36% in the reduced set against 33% in the full set. Now a concurrence model gave an adequate description of the genotype × environment interaction, that is, deviations from the concurrence model were not significant anymore. However, the rank order of the slopes showed some clear reversals in comparison to the rank order derived from the Finlay-Wilkinson analysis for the full set of environments. This means that if circumstances had been such that only low disease pressures had occurred, an interaction analysis would have led to a concurrence model and the ranking of genotypes for stability would not have been predictive for situations with higher disease pressures.

It was evident that a description of the interaction in tems of a Finlay-Wilkinson model for the full set was not satisfactory. An alternative was a model with Additive Main effects and Multiplicative Interaction effects, an AMMI model. This model may be written

$$Y_{ijk} = \mu + G_i + E_j + \sum_{n=1}^{N} \lambda_n a_{ni}\, b_{nj} + I_{ij}^* + e_{ijk},$$

where μ, G_i, E_j, I_{ij}^*, and e_{ijk} have the same interpretation as above, while λ_n^2 is the eigenvalue for axis n of the principal components analysis, and a_{ni} and b_{nj} are the corresponding genotypical and environmental scores. The a_{ni} may be interpreted as genotypical stabilities, while the b_{nj} may be seen as environment characterizations. N denotes the number of multiplicative terms necessary for an adequate description of the interaction. The model can be fitted by first calculating additive main effects for genotypes and environments, followed by a principal components analysis (singular value decomposition) of the matrix of the residuals (Gabriel 1978).

With respect to the assessment of N, two strategies are possible: (i) a strategy based on postdictive success, i.e., the ability of a model to fit its own data (e.g., traditional F-tests), and (ii) a strategy based on predictive success, the ability to predict validation data not used in constructing the model (Gauch 1988; Gauch and Zobel 1988). Because of the fact that in our experiment main- and subplot treatments changed over the years, assessment of predictive success was not straightforward. Therefore, model validation took place on postdictive grounds. Approximate F-tests were done after ascribing degrees of freedom to the eigenvalues following Mandel (1969) and calculating the corresponding mean squares. A summary of the ANOVA for the AMMI model is shown in Table 3. Three multiplicative terms seem necessary for an adequate description of the interaction. The

Table 3. Summary of the results from the AMMI analysis

Term	*df*	SS	MS
Genotype	16	12,368	773**
Environment	11	15,201	1,382**
Genotype × environment	176	13,867 (100%)	79**
Component 1	48[a]	6,199 (45%)	129**
Component 2	36	4,061 (29%)	113**
Component 3	27	1,756 (13%)	65*
Rest	65	1,851 (13%)	17
Error	>51		25

* Significant at $P \leq 0.01$
** Significant at $P \leq 0.001$
[a] The degrees of freedom for the components are calculated according to Mandel (1969)

Table 4. Genotypical scores ($\times 10^{-2}$) from the AMMI analysis, normalized at squared length 1

Genotype	Component		
	1	2	3
SVP 72017-17-5-10	−31	−13	1
SVP 77076-4	−21	−9	0
Arina	−23	−11	−1
SVP 77076-38	−12	−22	−3
SVP 77076-1	−27	−10	5
Saiga	−28	−7	12
SVP 77078-30	−23	−4	3
SVP 72003-4-2-4	−8	−18	−2
SVP 77079-15	4	6	−21
SVP 75059-28	−8	44[a]	−46[a]
SVP 73030-8-1-1	36	−13	−21
SVP 73016-2-4	23	−4	−32
SVP 73012-1-2-3	38	−27	−16
SVP 75059-46	5	57[a]	7
Nautica	32	−23	55[a]
SVP 75059-32	2	43[a]	49[a]
SVP 72005-20-3-1	38	12	11

[a] Genotypes with high scores used as a basis for component interpretation

rest of the terms is not significant when tested against the error estimate introduced above. The plot of residuals showed no gross failures of the assumptions. The interpretation of the components is as follows:

1. The first component provides genotypical scores, 'stabilities' (Table 4), and environmental scores (Table 5) that are closely correlated with the stabilities and scores from the Finlay-Wilkinson model. From the environmental scores in Table 5 it can be seen that the first component is the contrast between IPO 39-01 in 1986 and 1987, on the one hand, and the rest of the strain-year combinations, on the other hand. The proportion of variance explained by this component is 45% (for comparison, 41% in the Finlay-Wilkinson model). It can be concluded that the first multiplicative term is more or less equivalent to the Finlay-Wilkinson regressions.

Table 5. Environmental scores ($\times 10^{-2}$) from the AMMI analysis, normalized at squared length 1

Environment Strain	Year	Component		
		1	2	3
IPO 39-01	1986	−83[a]	−37	−8
IPO 329-01	1986	−10	−26	−19
IPO 348-01	1986	−2	−25	−11
IPO 436-01	1986	4	−2	35
IPO 39-01	1987	30[a]	79[a]	−31
IPO 329-01	1987	−25	−14	−31
IPO 348-01	1987	−28	−16	−31
IPO 436-01	1987	−15	21	−20
IPO 39-01	1988	0	13	63[a]
IPO 329-01	1988	−8	−2	12
IPO 348-01	1988	−19	6	24
IPO 436-01	1988	−10	1	18

[a] Environments with high scores used as a basis for component interpretation

2. The second component arises from nonadditivity of the genotypes SVP 75059-28, SVP 75059-32, and SVP 75059-46 inoculated with strain IPO 39-01 in 1987 (Tables 4 and 5). These three selections from the same cross had a far higher *Fusarium* head blight incidence in 1987 after inoculation with IPO 39-01 than may be expected from the genotypical and environmental main effects plus the Finlay-Wilkinson coefficients.

3. The third component results from genotype SVP 75059-28, with a far lower than expected incidence, and genotypes SVP 75059-32 and Nautica, with a higher than expected incidence. Again, this component is mainly due to an IPO 39-01 reaction, this time in 1988. The interpretation of this component is not easy. It probably represents merely noise but, as a consequence of a postdictive validation strategy, prone to lead to overfitting (Gauch 1988), is not identified as such. An estimate for the amount of noise in the overall treatment sum of squares is the product of the treatment degrees of freedom with the error estimate: $203 \times 25 = 5{,}075$. Acknowledging the fact that the noise will predominantly turn up in the higher axes, a strong argument for an interpretation of axis three in terms of noise is given.

The AMMI model thus provides a good description of the data, including the genotype × environment interaction, and uncovers some features we were not able to disclose before. Altogether the interaction may be said to consist primarily of a divergence of the incidences at higher disease pressures, modified by genotype-specific reactions in certain years. However, the modifications are on the whole not such that they heavily disrupt the rankings of the genotypes over the environments, al-

though incidental changes occur. The divergence is mainly caused by the highly pathogenic strain IPO 39-01.

A last point concerns the scale of the measurement. A percentage scale was used, as experience has shown that this is a convenient scale for resistance breeding research. For the purpose of genetic analyses the scale should preferably be one on which the analysis is as simple as possible, which means one on which interactions are small or absent. Various empirical transformations were tried. The most successful was the complementary log log transformation, which removed the genotype × strain × year interaction completely. However, genotype × year interaction and, to a lesser extent, genotype × strain interaction remained significant. The conclusions with respect to the status of resistance type, horizontal, did not change. The complementary log log transformation confers extra weight to the lower percentages. This seems unjustifiable in the light of the size of the measurement error. Therefore, the original percentage scale was retained. For nonremovable interactions, Mather (1971) remarked that "we must always be prepared to bring interaction explicitly into an analysis."

Conclusions

The three environments with the highest disease pressure were the combinations of one particular strain (IPO 39-01) with the 3 years. No evidence was found for strain-specific resistance. The *Fusarium* head blight resistance in this study can be described as horizontal resistance in terms of Vanderplank (1984), with the exception of the lines selected from cross SVP 75059, which showed a 'strain-year combination' dependent resistance, ineffective in 1987. For large-scale screening for resistance to *Fusarium* head blight using experimental inoculation, highly pathogenic strains should be used. The use of an AMMI model for the description of genotype × strain interaction over years allows conclusions not obtainable by the additive models used in the studies reported in the introduction. It provides a means to check whether the environmental 'lability' of interaction in the aforementioned studies was really part of the pattern in the data and hence merits agricultural interpretation, or whether it was merely noise.

Acknowledgements. Financial assistance from The Netherlands Grain Center (NGC) is gratefully acknowledged. Technical assistance from N. M. P. Pieterse was appreciated. The authors thank Dr. J. Jansen, Dr. A. P. M. den Nijs, Prof. Dr. J. E. Parlevliet, Dr. P. Stam, and two anonymous referees for their useful advice.

References

Atanasoff D (1924) *Fusarium* blight of the cereal crops. Agric Univ Wageningen Papers 27-4

Atlin GN, Enerson PM, McGirr LG, Hunter RB (1983) *Gibberella* ear rot development and zearalenone and vomitoxin production as affected by maize genotype and *Gibberella zeae* strain. Can J Plant Sci 63:847–853

Booth E (1971) The genus *Fusarium*. Commonwealth Mycological Institute, Kew, UK

Bradu D, Gabriel KR (1978) The diplot as a diagnostic tool for models of two-way tables. Technometrics 20:47–68

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant breeding programe. Austr J Agric Res 14:742–754

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

Gauch HG Jr (1988) Model selection and validation for yield trials with interaction. Biometrics 44:705–715

Gauch HG Jr, Zobel RW (1988) Predictive and postdictive success of statistical analyses of yield trials. Theor Appl Genet 76:1–10

Kempton RA (1984) The use of biplots in interpreting variety by environment interactions. J Agric Sci 103:123–135

Mandel J (1969) The partitioning of interaction in analysis of variance. J Res Nat Bur Stand Sect 73B:309–328

Mather K (1971) On biometrical genetics. Heredity 26:349–364

Mesterhazy A (1982) Resistance of corn to *Fusarium* ear rot and its relation to seedling resistance. J Phytopathol 103:218–231

Mesterhazy A (1984) A laboratory method to predict pathogenicity of *Fusarium graminearum* in field and resistance of wheat to scab. Acta Phytopathol Acad Sci Hung 19:205–218

Mesterhazy A (1988) Expression of resistance of wheat to *Fusarium graminearum* and *F. culmorum* under various experimental conditions. J Phytopathol 123:304–310

Mesterhazy A, Kovacs K (1986) Breeding corn against Fusarial stalk rot, ear rot and seedling blight. Acta Phytopathol Entomol Hung 21:231–249

Parry DW, Bayles RA, Priestley RH (1984) Resistance of winter wheat varieties to ear blight (*Fusarium culmorum*). J Nat Inst Agric Bot 16:465–468

Schroeder HW, Christensen JJ (1963) Factors affecting resistance of wheat to scab caused by *Gibberella zeae*. Phytopathology 53:831–838

Scott PR, Hollins TW (1977) Interactions between cultivars of wheat and isolates of *Cercosporella herpotrichoides*. Trans Br Mycol Soc 69:397–403

Snijders CHA (1987) Interactions between winter wheat genotypes and isolates of *Fusarium culmorum*. Meded Fac Landbouwweten Rijksuniv Gent 52:807–814

Snijders CHA (1990) Genetic variation for resistance to *Fusarium* head blight in bread wheat. Euphytica 50:171–179

Snijders CHA, Perkowski J (1990) Effects of head blight caused by *Fusarium culmorum* on toxin content and weight of wheat kernels. Phytopathology 80:566–570

Steel RGD, Torrie JH (1981) Principles and procedures of statistics. A biometrical approach. McGraw-Hill Int, Singapore

Vanderplank JE (1984) Disease resistance in plants, 2nd edn. Academic Press, New York

Zobel RW, Wright MJ, Gauch HG (1988) Statistical analysis of a yield trial. Agron J 80:388–393

Chapter VII

Multiplicative models for genotype-environment interaction in plantbreeding

F.A. van Eeuwijk
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

Statistica Applicata Vol. 4, n. 4, 1992

MULTIPLICATIVE MODELS FOR GENOTYPE–ENVIRONMENT INTERACTION IN PLANT BREEDING

F.A. van Eeuwijk

Centre for Plant Breeding and Reproduction Research (CPRO–DLO), P.O. Box 16, 6700 AA Wageningen, The Netherlands.

In plant breeding genotypes are evaluated under varying environmental circumstances. Genotype–environment interaction is said to occur when genotypes respond differently to changes in the environment. Insight into this phenomenon is essential for progress in plant breeding. Many techniques have been developed for its analysis, usually assuming the data to be arranged in a two–way table of genotypes by environments. In this paper multiplicative models for the interaction are presented, which have not yet found broad application in plant breeding. These models are compared to the popular regression on the environmental mean model.

1. INTRODUCTION

Plant breeding is the science devoted to the development of new varieties. Research is initiated in response to changed demands on the existing varieties, ranging from higher yield to broader resistance to pathogens. To comply with a new standard the genetical constitution of plants is manipulated. The genetical constitution, or genotype, only partly determines the observable characteristics. Properties of organisms always are functions of both genotype and environment. The environment includes everything that is not genetic. Macro–climatological factors are among the most influential environmental factors, and not open to human manipulation. This is the reason why plant breeders emphasize the genetical approach. When genotypes respond differently to a change in the environment the phenomenon of genotype–environment interaction is said to occur. Genotype–environment interaction is of major importance in plant breeding, because its consequence is that the relative merits of genotypes depend on the environmental circumstances. In its most extreme form it can mean that a genotype A has higher yield than a genotype B on location 1, whereas B outyields A on location 2. For each genotype plant breeders have to identify the most influential environmental factors, and determine their mode of action. This must be the basis for decisions on the adaptedness of newly developed genotypes to future commercial growing environments.

Because data from plant breeding experiments are naturally arranged in a two–way table of genotypes by environments, techniques for the analysis of genotype–environment interaction usually have a two–way classification as a starting point. Cluster and ordination techniques are not uncommon in applications, but model based techniques departing from regression and singular value decomposition are more popular. Undoubtably the most used technique in applied plant breeding, and element of the second class, is the so–called regression on the environmental mean, commonly attributed to Yates and Cochran (1938). It describes the interaction by a bilinear form. Notwithstanding the fact that it is said to suffice in many instances, its value seems overrated. One aim of this paper is to illustrate that it can easily lead to an incomplete or incorrect description of genotype–environment interaction. Another aim is to present some alternatives which complement the regression on the environmental mean in some cases, or remedy its problems in other cases. In the next section the regression on the environmental mean and some closely related alternatives will be described. Some illustrations of the shortcomings of the regression on the environmental mean model and the wider applicability of the alternatives will be given in an Examples section.

2. MODELS FOR INTERACTION IN TWO–WAY TABLES

By convention data are arranged in a two–way table of genotypes by environments, and both genotypes and environments are taken fixed. An entry in the table, $Y_{ij.}$, is the mean over replications, k (k=1...K), of an observed variable, say yield, for a genotype i (i=1...I), evaluated under the j–th set (j=1...J) of environmental circumstances (e.g. a combination of year and location). The model with the most general formulation for the interaction is

$$Y_{ijk} = \mu + \rho_i + \gamma_j + \eta_{ij} + \epsilon_{ijk} \qquad (1)$$

In the two–way ANOVA model (1), μ denotes the general mean, ρ_i the genotypical main effect, γ_j the large–scale macro–environmental conditions, and η_{ij} the genotype–environment interaction. The ϵ_{ijk}'s are distributed independently as normal with mean zero and variance σ^2. They reflect 3 sources of variability; small–scale micro–environmental variability due to differences in the external environment, developmental variability, and measurement error.

In (1) each cell has its own interaction parameter. More parsimonious models are possible by writing the interaction as a bilinear form. A first example is the **regression on the environmental mean** model, which represents the data as a bundle of non–parallel lines (Yates and Cochran, 1938; Mandel, 1961; Finlay and

Wilkinson, 1963). Yield for the separate genotypes is regressed on the mean yield of all genotypes in a particular environment: in (1) η_{ij} is regressed on γ_j giving $\eta_{ij} = \beta_i\gamma_j + \delta_{ij}$, with β_i a linear regression coefficient for genotype i and δ_{ij} a deviation. Model (1) becomes

$$Y_{ijk} = \mu + \rho_i + (1+\beta_i)\gamma_j + \delta_{ij} + \epsilon_{ijk} \qquad (2)$$

Interaction here is simply heterogeneity of regression slopes. The main requirements for this model to be useful are that the responses of the genotypes to the environmental main effect indeed are linear, and that the deviations from regressions are comparably small, or not significant. This model was attractive, because it allowed plant breeders to predict yield for environments not in the experiment, in the presence of interaction. Furthermore, the description of the relevant environmental circumstances by the mean yield of the genotypes in an environment was intuitively appealing and did not require extra measurements on environmental variables.

A special case of (2) is the so–called **concurrence** model, in which the lines all intersect in the same point (Mandel, 1961; 1969). The model can be tested by Tukey's test for non–additivity (Tukey, 1949), and reads

$$Y_{ijk} = \mu + \rho_i + \gamma_j + c\rho_i\gamma_j + \delta_{ij} + \epsilon_{ijk} \qquad (3)$$

The model follows from (2) by taking the environmental sensitivity, β_i, as a constant, c, times the genotypical main effect, ρ_i.

Models (2) and (3) provide simple descriptions of the interaction between genotypes and environments by means of a one–dimensional representation of the environments to which genotypes are supposed to differ in sensitivity. However, in many instances the interaction will have higher dimensionality. A straightforward extension of (2) is given by the **additive main effects and multiplicative interaction effects** model (**AMMI**–model), which is simply a combination of ANOVA main effects for genotypes and environments, and multiplicative interaction effects obtained from a singular value decomposition of the matrix of residuals (Mandel, 1969; Perkins, 1972; Gauch, 1988). The model is written as

$$Y_{ijk} = \mu + \rho_i + \gamma_j + \sum_{l=1}^{L} u_{li}v_{lj} + \delta_{ij} + \epsilon_{ijk} \qquad (4)$$

The multiplicative terms, or scores, for the genotypes are given by the u_{li}'s, those for the environments by the v_{lj}'s, while L is the number of multiplicative terms needed for an adequate description of the interaction. Models (2) and (3) can easily be shown to be special cases of (4). Testing for interaction therefore often starts by determining L, and if L=1, subsequently a test for the appropriateness of (2), or

possibly (3), follows (Krishnaiah and Yochmowitz,1980). If L=1 and (4) is not reducible to (2) or (3), one might try to localize the interaction within parts of the two–way table. Localization of interaction for the cases L=1 and L=2 is well described in the text–book of Milliken and Johnson (1989), while a recent generalization is due to Goodman and Haberman (1990). In this context the possibility of using diagnostic biplots for a first approximation to an adequate model must also be mentioned (Bradu and Gabriel, 1978; Gower, 1990).

The genotypical and environmental scores belonging to a dimension in (4) can roughly be said to have one of 3 interpretations. The most common interpretation is that the environmental scores represent an environmental factor to which genotypes respond linearly, their sensitivities being expressed by the genotypical scores. For this 'regression' interpretation to be valid the environmental scores have to be dispersed evenly over the range. When the environmental scores can be divided into a cluster of positive values on the one hand, and a cluster of negative values on the other hand, for the moment ignoring near–zero values, the 'contrast' interpretation is valid. The limiting situation in which the near–zero cluster contains all but a few environmental scores, until the point that one of both non–zero clusters has no members left, provides the 'outlier' interpretation, if most of the genotypical scores are near–zero as well. Gnanadesikan and Kettenring (1972) showed that outliers visible in earlier dimensions are outliers that inflate variances and covariances, whereas observations with deviating covariance structure can be identified in later dimensions.

Models (2), (3), and (4) extract descriptions for the environments from the measurements in the two–way table. This can be sensible in the absence of specific measurements on environmental variables. However, having available extra information on the environments in the form of soil or climatological variables a method to be preferred is **redundancy analysis**, because it allows direct incorporation of this extra information. The redundancy analysis model can be written as

$$Y_{ijk} = \mu + \rho_i + \gamma_j + \sum_{l=1}^{L} u_{li} \left(\sum_{q=1}^{Q} d_{lq} x_{qj} \right) + \delta_{ij} + \epsilon_{ijk} \qquad (5)$$

Here x_{qj} denotes the value of the q–th environmental variable (q=1...Q) in the environment j, which for convenience can best be centred, and d_{lq} is the coefficient for that variable in the l–th redundancy axis. The other parameters are analogous to those in (4). The difference with (4) can be found in the restriction placed on the environmental scores of having to be linear combinations of measured environmental

variables. The model was introduced in plant breeding by Hardwick and Wood (1972), but has found only limited application. For details see Izenman (1980), Davies and Tso (1982), Denis (1988) and Van der Leeden (1990).

3. EXAMPLES

The data used in this section come from experiments carried out at the Centre for Plant Breeding and Reproduction Research (CPRO–DLO) in Wageningen, the Netherlands, in the years 1986, 1987, and 1988.

3.1. SEED YIELD IN PERENNIAL RYEGRASS

Perennial ryegras is the most important cultivated grass species in NW Europe. Though much effort went into characteristics as forage yield and turf quality, breeding for seed yield received little attention. Recently this situation has changed. For the experiment in this example the question was to investigate the influence of soil type (sand or clay) and year on the seed yield of 9 cultivars (genotypes) of perennial ryegrass. This was done by evaluating the 9 cultivars in 12 trials (environments) over the 3–year period 1986–1988 (Elgersma, 1990).

The ANOVA test for interaction, derivable from model (1), indicated the existence of a highly significant interaction between cultivar and trial ($p<0.001$), even though the percentage of the total sum of squares due to interaction was low; 4.5%. Significancy of the genotype–environment interaction following from the ANOVA test is common, because usually the number of degrees of freedom for interaction is large. However, it is often better to concentrate on relevancy, which may be measured by the relative size of the interaction sum of squares in comparison to the sum of squares for the genotypical main effect; 7.5%. So the interaction certainly was relevant. The environmental main effect, 88.0%, was considered not interesting, because the environments were consciously manipulated to cover the most extreme circumstances. An AMMI–model, (4), with 2 multiplicative terms for the interaction, explaining respectively 51.1 and 19.4% of the interaction sum of squares, was found to describe the interaction well. The rule was used that only those multiplicative terms would be retained that accounted for more than the average amount of interaction sum of squares to be expected in the absence of interaction (100/number of non–zero eigenvalues; Jolliffe, 1986). The necessity of 2 multiplicative terms for the interaction ruled out the adequacy of the regression on the environmental mean model, (2), and the concurrence model, (3), which performed very poorly anyway (13.5 and 1.7% of the interaction, both not even significant).

The cultivar scores for the first axis of the AMMI–model could be interpreted as a contrast between early and late maturing cultivars. This was remarkable as the cultivars were chosen to be homogeneous with respect to maturity, as this physiological variable is known to cause interaction. Regressions of the trial scores on individual meteorological variables were found significant only for minimum temperature in the period just before ear emergence and temperature sum (= mean temperature x length of period) during full flowering. Both variables explained the trial scores well, and were highly correlated with each other. However, the observations for temperature sum appeared to consist of only 2 clusters (Fig. 1a), thereby making a check on the linearity of the observed relation impossible. Observations for minimum temperature were more evenly dispersed (Fig. 1b).

For genotype–environment problems variable selection problems are difficult to solve on merely statistical grounds. Subsidiary physiological arguments are essential. In this case temperature sum suited physiological explanations best. However, the clumpedness of its values formed a drawback from a statistical point of view. The indecisiveness was tried to be resolved by using a method to relate environmental scores to environmental variables that avoided the assumption of linearity (Van Eeuwijk and Elgersma, submitted). Scale the environmental variables as the environmental scores (e.g. with mean zero, squared length unity), take care of maximal alignment of signs, and then check the variable values on their position with respect to the confidence intervals for the scores. Construct the confidence intervals following the prescripts given by Goodman and Haberman (1990). In Fig. 1a and 1b the abscissa already contains appropriately scaled values for the environmental variables. The procedure then amounts to searching for intersection of the confidence intervals with the line with slope unity. It remains difficult to choose a best explanatory variable. Further discussion in Van Eeuwijk and Elgersma (submitted).

With the above described interval matching method, genotypical and environmental scores for the second dimension of the interaction in the AMMI–model, were found to be in accordance with the slopes and environmental effects of (2). Mainly responsible for this dimension were 2 cultivars, which showed above average yield in the lowest yielding environment, and below average yield in the highest yielding environment. These cultivars also exhibited lower sensitivities in (2). The predominance of just 2 environments, positive and negative, excluded a regression interpretation for the AMMI–axis, and equally so for the regression on the environmental mean. The genotypical scores, or slopes, should not be used for predicting performance in intermediate environments.

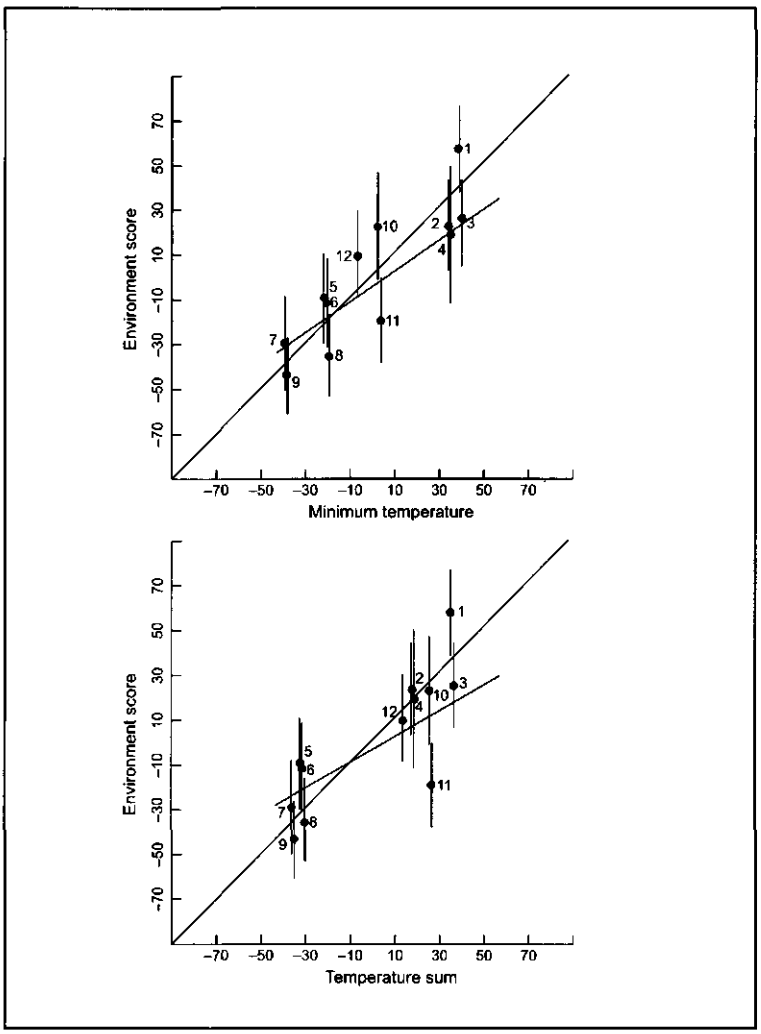


Fig. 1a: Scores and confidence limits for the environments 1 to 12, on the first axis, against scaled values of temperature sum. The line covering only part of the abscissa is the fitted regression line.

Fig 1b: As Fig. 1a but for minimum temperature.

3.2. NITRATE CONCENTRATION IN LETTUCE

High nitrate concentrations in lettuce are a problem when lettuce is grown under low light conditions. Inside the human body nitrate is converted into nitrite, which decreases the oxygen transporting capacity of the blood. To study the influence of annual variation of light intensity, 8 lettuce genotypes were evaluated in 18 trials (environments) from April 1987 until June 1988 (Reinink, 1991). An important question was whether genotypes interacted with time of the year, and how this interaction depended on a number of measured variables representing daylength, total global irradiation per day, and ion composition of the medium on which the lettuce was grown.

The ANOVA test for interaction was, as in the previous example, highly significant. Also the heterogeneity between regression lines in (2) now was significant ($p<0.001$). However, the deviations from regression were very significant too ($p<0.001$). Heterogeneity of slopes explained 19.8% of the interaction sum of squares, leaving 80.2% for the deviations. The latter was too high for considering regression on the environmental mean satisfactory. An AMMI–model with 2 multiplicative terms, 61.0% and 16.0%, did clearly better.

Redundancy analysis, (5), was used for further analysis of the interaction in relation to the measured variables (Van Eeuwijk, 1992). A forward selection procedure delivered 2 environmental variables as significantly contributing to the interaction. Maximum likelihood – (Van der Leeden, 1990) and randomization tests (Ter Braak, 1988) both led to the conclusion of significancy of 2 redundancy components, explaining respectively 54 and 8% of the interaction sum of squares. So, only the full rank model was adequate. Though a full rank multivariate multiple regression model was diagnosed, rewriting it in the form of redundancy components certainly facilitated interpretation of the interaction. The first redundancy component was the sum of daylength a month before harvest and daylength at harvest, the second the difference between them. Environmental scores for the first component were thus extreme in summer and winter, for the second in spring and autumn. In the biplot of Fig. 2, giving genotypical and environmental scores (scaled both to sum of squared scores equal to the corresponding singular value), it is easily seen that the interaction has a cyclical nature. The environments are numbered consecutively 1 to 18, their order in time, the genotypes **a** to **h**. As the experiment ran from spring 1987 until summer 1988, there were replications for the observations in spring and summer, and an internal check on the model was possible. Environments representing similar periods in 1987 and 1988 had to be situated near each other in the biplot. Correspondence was remarkably good, e.g. 3 was harvested on the

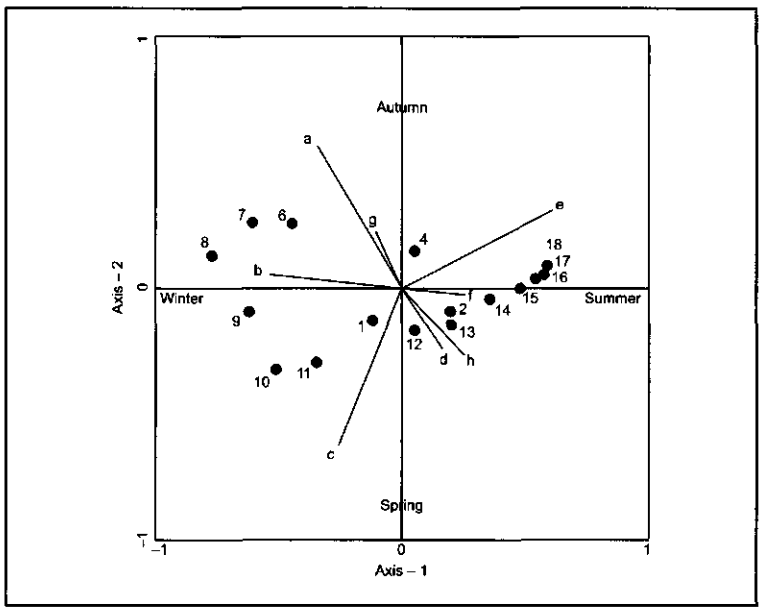


Fig. 2: Redundancy analysis biplot. Environments are indicated by the figures 1 to 18, genotypes by the letters a to h.

3rd of July 1987, whereas 16, 17 and 18 were from the 14th of June, the 20th of June, and the 3rd of July 1988. Genotype e was selected earlier by plant breeders for its extremely low nitrate concentrations under low light conditions (winter). This is in accordance with its position in the biplot. Redundancy analysis offers a generally appealing means to interpret genotype–environment interaction in the light of extra environmental information.

3.3. FUSARIUM HEAD BLIGHT INCIDENCE IN WINTER WHEAT

Fusarium head blight is a fungal disease causing not only head blight, but also root, foot and stem rot in a number of plant species like cereals, peas, and alfalfa. In winter wheat no complete resistance is known, meaning that all genotypes will show symptoms at (artificial) infection, but genetic variation in resistance has been demonstrated. The question for this example was whether resistance was horizontal or not. A definition of horizontal resistance is given by Vanderplank (1984). Here it suffices to formulate the concept as the non–existence of changes in rank order of

genotypes evaluated under various circumstances. Data on Fusarium head blight incidence (%) were collected on 17 winter wheat genotypes in the years 1986, 1987, and 1988. Each year the genotypes were artificially infected with 4 strains of the disease causing fungus. A two–way table of means over replications was constructed with the environment having 3 (years) x 4 (strains) = 12 levels.

Interaction was very prominent, 33.5% of the total sum of squares, a consequence of the retainment of the percentage scale for the measurements. A logit transformation, indicated by a score test for transformation within the Aranda–Ordaz family (Atkinson, 1985, chpt. 7), decreased the percentage interaction to 18.7%. However, after transformation still 3 components of the singular value decomposition remained necessary, and in this sense no reduction of the problem occurred. Besides, a logistic transformation would confer high weights to low percentages, which were hard to reconcile with the relatively great inaccuracies in the measuring procedure (Snijders and Van Eeuwijk, 1991). For these reasons the analysis was initially continued on the percentage scale.

Heterogeneity between genotypical regression slopes, (2), accounted for a high percentage of interaction, 41.4%, and was found to be equivalent to concurrence, (3). However, deviations from regression were also significant, indicating the need for a higher dimensional model. Three AMMI interaction components were found necessary. The first component resembled the interaction part of the concurrence model, and described the divergence of the genotypical responses. The second component modelled an interaction between a particular group of genotypes with a common ancestor and the most agressive strain of the fungus in 1987. The last component was not very interpretable, except as modelling a multiple outlier. Because the interaction was mainly due to concurrence, meaning no rank changes of genotypes, the question with respect to the type of resistance as horizontal could be answered affirmatively.

Analysis of variance on the logit transformed data, with a full factorial model including genotypes, years, and strains, showed no significant interaction to exist between genotypes and strains, whereas interaction between years and strains, and genotypes and years was significant. The conclusion of horizontal resistance again follows.

4. DISCUSSION

The arguments against the regression on the environmental mean have been summed up time and again, e.g. Westcott (1986). For it to work heterogeneity between slopes needs to be significant, and a substantial amount of the interaction sum of squares must be explained, while the deviations from regression should be relatively small, or not significant. In addition, the environmental scores must be evenly dispersed over the range. For the data sets in this paper, the model described part of the interaction in the ryegrass and wheat example, though it could not be conferred the regression status in the case of ryegrass, wheras it was of no value at all for the lettuce set. No problems arise as long as regression on the environmental mean is treated as a special case of the AMMI–model; an axis with environmental scores equal to the environmental main effect. The same rules for interpretation as described for the AMMI–model pertain. There is nothing inherently wrong with the model, it is just almost never the complete story. This was noted earlier by Perkins (1972) and Freeman and Dowker (1973), who also identified an AMMI–axis as the regression on the environmental mean. In both cases it concerned the first of a number of axes.

Another recurring theme in genotype–environment interaction is the choice of scale. Plant breeders tend to be very reluctant to accept other scales than the one on which the measurements have been collected. Retainment of a scale on which the measurements are non–normal can induce heterogeneity of variance. The problem then is how to disentangle interaction from heterogeneity of variance. An approach as for the wheat example, incorporating heterogeneity by multiplicative interaction terms was defended by Snee (1982), and Milliken and Johnson (1989). Others have the opinion that the only thing of importance in plant breeding is the occurrence of rank changes between genotypes over environments (Baker 1988), and emphasized the use of tests on rank reversal, as the one from Azzalini and Cox (1984). Use of individual scaling methods is not common in plant breeding, but would certainly merit more attention. At present the most promising direction seems independent modelling of interaction and variance in the form of generalized linear models (Pettitt, 1989).

Because of the emphasis on bilinear models for the analysis of genotype–environment problems 2 other main classes of analyses were done some injustice. Firstly the approach by variance components must be mentioned, whose origin may be attributed to Sprague and Federer (1951). In general one may be doubtful about fulfilment of its assumptions in plant breeding experiments, for instance sampling of years is done sequentially, so never really random, at best typically. In

addition, genotypes are often selected to some degree. The type of conclusions admitted by the variance components approach is also rather coarse. There still remains a long way to go if one only knows that 20% of the variance is due to genotype–environment interaction. Explicit modelling of interaction as in (reduced rank) regression models seems more prospectful, though more attention to the random character of particular environmental factors certainly is necessary. The second class of methods not dealt with concerns clustering methods. These are useful in cases where one does not want to assume a particular structure for the interaction. This approach seems especially suited to difficult to model interactions, such as occur frequently within resistance breeding.

For the analysis of genotype–environment interaction problems singular value based techniques have been shown to be a powerful and flexible tool, provided that the distinction between regression, contrast, and outlier interpretation of the dimensions is sufficiently appreciated. Incorporation of environmental variables can be done by means of a redundancy analysis. Alternatively, an AMMI analysis can be supplemented by interval matching of hypothetical scores to confidence intervals of the genotypical and/or environmental scores. The present paper may be seen to meet demands for more research on the use of environmental variables made by most notably Freeman (1973) and Westcott (1986). However, many problems are still unsolved, such as the selection of relevant environmental variables, as illustrated by the ryegrass example. Plant characteristics are the results of highly complex functional integrations of environmental circumstances over time, depending on the genetical code present in individual plants. In the development of characteristics, physiological periods can be distinguished. To unravel the causes of genotype–environment interaction critical periods need to be pinpointed in combination with critical variables within these periods. To model the development of a characteristic for a number of genotypes, taking into account physiological periods and critical variables, LISREL models seem to be the natural approach (Jöreskog and Sörbom, 1979). More attention should therefore be directed to these models, and, as stated above, to independent modelling of random and systematic components in the interaction. A synthesis of generalized linear models with multiplicative interaction models would be welcomed.

ACKNOWLEDGEMENTS

The author wishes to thank Ir. J. Jansen and Drs. T.J.J. Denteneer for critically reading earlier drafts of the manuscript, and Drs. L.C.P. Keizer for preparation of the manuscript.

REFERENCES

Atkinson A.C., 1985, *Plots, transformations, and regression.* Clarendon Press, Oxford.

Azzalini A., Cox D.R., 1984, Two new tests associated with analysis of variance. *J. R. Statist. Soc. B*, 46, 335–343.

Baker R.J., 1988, Tests for crossover genotype–environmental interactions. *Can. J. Plant Sci.*, 68, 405–410.

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

Davies P.T., Tso M. K–S., 1982, Procedures for reduced–rank regression. *Appl. Stat.*, 31, 244–255.

Denis J.B., 1988, Two way analysis using covariates. *Statistics*, 19, 123–132.

Elgersma A., 1990, Genetic variation for seed yield in perennial ryegrass, *Lolium perenne L. Plant Breeding*, 105, 117–125.

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

Freeman G.H., 1973, Statistical methods for the analysis of genotype–environment interactions. *Heredity*, 31, 339–354.

Freeman G.H., Dowker B.D., 1973, The analysis of variation between and within genotypes and environments. *Heredity*, 30, 97–109.

Gauch H.G. Jr., 1988, Model selection and validation for yield trials with interaction. *Biometrics*, 44, 705–715.

Gnanadesikan R., Kettenring J.R., 1972, Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics*, 28, 81–124.

Goodman L.A., Haberman S.J., 1990, The analysis of nonadditivity in two–way analysis of variance. *J. Am. St. Ass.*, 85, 139–145.

Gower J.C., 1990, Three–dimensional biplots. *Biometrika*, 77, 773–785.

Hardwick R.C., Wood J.T., 1972, Regression methods for studying genotype–environment interactions. *Heredity*, 28, 209–222.

Izenman A.J., 1980, Assessing dimensionality in multivariate regression. In: Krishnaiah P.R., ed., *Handbook of statistics*, Vol. 1. North–Holland Publishing Company, pp 571–591.

Jolliffe I.T., 1986, *Principal components analysis*, Springer Verlag, New York.

Jöreskog K.G., Sörbom D., 1979, *Advances in factor analysis and structural equation models*, Abt Books, Cambridge.

Krishnaiah P.R., Yochmowitz M.G., 1980, Inference on the structure of interaction in two–way classification model. In: Krishnaiah P.R., ed., *Handbook of statistics*, Vol. 1. North–Holland Publishing Company, pp 973–994.

Mandel J., 1961, Non–additivity in two–way analysis of variance. *J. Am. Stat. Ass.*, 56, 878–888.

Mandel J., 1969, The partitioning of interaction in analysis of variance. *J. Res. NBS.*, 73B, 309–328.

Milliken G.A., Johnson D.E., 1989, *Analysis of messy data*, Vol. 2, Nonreplicated experiments, Van Nostrand Reinhold, New York.

Perkins J.M., 1972, The principal components analysis of genotype environmental interactions and physical measures of the environment. *Heredity*, 29, 51–70.

Pettitt A.N., 1989, One degree of freedom for nonadditivity: applications with generalized linear models and link functions. Biometrics, 45, 1153–1162.

GENOTYPE BY ENVIRONMENT INTERACTION; BASIC IDEAS AND SELECTED TOPICS

Abstract

A reappraisal of concepts and models as used within quantitative genetics is given. The motivation for this reappraisal stems from the acknowledgement of a growing interest in the role of the environment in creating differences between genotypes. Heritability, being the central notion in classical quantitative genetics, is dealt with first. Three versions are distinguished, each with its own assumptions and range of applicability. The conclusion will be drawn that the usefulness of heritability depends too heavily on the over-simple traditional additive model, which sees the phenotypic value as just the sum of a genotypic and an environmental value. As an alternative, models based on the concept of reaction norm, the functional relationship between phenotype and environment over a range of environments for a particular genotype, are proposed. Recent literature on reaction norms is reviewed. Finally some consequences for human genetics are given. Throughout reference is made to the IQ debate to illustrate use, range, and merits of both the classical and the alternative approach.

Index terms: genotype by environment interaction, reaction norm, heritability, genetic cause, IQ.

1. Introduction

Traditionally quantitative genetics studies the inheritance of metric characters with a continuous range of variability, assumed to be created by a large number of genes with individually small effects in interaction with a broad spectrum of environmental influences. The predominant concept within classical quantitative genetics is the concept of heritability, for the moment to be defined as the proportion of observable variation in a character, i.e. the phenotypic variation, that is due to genetic differences between the individuals of a particular population, as opposed to differences due to environmental differences. Heritability is conditional on a reference population of genotypes on the one hand and a reference population of environments on the other hand. This follows from the fact that the value of a genotype is defined by the mean of its phenotypic values over the reference population of environments, whereas the value of an environment is defined by the mean of the phenotypic values over the reference population of genotypes. Heritability estimates are only meaningful with respect to these reference populations. Furthermore, there is the tacit assumption that genotypes translate changes in the environment in a similar way to their phenotypes. This is equivalent

to saying that genotype and environment act additively in their determination of the phenotype. The model underlying this assumption is P = G + E, with P for phenotypic value, G for genotypic value, and E for environmental value. The genotypic and environmental value are classically considered to be normally distributed and to be independent of each other.

The extremely simplifying nature of the additivity assumption for the genotypic and environmental value will be evident. In recent years the utility of a heritability concept based on additivity has been contested from two directions. Firstly, from those researchers who want to emphasize more heavily the importance of developmental processes for a good understanding of evolutionary processes. Secondly, from those who want more attention to be paid to the role of the environment in connection to the occurrence of differential reactions of genotypes to changes in the environment. The distinction between these two groups is not exclusive, in fact developmental influences are often interpreted or modelled as a special type of environmental influences.

A recent model which explicitly incorporated developmental and environmental effects was given by Cowley & Atchley (1992). Their model described a vector of phenotypic characters as an integration over time of the effects of four classes of controlling factors. Besides intrinsic genetic factors and environmental effects, which loosely resemble the classical genetic and environmental factors (Falconer, 1981), two developmental types of factors were taken into account; epigenetic factors and maternal factors. The first class concerns all those processes relating to the expression and the interaction of genetic materials; the influence of genes regulating one group of cells on a different group of cells. An example is embryonic induction. (A gene is roughly equal to a functional unit at the chromosome, i.e. a sequence of nucleotides at a certain place, locus, that codes for a certain product. An example is the gene for eye-colour. A gene can have several appearances, which are called alleles. Examples are the alleles for blue and brown eye-colour.). Maternal factors consist of the influences of a mother on her progeny beyond the direct transmission of her genes. The model of Cowley & Atchley makes explicit the dependence of the genetic and phenotypic covariance structure on developmental and environmental circumstances.

An example in the human context of a model that shows the dependency of these structures on the acculumulating effects in time of genes and environments was given by Eaves et al. (1988). They used a linear, multivariate, first order time series model and modelled genetic and phenotypic variances and covariances as a function of age. A population genetic model for the dependency of these variances and covariances on the environment was developed by De Jong (1990a). Individual allelic effects were modelled on the supposition that allelic effects contributing to any two traits were linear. The matrix of additive genetic variances and covariances then becomes a quadratic function of the environment.

Quantitative geneticists in various fields of application have started to integrate more realistic representations of development and the environment in their models. A major change in comparison to the classical additive model is the interest for the differences between genotypes in their development and reactions to the environment. The additive model emphasized similarity of genotypes, while nowadays the attention has shifted to differences. Genotypes exhibit differential developmental patterns and differential reactions in response

to changes in the environment. This conglomerate of differences can be caught in the concept of genotype by environment (GE) interaction, in which environment must be interpreted broadly as also encompassing a number of developmental processes. As a consequence of this shift in attention a reappraisal of basic concepts within the domain of quantitative genetics is necessary. An effort in that direction will be undertaken in this contribution. First a critical discussion of heritability will be given along a line indicated by Jacquard (1983). Then the concept of the reaction norm will be introduced. The reaction norm of a genotype may be interpreted as the functional dependence of the phenotype on the environmental circumstances for that particular genotype. It is a function specific to a genotype that translates the distribution of relevant environmental circumstances to that of the corresponding phenotypes. Subsequently the case for a quantitative genetics of reaction norms will be made followed by a birds-eye view of the state of the art on this topic. Finally some consequences of the use of models based on reaction norms for the area of human genetics are discussed.

2. Heritability

In a very illuminating paper Jacquard (1983) showed that the word heritability is used for three concepts. First there is heritability as a measure for empirical resemblance: biometrical heritability. It characterizes the relation between observations made on offspring and parents, and allows the prediction of offspring values from those of parents. Let X be a variable measurable on parents and offspring, which takes the value x_i for a set of fathers. For their sons X follows a distribution with a conditional mean $\overline{X}_i$ and a conditional variance V_i . In this case there is *resemblance* between fathers and sons if $\overline{X}_i$ lies between x_i and μ, the overall mean of X in the parent population. Algebraically we could write $\overline{X}_i = \mu + k\,(x_i - \mu)$, in which the conditional mean $\overline{X}_i$ is supposed to be a linear function of x_i, with k the measure of resemblance. For the offspring-one parent case the biometrical heritability, h^2, is equal to 2k. Knowledge of this type of heritability permits calculations concerning the conditional mean and the conditional variance of X. Jacquard gives the following example. Say for the character IQ the heritability is 0.80, the mean is 100 and the variance 225 (making a standard deviation of 15). Children of a parent with an IQ of 120 then have a phenotypic distribution with a conditional mean of $100 + \frac{1}{2}(0.80)(120 - 100) = 108$. The conditional variance will be $225\,[1 - \{\frac{1}{2}(0.80)\}^2] = 189$. The conditional standard deviation is 13.6, and is only slightly less than the unconditional value of 15 in the population as a whole, despite the high heritability. Though the accuracy of the prediction of the offspring value from the parent value increases with the value of h^2, it is easily overestimated by just looking at the size of h^2, as will be evident from the above example. There is absolutely nothing genetic about the model underlying the heritability here. No explanation whatsoever is offered for the observed resemblance.

A second type of heritability to be distinguished is the so called heritability in the broad sense, h_b^2, to be contrasted with the heritability in the narrow sense described below. Broad sense heritability purports to assess the relative contribution of genetic differences between individuals to the total phenotypic variation between individuals. The underlying idea

is that variation between individuals is partly due to genetic differences and partly to environmental differences: $\sigma_P^2 = \sigma_G^2 + \sigma_E^2$, with σ_P^2 the phenotypic variance, σ_G^2 the genotypic variance, and σ_E^2 the environmental variance. The broad sense heritability is not an attribute proper to a character as such, but is defined only for a character conditional on the reference set of genotypes and environments, and so may differ from one population to another, or from one environment to another. Furthermore, it is only meaningful in the absence of GE interaction and under the independence of genotypic and environmental values. Genotype-environment interaction occurs whenever genotypes react differentially to changes in the environment. This situation is quite common so that the condition of no interaction is often violated. The second condition of independence of genotypic and environmental value is also difficult to fulfil. It is violated for example, whenever genotypes with a high genotypic value also experience better environmental circumstances. In practice both conditions are hardly ever met. Nevertheless, it is this type of heritability which showed up in numerous discussions about the heritability of IQ in connection with the role of genes in the variance of this character. However, as we will see below, even when the reference populations of genotypes and environments are clearly defined and both conditions are fulfilled it still is not true that the separation of the phenotypic variance in genotypic and environmental variance actually separates the *causes* of variation.

There seems to be confusion about what constitutes a genetic cause. A distinction made by Gifford (1990) concerning two senses of the word 'genetic' may help to elucidate this confusion. The first and central sense is as follows: a trait is genetic if genetic *differences* between the individuals in a given population account for the phenotypic *differences* in that trait among members of that population. A second sense says that the trait must be the *specific* effect of some genetic cause, the trait must be described or individuated in such a way that it is properly matched to what the gene causes specifically.

In the IQ debate the second sense is intended when in fact the first sense is the only one accessible via broad sense heritability. One should note that high h_b^2, which refers to 'genetic' in the first sense, does not imply that a trait is unaffected by the environment, which is closer to 'genetic' in the second sense. High heritability does not imply unchangeability. The most bold statement in this context was made by Kempthorne (1978) who remarked: 'Heritability does not even exist in the human IQ context. Why then argue about the magnitude of an imaginary number?'

A third and most important concept corresponding to the word heritability is the so-called heritability in the narrow sense. It is used in animal and plant breeding to guide the choice of selection techniques in improving particular traits in the reference population. Again it does not provide insight in the biological mechanisms at work. In comparison to the situation above for the broad sense heritability, the genetic variance is subdivided into components attributable to the main effects of the contributing genes and interaction effects between these genes. Only the part of the genetic variation due to the gene main effects, the additive genetic variance, can be utilized within breeding programmes for improvement of the reference population of genotypes. Narrow sense heritability, h_n^2, is defined as the proportion of the additive genetic variance in the phenotypic variance. Broad sense heritability is mainly

useful as a step-up to narrow sense heritability. For illustrating the dramatical differences that may occur between h_n^2 and h_b^2 recessive genetic diseases can serve as an example. Since these are completely genetically determined $h_b^2 = 1$, whereas $h_n^2 = 2q / (1 + q)$, with q the frequency of the gene causing the trait. Though the narrow sense heritability is a popular device in breeding, for its application the same kind of conditions have to be fulfilled as for broad sense heritability. Especially the necessity of absence of GE interaction seems hard to reconcile with the practical situation. Therefore we will now turn our attention to the phenomenon of GE interaction, how it can be detected and modelled, and how it affects heritability. The central concept in connection to the phenomenon of GE interaction is the concept of reaction norm.

3. Reaction norms

In the Introduction reaction norm was already defined as the functional dependence of the phenotype for a particular genotype on the environment, or as the translation of the environmental distribution to the phenotypic distribution. Norm is to be understood here as derived from normative; how an individual 'should' behave according to its genetic make-up in reaction to the environment. There is no reference to norms in a statistical/mathematical sense. The environment is usually represented as a unidimensional quantity in the literature on reaction norms. The reason for this is mainly historical. It is a consequence of the traditional quantitative genetic subdivision of the phenotype in a genetic and environmental contribution. In the absence of more detailed information about the environmental circumstances this environmental contribution is then taken as a description of the environment. The most popular model for linear reaction norms consists of linear regressions of phenotypic responses for individual genotypes on the environmental contribution proper to an environment (Yates & Cochran, 1938; Finlay & Wilkinson 1963). There is, however, no need to restrict the environment to a unidimensional quantity and there are enough indications that this is an oversimplification (van Eeuwijk, 1992a, 1993). Nevertheless, as the arguments to be given below do not depend on the dimensionality of the environment, for ease of exposition unidimensionality is assumed unless stated otherwise.

The concept of reaction norm will now be dealt with in somewhat greater detail. In Figure 1 the reaction norms are given for two genotypes, G1 and G2. Both norms are linear and genotype G2 can be seen to be less susceptible to environmental changes than G1. Because neither genotype is superior over the entire environmental range, it can be concluded that GE interaction is present. The case in which both reaction norms would be parallel corresponds to additivity, a change in the environment would then cause an equivalent change in both phenotypes irrespective of the initial environmental circumstances. In Figure 1a the population of genotypes is assumed to consist of G1 and G2 in equal frequencies. The distribution of phenotypes is clearly bimodal and because the difference in phenotypic means between the genotypes, a measure for the genetic variance, is large in comparison to the average width of both the underlying phenotypic distributions, a measure for the environmental variance, h_b^2 is high. In Figure 1b the environmental distribution has been shifted to the left. The difference in means between the genotypes has become less. Also the

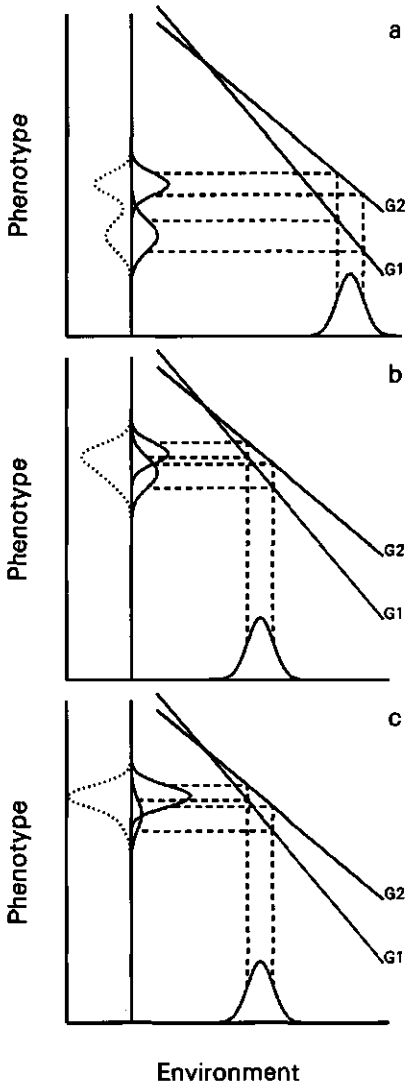


Figure 1. Illustration of how genotypes interpreted as reaction norms translate the environmental distribution to the phenotypical distribution. (a) Bimodality of phenotypical distribution, genotypes in equal frequencies, high broad sense heritability. (b) Shift in environmental distribution lowers genetic variance, genotypes still in equal frequencies. (c) Frequency of most stable genotype, G2, is increased, phenotypic variance decreases as a consequence.

amount of GE interaction will be less, because the reaction norms diverge less in this range. Important is that the amount of *genetic* variance has been altered by a shift in the *environment*. When we change the distribution of genotypes in the situation of Figure 1b by increasing the proportion of G2, the amount of environmental variance in the population as a whole decreases, because we enrich the population with the phenotypically more stable genotype. So we can alter the *environmental* variance by changing the *genotypic* distribution. This is what happens in Figure 1c. The lesson to be learnt from the simple pictures of Figure 1 is that genotype and environment produce phenotype in a continuous interaction over time and space, and there is no way in which an analysis of variance could identify components of variance with the actual causes of variance. In fact the relative size of genetic and environmental components of variance is dependent on (a) the form of the reaction norm, (b) the distribution of genotypes, (c) the distribution of environments, and (d) the correlation between genotypes and environments.

Though the dependency of genetic and GE interaction variance on the environment is most easily demonstrated with the help of reaction norms, the same kind of results can also be derived from considerations with respect to suitably defined reference populations of genotypes and environments. Comstock & Moll (1963) showed how restriction of the environmental range over which genotypes are tested can lead to an increase of the genetic variance and a decrease of the GE interaction variance. In contrast, widening of the environmental range can decrease genetic variance and increase GE interaction variance.

4. Quantifying non-parallelism of reaction norms

Non-parallelism of reaction norms can be encountered frequently in practice. Non-parallelism (unidimensional) can either be described in terms of convergence, divergence and intersection of reaction norms, or in terms of changes of rankings of genotypes over environments in combination with changes in the amount of variation (variance) between the genotypes. Whether one emphasizes the first or second type of description is dependent on the research question. Two situations have to be distinguished:

(a) One is interested in the performances of a number of selected genotypes under a set of unequivocally defined environmental circumstances. This situation is typical for the last stage of plant breeding programmes in which a number of interesting genotypes are evaluated under future commercial growing circumstances. The issue is then to describe as accurately as possible the individual reaction norms. Differences between reaction norms are expressed in terms of convergence, divergence and intersection.

(b) One is interested in a population of genotypes, and not so much in the performance of particular genotypes, with respect to a tightly or loosely defined set of environments. This situation occurs in the earlier stages of plant breeding programmes, in animal breeding, human genetics, and evolutionary biology. Only samples of the genotype population can be observed. It might still be profitable to describe individual reaction norms as in the above case, but this is seldom possible because of an overload of work. As a practical compromise the changes in rank order of the genotypes from one environment to another can be investigated. Because

it is also less clear in this situation how environments should be defined unequivocally, allowing at best nominal descriptions, an approach via the comparison of rank orders is the maximum achievable anyway.

Statistically, non-parallelism of reaction norms is equivalent to the existence of interaction in a two-way table of genotypes by environments. All methods appropriate for detecting interaction in such tables can thus also be used for detecting non-parallelism of reaction norms. For the situation in which the aim is to assess the individual norms of reaction for a selected set of genotypes under clearly defined circumstances, the usual start of an analysis is the two-way analysis of variance (ANOVA) of genotypes by environments, with both factors fixed. A test for non-parallelism then consists of the F-test of the mean square for interaction over an estimate for the error. Other tests for the interaction are possible depending on what model one wants to assume for the interaction. The two-way ANOVA model with interaction is the most simple but simultaneously least parsimonious model. More parsimonious and more complicated models for the interaction, uni- and higher dimensional, include regression models and models containing terms from a singular value decomposition of the observation matrix corrected for main effects. Reviews of these models are given in Westcott (1986) and van Eeuwijk (1992b), a review of tests is given in Krishnaiah & Yochmowitz (1980), approximate F-tests for a number of more complicated cases were derived by Goodman & Haberman (1990).

For the situation in which inferences have to be made for a population of genotypes, non-parallelism of reaction norms is preferentially reduced to changes in rank order of genotypes over environments. The same trait as expressed in different environments is thereby interpreted as a different trait in each environment: character-states. The (additive) genetic correlation, r_G, between the character-states in two environments reflects the common genetic basis of the expression of the trait in both environments. The genetic correlation between two character-states can be calculated from the genetic variances within the environments and the corresponding genetic covariance. However, when more than two environments are present this method can become quite cumbersome. An alternative is to calculate the so-called intra-class correlation, r_G', from a two-way ANOVA, which represents a lower limit to r_G (Yamada et al., 1988). The intra-class correlation coefficient is a function of the genetic and GE variance of the two-way table analysis. For the situation in which inferences have to be made for a population of genotypes and a population of environments, r_G' can be estimated by $\hat{\sigma}_G^2 / (\hat{\sigma}_G^2 + \hat{\sigma}_{GE}^2)$, with $\hat{\sigma}_G^2$ and $\hat{\sigma}_{GE}^2$ the estimates for the genetic and GE variance from the two-way ANOVA with both factors random.

When the environments represent a specific set, i.e. a fixed factor instead of random, inference and estimation is severely complicated by the fact that there are two commonly used competing two-way ANOVA models which can lead to different conclusions for the same data, when applied without caution. The first model, Model I, is generally attributed to Scheffé (1959) and has as its basic assumption that every genotype generates a multivariate normal vector variable with a specific mean in every environment and no restrictions on the covariance matrix. Main effects and interaction effects are defined in terms of the genotypic vector variables. On the environmental main effect, which is considered fixed, a sum zero

constraint is imposed. Let I be the number of environments, then the covariance matrix of the genotypic vector variables, Σ , is of the order I x I with the typical element $\{ \sigma_{ij} \}$ (i,j = 1,...,I). The genetic variance, σ_G^2 , is now defined as the mean entry of the covariance matrix, $\bar{\sigma}_{..}$. The interaction effects in Scheffé's model are correlated and the GE interaction variance, σ_{GE}^2 , is defined by $\sum_{I} (\sigma_{ii} - \bar{\sigma}_{..}) / (I - 1)$. Genotypic and GE interaction variance are thus defined to be linear functions of Σ.

The second model, Model II, can be derived from the first by imposing a structure upon Σ of the form $\sigma_{ii} = \sigma_g^2 + \sigma_{ge}^2$ and $\sigma_{ii'} = \sigma_g^2$ ($i \neq i'$), in which σ_g^2 denotes the variance as defined for the random variable g_j (j=1,...,J), and σ_{ge}^2 that for the random variable ge_{ij}. The random variables g_j and ge_{ij} are defined uncorrelated and having zero mean. The relation between the variance components in Model I and Model II is $\sigma_G^2 = \sigma_g^2 + \sigma_{ge}^2/I$ and $\sigma_{GE}^2 = \sigma_{ge}^2$ (Hocking, 1973).

A difference between both models to which ample attention has been paid in the past is the difference in the expected mean squares (EMS) for the genotypes. For the Scheffé model the EMS for the genotypes contains only the variance components for error and genotypes and not the component for the GE interaction, whereas in Model II the latter is present as well. The consequence is that in Model I the mean square (MS) for the genotypic main effect is tested over the interaction mean square, whereas in Model II the error mean square is used. Depending on the model chosen one can then find significance or non-significance for the genotypic variance in the same data (Ayres & Thomas, 1990).

However, the choice between Model I and II should not be arbitrairy or based upon the statistical package available. Whether one chooses Model I or Model II depends first on the fulfilment of the structural condition for Σ, which is necessary before application of Model II. Secondly, and not less important, it depends on whether σ_G^2 or σ_g^2 is the parameter of interest with respect to the research question one is trying to answer.

Recently Fry (1992) showed that the tests for the genotypic main effects in both models should correspond to different research questions. The genetic variance, σ_g^2 , of Model II must be read as the *covariance* in means for a particular genotype across environments, which can take on positive as well as negative values. Writing the EMS table in terms of the genetic correlation over environments following Robertson (1959), Fry showed how the F-test of MS_g over MS_{ge} provides a two-tailed test for testing whether $\sigma_g^2 > 0$ and $\sigma_g^2 < 0$, the null hypothesis being $\sigma_g^2 = 0$. This is equivalent to a test for rejecting $r_G' = 0$. The F-test of MS_{ge} over MS_{error} tests the hypothesis whether r_G' deviates from 1, which represents perfect additivity of genotype and environment and parallelism of reaction norms. For two environments it can be shown that $\sigma_{ge}^2 = \sigma_g^2 (1 - r_G)$ (Yamada, 1962), assuming the genetic variance to be equal in both environments (for two environments $r_G = r_G'$). The GE interaction variance is a monotone function of r_G. When $r_G = 1$, GE interaction is zero. The GE interaction increases via $\sigma_{ge}^2 = \sigma_g^2$ for $r_G = 0$, to the maximum $\sigma_{ge}^2 = 2\sigma_g^2$ for $r_G = -1$.

The opposition against the use of the GE interaction as a measure for the non-parallelism of reaction norms originates from the fact that it does not provide an accurate picture of the potential for evolution of reaction norms (Via, 1987). The greatest opportunity for change exists when the genetic correlation is zero (Figure 2b), GE interaction is then at

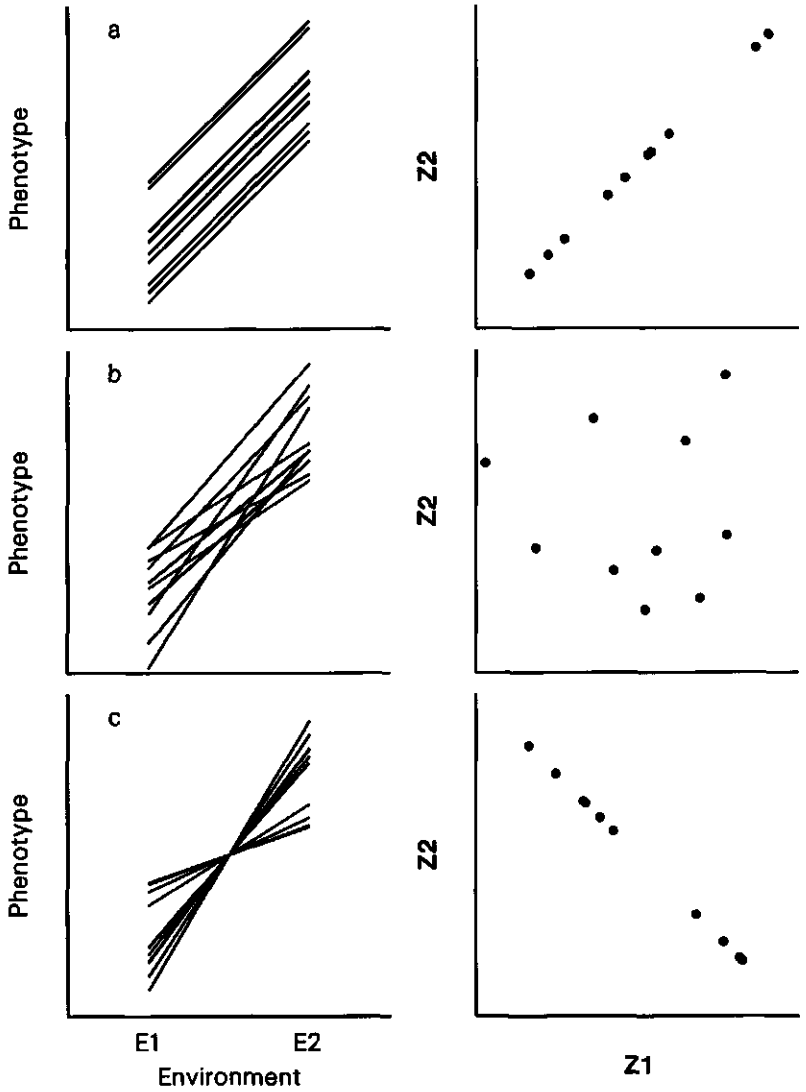


Figure 2. Relation between GE interaction and genetic correlation between environments (left: reaction norms for a number of genotypes across the environments E1 and E2, reflecting interaction; right: phenotypical values in E2, i.e. Z2, plotted against those in E1, i.e. Z1, reflecting genetic correlation): (a) additivity, no interaction, genetic correlation 1, (b) interaction is present, genetic correlation 0, (c) maximal interaction, genetic correlation -1.

an intermediate value. Independent evolution within environments becomes impossible when the genetic correlation reaches the values +1 (Figure 2a) or -1 (Figure 2c). The GE interaction is, however, at a maximum for $r_G = -1$. Returning to the genetic variance of the Scheffé model, σ_G^2, Fry remarked that this is the natural quantity to look at when a population of genotypes is evaluated in a defined mixture of environments, and one is interested in the variance of the marginal means, the genotypic means, over this particular mixture. An example is the evaluation of the genetic variance of a plant pathogen with respect to a specified set of host plants. By taking into account the biological question, a sensible choice can be made between the two contesting mixed model formulations. Note that the estimate for the total phenotypic variance also depends on the choice of model and with it the estimate for the heritability.

The test for the genetic effect in the random two-way ANOVA can be interpreted as a test of whether genotypes differ in their marginal means over the entire population of environments from which the environments in the study are sampled. Under the assumptions of the fully random model the variance of the marginal means, the genetic variance, must be equal to the genetic covariance over all pairs of environments. This assumption surely needs checking before application of the model (Fry, 1992).

5. Evolutionary models for reaction norms

The form of the reaction norm itself can be the target of evolutionary forces. Traditional quantitative genetics is built around the notion of heritability, with the basic assumption of additivity of genetic and environmental contribution in the phenotype. Additivity means parallelism in terms of reaction norms. Parallelism may be expected to be the exception, non-parallelism, and so interaction between genotype and environment, is probably far more common. Quantitative genetic models which purport to predict future distributions of genotypes and phenotypes on basis of empirical information about natural or experimental populations must therefore provide easy means of incorporating a diversity of reaction norms.

A few models developed towards this aim will be treated. A straightforward multivariate generalization of the classical univariate evolutionary response model, $R = h^2S$, with R the response to selection (difference in mean between offspring of selected and unselected parents), h^2 the heritability, and S the selection differential (difference in mean between selected and unselected parents), was given by Lande (1979) as $\delta \mathbf{z} = \mathbf{G}\mathbf{P}^{-1}\mathbf{s}$ (Lande's notation), with $\delta \mathbf{z}$ a vector of responses to selection for a number of variables, $\mathbf{G}$ the genetic covariance matrix, $\mathbf{P}^{-1}$ the inverse of the phenotypic covariance matrix, and $\mathbf{s}$ a vector of selection differentials. For the genotype by environment case the character-states are introduced in this multivariate equation as if they were characters of their own. The genetic correlations automatically find their place in the right hand side of the equation (Via & Lande, 1985). The response to selection for a particular character-state is the sum of direct selection on the character in that specific environment, a function of the heritability in that environment, plus indirect selection on the character in the other environments, a function of the

heritabilities in those environments together with the genetic correlation between the other environments and the target environment. Negative correlations between environments can retard evolution and keep the population away from an equilibrium. The approach of Via and Lande is especially suitable for discrete environments.

Gomulkiewicz & Kirkpatrick (1992) developed a model that can deal with continuous environments. Their infinite-dimensional model contains covariance functions instead of covariances, but is otherwise very similar to Lande's multivariate model. Also the approach chosen by Gavrilets & Scheiner (1993a, 1993b) is not essentially different from that of Via & Lande (1985). They describe reaction norms by polynomials and input the coefficients of these polynomials in Lande's multivariate evolution model. Really different is the approach by De Jong (1990a, 1990b). She introduces population genetic models in which the allelic effects themselves are linearly dependent on an environmental variable. This in contrast to the Falconer approach (1952) used by, among others, Via & Lande (1985), where alleles are either switched on or off, depending on the environment. The genetic correlation between two environments is then the reflection of the amount of alleles that are switched on in both environments. For De Jong's models the genetic correlation is a function of the intercepts and slopes that characterize the individual allelic effects.

6. Some consequences for human genetics

In human genetics the problem of GE interaction is very prominent. It is hardly imaginable that the complex interactions (in the every day sense) of people with their social and physical environments would not create differential reactions in their phenotypes. People very probably possess wildly different reaction norms for a broad spectrum of behavioural traits as, for example, IQ. The main issue of the IQ debate is to what extent educational intervention programmes can improve the intellectual capacities of certain retarded groups. This is mainly a question about the form of the reaction norms for such groups; their mean reaction norms and the amount and structure of variation within the groups. Quite mistakenly the issue has been tried to be settled by research in the heritability of IQ. Besides the fact that great variation in reaction norms plus equivocality of the reference population and almost surely existing genotype-environment correlation make the heritability for IQ a meaningless quantity anyway, it is the *flexibility* within one generation that we are after and not the inheritance from one generation to the next.

Quantitative genetic models that include vertical transmission for human behavioural traits typically consist of path models in which the phenotypes of the offspring are partly determined by the parental phenotypes, by means of their influence on the environment for the offspring (see, among others, Carey, 1991, and Cardon, Fulker, & Jöreskog, 1991). A simple extension of the classical model $P_o=G_o+E_o$, with P_o , G_o , and E_o the phenotype, genotype, and environment for the offspring can serve as an example. Just by making E_o a linear function of the maternal phenotype, P_m , and the paternal phenotype, P_f , we obtain the desired result: $E_o=b_mP_m + b_fP_f$, with b_m and b_f regression weights for mother and father. The genotype for the offspring is, of course, $G_o= 0.5G_m + 0.5G_f$, with G_m and G_f the maternal and

paternal genotype. Genotype by environment interaction for the offspring is incorporated in this model via the covariance between genotype and phenotype for the parents. Though deviating in an important sense from the classical additive genetic model, this type of model is still more related to the additive model than to models based on reaction norms. Criticisms expressed to it come from two directions:

(a) In general no explicit characterizations of the environment are taken into account. The only data that enter the model are phenotypic observations on the character under study. No effort is made to link the values for that character with other environmental or phenotypic variables. This can lead to severe interpretation problems (van Eeuwijk, 1992b).

(b) All criticisms that are expressed towards path models in general apply to the specific class of quantitative genetic path models as well. Most notably Kempthorne (1978) denies these models any causal interpretation. Causality should only be inferred from conscious intervention (see also Holland, 1986).

Within human genetics the need for a quantitative genetics of reaction norms has already been recognized. For example Eaves et al. (1988) asked for models which allow for the developmental genetic control of sensitivity (reaction norm) to environmental input. There seems to be no good argument why the analysis of human behavioural traits should not also be grounded in a study of reaction norms.

Acknowledgement

Paul Keizer is thanked for his help in the preparation of the manuscript.

References

Ayres, P.A. & Thomas, D.L. (1990) Alternative formulations of the mixed-model anova applied to quantitative genetics. *Evolution, 44,* 221-226.

Cardon, L.R., Fulker, D.W. & Jöreskog, K.G. (1991) A LISREL 8 model with constrained parameters for twin and adoptive families. *Behavior Genetics, 21,* 327-350.

Carey, G. (1991) Evolution and path models in human behavioral genetics. *Behavior Genetics, 21,* 433-444.

Comstock, R.E. & Moll, R.H. (1963) Genotype-environment interactions. *Statistical Genetics and Plantbreeding* (pp. 164-196). Nat. Acad. of Sc. Pub. 982.

Cowley, D.E. & Atchley, W.R. (1992) Quantitative genetic models for development, epigenetic selection, and phenotypic evolution. *Evolution, 46,* 495-518.

De Jong, G. (1990a) Quantitative genetics of reaction norms. *Journal of Evolutionary Biology 3,* 447-468.

De Jong, G. (1990b) Genotype-by-environment interaction and the genetic covariance between environments: multilocus genetics. *Genetica 81,* 171-177.

Eaves, L.J., Hewitt, J.K. & Heath, A.C. (1988) The quantitative genetic study of human developmental change: A model and its limitations. In B.S. Weir, E.J. Eisen, M.M. Goodman & G. Namkoong (Eds.), *Proceedings of the Second International Conference on Quantitative Genetics* (pp. 279-311). Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Falconer, D.S. (1952) The problem of environment and selection. *American Naturalist 86,* 293-298.

Falconer, D.S. (1981). *Introduction to quantitative genetics* (pp. 340). 2nd. edn. Longman, New York.

Finlay K.W. & Wilkinson, G.N. (1963) The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research, 14,* 742-754.

Fry, J.D. (1992) The mixed-model analysis of variance applied to quantitative genetics: Biological

meaning of the parameters. *Evolution, 46*, 540-550.

Gavrilets, S. & Scheiner, S. (1993a) The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *Journal of Evolutionary Biology*, in press.

Gavrilets, S. & Scheiner, S. (1993b) The genetics of phenotypic plasticity. VI. Theoretical predictions for directional selection. *Journal of Evolutionary Biology*, in press.

Gifford, F. (1990) Genetic traits. *Biology and Philosophy 5*, 327-347.

Gomulkiewicz, R. & Kirkpatrick, M. (1992) Quantitative genetics and the evolution of reaction norms. *Evolution, 46*, 390-411.

Goodman, L.A. & Haberman, S.J. (1990) The analysis of nonadditivity in two-way analysis of variance. *Journal of the American Statistical Association, 85*, 139-145.

Hocking, R.R. (1973) A discussion of the two-way mixed model. *The American Statistician, 27*, 148-152.

Holland, P.W. (1986) Statistics and causal inference. *Journal of the American Statistical Association, 81*, 945-960.

Jacquard, A. (1983) Heritability: One word, three concepts. *Biometrics 39*, 465-477.

Kempthorne, O. (1978) Logical, epistemological and statistical aspects of nature-nurture data interpretation. *Biometrics 34*, 1-23.

Krishnaiah, P.R. & Yochmowitz, M.G. (1980) Inference on the structure of interaction in two-way classification model. In P.R. Krishnaiah (Ed.) *Handbook of statistics* (Vol. 1. pp.973-994). North-Holland Publishing Company.

Lande, R. (1979) Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution 33*, 403-416.

Robertson A. (1959) The sampling variance of the genetic correlation coefficient. *Biometrics 15*, 469-485.

Scheffé, H. (1959) *The analysis of variance*. John Wiley and Sons, New York.

van Eeuwijk, F.A. (1992a) Interpreting genotype-by-environment interaction using redundancy analysis. *Theoretical and Applied Genetics, 85*, 89-100.

van Eeuwijk, F.A. (1992b) Multiplicative models for genotype - environment interaction in plant breeding. *Statistica Applicata, 4*, 7-20.

van Eeuwijk, F.A. & Elgersma, A. (1993) Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. *Heredity, 70*, 447-457.

Via, S. (1987) Genetic constraints on the evolution of phenotypic plasticity. In V. Loeschcke (Ed.), *Genetic Constraints on Adaptive Evolution* (pp. 47-71). Springer-Verlag, Berlin.

Via, S. & Lande, R. (1985) Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution 39*, 505-523.

Westcott, B. (1986) Some methods of analysing genotype-environment interaction. *Heredity, 56*, 243-253.

Yamada, Y. (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. *Japanese Journal of Genetics, 37*, 498-509.

Yamada, Y., Itoh Y. & Sugimoto, I. (1988) Parametric relationships between genotype x environment interaction and genetic correlation when two environments are involved. *Theoretical and Applied Genetics, 76*, 850-854.

Yates F. & Cochran, W.G. (1938) The analysis of groups of experiments. *Journal of Agricultural Science Cambridge, 28*, 556-580.

Chapter IX

On the use of diagnostic biplots in model screening for genotype by environment tables

F.A. van Eeuwijk & L.C.P. Keizer
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

tatistics and Computing (1995) 5, 141–153

On the use of diagnostic biplots in model screening for genotype by environment ables

. A. VAN EEUWIJK and L. C. P. KEIZER

LO-Centre for Plant Breeding and Reproduction Research, (CPRO-DLO), P.O. Box 16, 'ageningen, The Netherlands

ubmitted August 1994 and accepted August 1994

opular rank-2 and rank-3 models for two-way tables have geometrical properties which can be ed as diagnostic keys in screening for an appropriate model. Row and column levels of two-ay tables are represented by points in two or three dimensional space, whereupon collinearity d coplanarity of row and column points provide diagnostic keys for informal model choice. oordinates are obtained from a factorization of the two-way table **Y** in the matrix product $\mathbf{V}^{\mathrm{T}}$. The rows of **U** then contain row-point coordinates and the rows of **V** column-point ordinates. Illustrations of applications of diagnostic biplots in the literature were restricted data from chemistry and physics with little or no noise. In plant breeding, two-way tables ntaining substantial amounts of noise regularly arise in the form of genotype by environ-ent tables. To investigate the usefulness of diagnostic biplots for model screening for geno-pe by environment tables, data tables were generated from a range of two-way models der the addition of various amounts of noise. Chances for correct diagnosis of the generat-g model depended on the type of model. Diagnostic biplots on their own do not seem to pro-de a sufficient means for model selection for genotype by environment tables, but in mbination with other methods they certainly can provide extra insight into the structure the data.

ywords: AMMI, biadditive model, bilinear model, concurrence model, genotype by environ-nt interaction, multiplicative interaction, row regression

Introduction

statistical practice two-way tables of data show up equently as temporary or final summaries of research. plant breeding two-way tables even form a major means data summarization, as research here consists to a large tent in the evaluation of a set of genotypes under a iscretized) range of environmental conditions. Statistical odels for genotype by environment two-way tables that not include genotypic or environmental covariates dom exceed rank 3. These models are all members of a ss of models to which preferably the predicate additive is attributed (Denis and Gower, 1992, 1994), though bilinear may also be acceptable (Denis, 1991). e structure of the models allows the use of graphical ethods to assist model selection, this as an alternative to e exhaustive fitting of a sequence of models. For the latter, Seyedsadr and Cornelius (1992a) provide inferential methodology.

A minimal model for the two-way table is the additive model $\mathscr{E}(\mathbf{Y}_{ij}) = \mu + a_i + b_j$, where $\mathscr{E}(\mathbf{Y}_{ij})$ is the expectation of the random variable **Y** for the combination of the ith row factor level (genotype i) with the jth column factor level (environment j), μ is the general mean, and a_i and b_j are the main effects for row i and column j (with sum-to-zero identification constraints). One way of graphically diagnosing an additive model is to plot the observation y_{ij} against $\hat{a}_i$ (or $y_{i\cdot}$, the corresponding mean) for each j. Parallel lines then indicate additivity. Plotting y_{ij} against $\hat{a}_i(y_{i\cdot})$ for each j and $\hat{b}_j(y_{\cdot j})$ for each i, enables identification of a number of rank-2 models for two-way tables (see Milliken and Johnson, 1989).

A graphical method that uses special geometrical properties of two-way models was introduced by Bradu and

Gabriel (1976) under the name of diagnostic biplots. A $p \times q$ table of observations, $\mathbf{Y}$, is factorized as a matrix product $\mathbf{UV}^T$. The rows of U and V contain coordinates, referring to orthogonal axes, of points representing the rows and columns, respectively, of Y. The observation y_{ij} can be reconstituted from the inner product between the ith row of U and the jth row of V. The factorization used is immaterial as for every non-singular matrix T, Y can be factorized as $(\mathbf{UT})(\mathbf{T}^{-1}\mathbf{V}^T)$. The row and column coordinates determine the position of the points in two- or three-dimensional displays. For rank-2 models collinearity of row and/or column points provides diagnostic keys to the kind of models meriting closer consideration. For rank-3 models this is done using coplanarity in addition to collinearity (Bradu and Gabriel, 1978; Gabriel and Odoroff, 1986). Explicit equations for relevant lines and planes were derived by Gower (1990).

Though the theory behind the use of diagnostic biplots is relatively straightforward and the interpretational rules easy, their use in practice seems limited. In this paper their use will be investigated for the diagnosis of models for genotype by environment tables as occurring in plant breeding. The necessity for this research is created by the fact that without exception illustrations of diagnostic biplots are either given for data generated by exactly the model to be diagnosed, or for practical data stemming from astronomy, chemistry and physics, with little noise. The absence of serious noise contrasts sharply with the usual situation in plant breeding. For evaluating the utility of diagnostic biplots in this domain, two-way tables were simulated with a distribution of variation over main effects, interaction, and noise thought to be representative of plant breeding practices. The size of the simulated tables was throughout chosen to be 10 rows (genotypes) by 8 columns (environments).

2. Models and keys

Table 1 summarizes the models that were investigated, together with their diagnostic keys (Gabriel and Odoroff, 1986). Reformulations demonstrate more clearly the rank of the models. Greek letters denote scalar parameters, and bold lower-case letters denote vectors of parameters. The symbol $\mathbf{1}_p(\mathbf{1}_q)$ denotes a vector of $p(q)$ ones. Identification constraints imposed are the standard ones of $\mathbf{1}_p^T\mathbf{a} = \mathbf{1}_q^T\mathbf{b} = \mathbf{1}_p^T\mathbf{c} = \mathbf{1}_q^T\mathbf{d} = 0$, $\mathbf{1}_p^T\mathbf{g} = p$, and a restriction on the length of either or both of $\mathbf{c}$ and $\mathbf{d}$, e.g. $\mathbf{d}^T\mathbf{d} = 1$. Furthermore, for the row-regression-plus-one model $\mathbf{g}^T\mathbf{c} = \mathbf{b}^T\mathbf{d} = 0$. In principle, the vectors $\mathbf{r}$, $\mathbf{s}$, $\mathbf{u}$, and $\mathbf{v}$ contain unconstrained parameters, although length restrictions may be imposed on one of both vectors within the pairs $\mathbf{r}$ and $\mathbf{s}$, and $\mathbf{u}$ and $\mathbf{v}$. As the sizes of matrices and lengths of vectors throughout can be derived from the context, they will only be given when confusion could arise.

Table 1. *Models for two-way tables, reformulations to show rank of the model, and diagnostic keys based on collinearity* (L) *and coplanarity* (P) *of the row and/or column points.*

Model (re)formulation(s)	Rows	Columns	Relation
Additivity	L	L	orthogonal
$\mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{1}_p\mathbf{b}^T$			
$(\mu\mathbf{1} + \mathbf{a})\mathbf{1}^T + \mathbf{1}\mathbf{b}^T$ or $\mathbf{1}(\mu\mathbf{1}^T + \mathbf{b}^T) + \mathbf{a}\mathbf{1}^T$			
Concurrence	L	L	non-orthogonal
$\mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{1}_p\mathbf{b}^T + \nu\mathbf{a}\mathbf{b}^T$			
$(\mu\mathbf{1} + \mathbf{a})\mathbf{1}^T + (\mathbf{1} + \nu\mathbf{a})\mathbf{b}^T$ or $\mathbf{1}(\mu\mathbf{1}^T + \mathbf{b}^T) + \mathbf{a}(\mathbf{1}^T + \nu\mathbf{b}^T)$ or $\lambda\mathbf{1}\mathbf{1}^T + \mathbf{r}\mathbf{s}^T$			
Row regression	–	L	–
$\mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{g}\mathbf{b}^T$			
$(\mu\mathbf{1} + \mathbf{a})\mathbf{1}^T + \mathbf{g}\mathbf{b}^T$			
Additivity-plus-one	P	P	orthogonal
$\mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{1}_p\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$			
$(\mu\mathbf{1} + \mathbf{a})\mathbf{1}^T + \mathbf{1}\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$ or $\mathbf{1}(\mu\mathbf{1}^T + \mathbf{b}^T) + \mathbf{a}\mathbf{1}^T + \mathbf{c}\mathbf{d}^T$			
Concurrence-plus-one	P	P	non-orthogonal
$\mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{1}_p\mathbf{b}^T + \nu\mathbf{a}\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$			
$(\mu\mathbf{1} + \mathbf{a})\mathbf{1}^T + (\mathbf{1} + \nu\mathbf{a})\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$ or $\mathbf{1}(\mu\mathbf{1}^T + \mathbf{b}^T) + \mathbf{a}(\mathbf{1}^T + \nu\mathbf{b}^T) + \mathbf{c}\mathbf{d}^T$ or $\lambda\mathbf{1}\mathbf{1}^T + \mathbf{r}\mathbf{s}^T + \mathbf{u}\mathbf{v}^T$			
Row-regression-plus-one	–	P	–
$\mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{g}\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$			
$(\mu\mathbf{1} + \mathbf{a})\mathbf{1}^T + \mathbf{g}\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$			

2.1. *Rank-2 models*

The reformulations of the additive model in Table 1 show its rank-2 character. This entails that the table $\mathbf{Y}_{pq}$ can be written as the product of a row coordinates matrix $\mathbf{U}_{p\times 2}$ with the transpose of a column coordinates matrix $\mathbf{V}_{q\times 2}$, leading to a planar representation of Y. Choosing $\mathbf{U}_{p\times 2} = (\mu\mathbf{1}_p + \mathbf{a}, \mathbf{1}_p)$ and $\mathbf{V}_{q\times 2} = (\mathbf{1}_q, \mathbf{b})$, the row points and the column points can be seen to be collinear on two orthogonal lines. Another more elaborate argument leading to the same conclusion is the following. From the reformulations of the additive model in Table 1 it will be evident that the row space of Y contains $\mathbf{1}_q$, and the column space of Y must contain $\mathbf{1}_p$. The spaces spanned by the rows of Y and the columns of V must be the same, just as the space spanned by the columns of Y is equal to that spanned by the columns of U. Thus there exists a vector $\mathbf{m}$ and a scalar ϕ, so that $\mathbf{V}\mathbf{m} = \phi\mathbf{1}_q$, being the equation for a line in $\mathbb{R}^2$. Hence, the column points will be on a line. Analogously, there must exist a vector $\mathbf{n}$ and a scalar φ so that $\mathbf{U}\mathbf{n} = \varphi\mathbf{1}_p$, implying collinearity of the row points. The relation between the row-points line and the column-points line can be derived from the knowledge that in the case of additivity the following relation holds for any quadruple i, k, j, l; $y_{ij} - y_{il} - y_{kj} + y_{kl} = 0$. As the inner-product

between the ith row of $\mathbf{U}$ and the jth row of $\mathbf{V}$ should recover y_{ij}, $y_{ij} = \mathbf{u}_i \cdot \mathbf{v}_j$, also the relation $(\mathbf{u}_i - \mathbf{u}_k) \cdot (\mathbf{v}_j - \mathbf{v}_l) = 0$ must hold. Which means that the line segment joining any two row points is orthogonal to the line segment joining any two column points. Because all row points are on a line and all column points are on another line, it follows that the diagnostic key for an additive model consists in the observation of two orthogonal lines.

The concurrence model (Table 1) can be interpreted as the regression formulation of Tukey's one-degree-of-freedom for non-additivity test (Tukey, 1949). Using the most common model parametrization, $\mathcal{E}(\mathbf{Y}_{ij}) = \mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{1}_p\mathbf{b}^T + \nu\mathbf{a}\mathbf{b}^T$, the diagnostic key for the concurrence model follows as row- and column-points collinearity, just as for the additive model, in combination with non-orthogonal intersection. Recently interest has arisen in a constraints-free reformulation of the concurrence model; $\mathcal{E}(\mathbf{Y}_{ij}) = \lambda\mathbf{1}_p\mathbf{1}_q^T + \mathbf{r}\mathbf{s}^T$ (Cornelius *et al.*, 1992; Denis and Gower, 1992; Seyedsadr and Cornelius, 1992a, b). An appealing decomposition of $\mathbf{Y}$ would have $\mathbf{U}$ taking the form $(\lambda_1^*\mathbf{1}_p, \mathbf{r})$, and $\mathbf{V}$ the form $(\lambda_2^*\mathbf{1}_q, \mathbf{s})$, with $\lambda_1^*\lambda_2^* = \lambda$. In that case the diagnostic key would consist either in parallel lines ($\lambda_1^* \neq \lambda_2^*$), or coinciding lines ($\lambda_1^* = \lambda_2^*$). Because this formulation of the concurrence model requires a rather different strategy of fitting from the other models, it was not considered in detail (but see Section 5.3).

A widely used rank-2 model in plant breeding is the row regression model (Yates and Cochran, 1938; Mandel, 1961; Finlay and Wilkinson, 1963). As $\mathbf{1}_q$ is in the row space of $\mathbf{Y}$ (see Table 1), the column points will be collinear. For the row points, no special relation holds; they are expected to be scattered all over the plot. The counterpart of the row regression model is the column regression model, which follows from interchanging rows and columns.

2.2. *Rank-3 models*

Models with additive main effects supplemented by one or more biadditive multiplicative interaction terms constitute a major means for the analysis of two-way genotype by environment tables. Papers introducing the model are due to Gilbert (1963), Gollob (1968), and Mandel (1969, 1971). Early plant breeding applications can be found in Perkins (1972), and Freeman and Dowker (1973). Later applications include Gauch (1988), Snijders and van Eeuwijk (1991), and van Eeuwijk and Elgersma (1993).

Gower (1990) showed how a decomposition of a table $\mathbf{Y}$ (corrected for the general mean) constructed from an additive model with one multiplicative interaction term, say an additivity-plus-one model (Table 1), always leads to a $\mathbf{U}$ of the form $(\mathbf{a}, \mathbf{1}_p, \mathbf{c})$, and a $\mathbf{V}$ of $(\mathbf{1}_q, \mathbf{b}, \mathbf{d})$. Row points are located on a row plane, column points on a column plane, and both planes are orthogonal. The main effects parameters in $\mathbf{a}$ and $\mathbf{b}$ can be approximated by projection on suitable directions in the row and column plane, respectively. Similarly, the interaction parameters in $\mathbf{c}$ and $\mathbf{d}$ can be found by projection on to the intersection of the row and column plane.

The concurrence-plus-one model (Table 1) is not a regularly encountered model. Considering the 'classical' parametrization for the concurrence-plus-one model, $\mathcal{E}(\mathbf{Y}_{ij}) = \mu\mathbf{1}_p\mathbf{1}_q^T + \mathbf{a}\mathbf{1}_q^T + \mathbf{1}_p\mathbf{b}^T + \nu\mathbf{a}\mathbf{b}^T + \mathbf{c}\mathbf{d}^T$, the model can be diagnosed from coplanar row and column points with the planes intersecting non-orthogonally. For the constraints free parametrization, $\mathcal{E}(\mathbf{Y}_{ij}) = \lambda\ \mathbf{1}_p\ \mathbf{1}_q^T + \mathbf{r}\ \mathbf{s}^T + \iota\ \mathbf{v}^T$, a factorization $\mathbf{U} = (\lambda_1^*\mathbf{1}_p, \mathbf{r}, \mathbf{u})$, and $\mathbf{V} = (\lambda_2^*\mathbf{1}_q, \mathbf{s}, \mathbf{v})$ with $\lambda_1^*\lambda_2^* = \lambda$, would result in either parallel plane ($\lambda_1^* \neq \lambda_2^*$) or coinciding planes ($\lambda_1^* = \lambda_2^*$) being the indicator for this model having generated the data. Again, because of the deviating fitting procedure the constraints-free form was not studied (see Section 5.3).

As with the concurrence-plus-one model, the row-regression-plus-one model is rarely applied in practice (Table 1), although its utility is probably greater (van Eeuwijk and Elgersma, 1993). The row-regression-plus-one model is characterized by coplanarity of the column points.

3. Simulation and graphical display

3.1. *Representation*

Data tables with 10 rows (genotypes) and 8 columns (environments) were simulated from the models listed in Table 1. Data tables from additive models were constructed using the model $\mathbf{Y}_{ij} = \mu + a_i + b_j + \epsilon_{ij}$, where ϵ_{ij} represents a random error drawn from a normal distribution. Total variation (sum of squares) around the general mean was set to 1000. Additive row and column effects were chosen equidistant and scaled in such a way that the corresponding sum of squares equalled 'typical' plant breeding values. Errors were drawn from a normal distribution and then scaled to let the error sum of squares (noise) be equal to {1000 – sum of squares due to main effects}. Because errors were not constrained to add to zero over rows and columns, the actual variation of the simulated tables deviated slightly from 1000. (Imposing sum-to-zero constraints did not influence conclusions.) For data tables from models with interaction, an analogous procedure was followed as for additive models. Total error variation then was {1000 – variation due to main effects and interaction}.

To mimic real-life data the general mean was chosen to be 10, so that coefficients of variation (standard deviation of noise divided by general mean) were obtained between 0 and 22%. Tables usually should be corrected for the general mean before (diagnostically) biplotting to prevent deceptive multiplicativity for tables with low deviations from the general mean relative to that mean (Bradu and

Gabriel, 1978). In our case we could have chosen to take the general mean equal to zero, making this correction redundant. However, we preferred to include a general mean to be able to investigate the effect of not centering the table on the diagnosis of concurrence (-plus-one) models with unconstrained parametrizations.

After creation, data tables were factorized, and diagnostic biplots in two and three dimensions were made. Programmes were written in Genstat (1993). Two-dimensional diagnostic biplots do not require special facilities; simply plotting row and column coordinates after factorization, on whatever kind of device, suffices. For the three-dimensional plots a special Genstat procedure was written that allowed the inspection of the three-dimensional configuration of the points. (The Genstat code is available upon request from the authors.) Gower (1990) was followed for calculating best-fitting row and column planes, as well as projection directions for finding approximations to the additive main effects and the multiplicative interaction effects. The input for the graphical part of the procedure consisted of: the starting configuration of the row points and the column points, as derived from a factorization of the table (see below); the best-fitting row and column plane; the projection directions for the main effects in the row and the column plane; and the intersection line between the planes for the reconstitution of the interaction effects. Features were available for rotating the total configuration in arbitrary directions, removing points, planes, and projection directions, drawing projection lines from the points to relevant vectors in order to find approximations to main effects or interaction effects, and drawing residuals of the points to the best fitting plane. For the graphical inspections and manipulations a CA-DISSPLA (1988) module was also used, which was considerably faster than the Genstat procedure.

3.2. *Factorization and scaling*

Three types of factorization/scaling were used for the row and column coordinates. The first type was advocated by Bradu and Gabriel (1978). The table $\mathbf{Y}$ is decomposed via a singular value decomposition in $\mathbf{N}\,\mathbf{D}\,\mathbf{M}^{T}$, with $\mathbf{N}$ and $\mathbf{M}$ orthogonal matrices containing respectively the left and right singular vectors, and $\mathbf{D}$ a diagonal matrix holding the singular values. Row and column coordinates for the points of the diagnostic biplot are simply $\mathbf{U} = \mathbf{N}_s\mathbf{D}_s^{1/2}$ for the rows, and $\mathbf{V} = \mathbf{M}_s\mathbf{D}_s^{1/2}$ for the columns. The subscript s indicates the chosen rank of approximation, $s = 2$ for two-dimensional plots, and $s = 3$ for three-dimensional plots. The largest s values of $\mathbf{D}$ plus corresponding columns of $\mathbf{N}$ and $\mathbf{M}$ are retained. The quality of the representation depends on the amount of variation that is captured by the first s dimensions.

The second type of factorization/scaling is due to Gower (1990), and is exclusively for three-dimensional biplots. It also starts with a singular value decomposition of $\mathbf{Y}$, but includes a scaling which forces the solution into the direction of the additivity-plus-one-model by using a 'fitted values' variant of $\mathbf{U}$ and $\mathbf{V}$. $\mathbf{U} = \mathbf{N}_3\mathbf{D}_3^{1/2}$ and $\mathbf{V} = \mathbf{M}_3\mathbf{D}_3^{1/2}$ are, loosely speaking, regressed on the coordinates that $\mathbf{U}$ and $\mathbf{V}$ would contain were the data generated by the additivity-plus-one model that fits the data best.

The third type of factorization/scaling is a hybrid of the first two. As regards calculation it is a small variation on the first factorization. However, just as the second type of factorization it is especially suitable for the diagnosis of the additivity-plus-one model. The table is again decomposed by an ordinary singular value decomposition. A left singular vector representing the row main effects is then chosen on basis of the absolute correlation of its corresponding right singular vector with the unit vector, 1_p. By a symmetrical argument a right singular vector representing the column main effects can be found. For these pairs of left and right singular vectors the unit vectors are left unchanged, while the main effects vectors are scaled to a length equal to the singular vector. The first pair of singular vectors not representing main effects, usually the third pair, is given equal length. Automatic identification of singular vectors may proceed by imposing appropriate restrictions within an alternating row and column regression for the decomposition of $\mathbf{Y}$ (Gabriel and Zamir, 1979; van Eeuwijk, 1995). For example, to fix unit vectors, one right and one left singular vector, corresponding to different singular values, could be regressed on a unit vector within each iteration of the alternating least squares algorithm.

4. Results

A large number of simulated data sets was inspected by means of two- and three-dimensional diagnostic biplots to see whether the underlying generating model could be assessed correctly. A (representative) sample of all these simulations was selected for this paper to illustrate the general conclusions. The corresponding diagnostic biplots are given in Figs 1–6. Row and column main effects were taken equidistant. Rows are denoted by upper-case letters, with the row main effect increasing from A to J. Columns are denoted by lower-case letters, with the column main effect increasing from a to h. For every one of the figures displaying diagnostic biplots, corresponding analysis of variance tables are given in Table 2. Degrees of freedom in Table 2 were attributed as if the analyses were done on two-way tables of means. Of the concurrence-(plus-one) models only the forms with constraints were considered.

4.1. *Additivity*

A series of data sets was generated based on additivity with

Table 2. *Analysis of variance tables for the simulated data sets biplotted in Figs 1–6*

Additivity

Figure		1a		1b		1c		1d	
Source	d.f.	SS	MS	SS	MS	SS	MS	SS	MS
Rows	9	300	33.3	270	30.0	240	26.7	210	23.3
Columns	7	700	100.0	630	90.0	560	80.0	490	70.0
Rows. Columns	–	–	–	–	–	–	–	–	–
Noise	63	–	–	100	1.6	200	3.2	300	4.8

Figure		2a		2b, 2c, 2d	
Source	d.f.	SS	MS	SS	MS
Rows	9	350	38.9	300	33.3
Columns	7	650	92.9	600	85.7
Rows. Columns	–	–	–	–	–
Noise	63	–	–	100	1.6

Concurrence

Figure		3a		3b	
Source	d.f.	SS	MS	SS	MS
Rows	9	210	23.3	210	23.3
Columns	7	490	70.0	490	70.0
Concurrence	1	300	300.0	75	75.0
Noise	62	–	–	225	3.6

Row regression

Figure		4a		4b		4c	
Source	d.f.	SS	MS	SS	MS	SS	MS
Rows	9	195	21.7	195	21.7	195	21.7
Columns	7	455	65.0	455	65.0	455	65.0
Slopes	9	350	38.9	262.5	29.2	175	19.4
Noise	54	–	–	87.5	1.6	175	3.2

Additivity-plus-one

Figure		5a		5b	
Source	d.f.	SS	MS	SS	MS
Rows	9	195	21.7	195	21.7
Columns	7	455	65.0	455	65.0
Bil. Int.	15	350	38.9	175	11.7
Noise	48	–	–	175	3.6

Row regression-plus-one

Figure		6a		6b	
Source	d.f.	SS	MS	SS	MS
Rows	9	200	22.2	200	22.2
Columns	7	400	57.1	400	57.1
Slopes	9	250	27.8	200	22.2
Bil. Int.	14	150	10.7	100	7.1
Noise	40	–	–	100	2.5

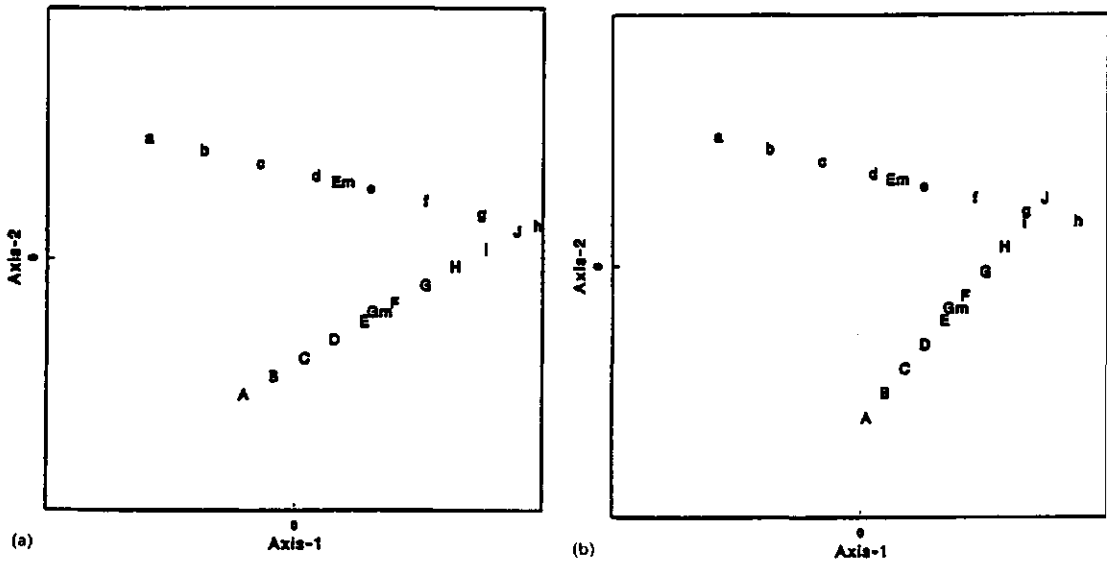


Fig. 3. *Two-dimensional plots showing concurrence with* (a) 30% *concurrence and no noise* (b) *with* 7.5% *concurrence and* 22.5% *noise. Gm denotes row-points mean, Em for column-points mean*

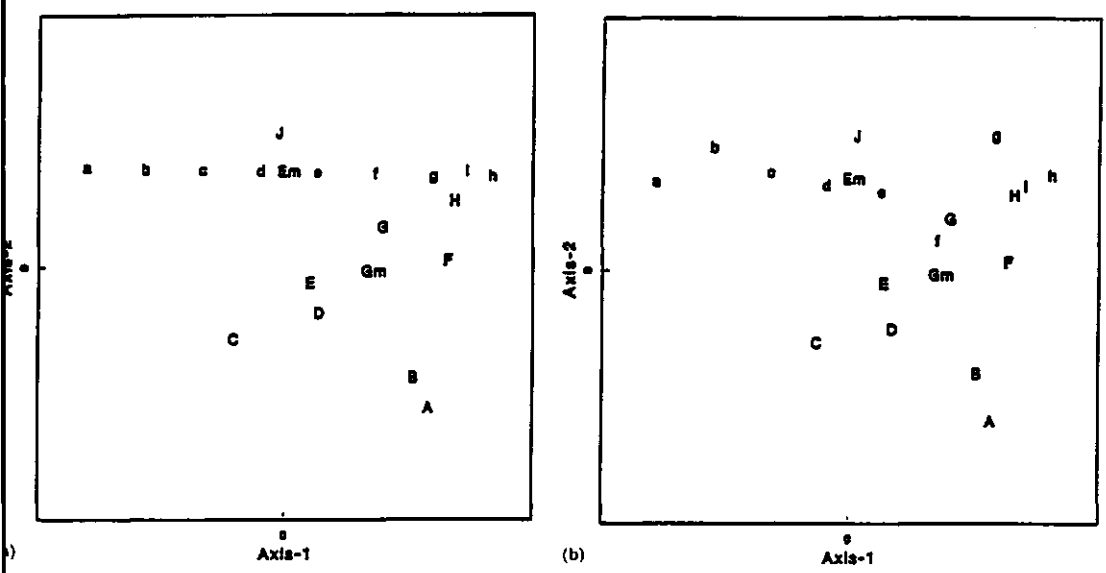
column plane.) Row and column points are driven away from the additivity lines, within the planes, as well as outside the planes, designated by the dotted residual lines. Also the row and column planes no longer intersect orthogonally. From this biplot a number of mistaken diagnoses are easily imaginable. The only feature of the biplot that survives in this scaling is that projection of the row and column points on the appropriate vectors still delivers a correct approximation to the additive effects. Figure 2c is constructed using the Gower type of scaling. More of the additive model survives now. Row and column plane orthogonality is restored and residuals from the row and column plane are small. The only problem that remains is that projections on the intersection of row and column plane clearly are not null, and thus suggest the presence of multiplicative interaction. The diagnosis would be an additivity-plus one model. In Fig. 2d the third type of scaling is used. A satisfactory result emerges; orthogonality of the planes, no residuals from the planes, and reasonable collinearity.

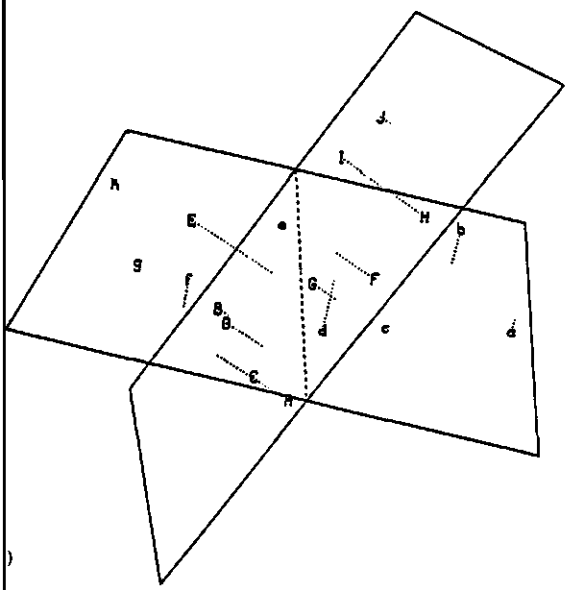
4.2. *Concurrence*

The series of concurrence models with increasing noise showed a remarkable phenomenon. Noise was increased from 0, via 7.5, 15, 22.5 to 30% of the total sum of squares. At the same time variation due to concurrence decreased from 30, via 22.5, 15, 7.5 to 0%. Row main effects were kept at 21%, column main effects at 49%. In Figs 3a and 3b the two-dimensional biplots are given for perfect concurrence, 30%, and consequently no noise, and for a case with 22.5% noise and thus 7.5 concurrence (see Table 2 for analyses of variance). It can be seen that row as well as column points stay perfectly collinear, but that their angle becomes a right angle. At orthogonality, i.e. additivity, the pattern of Fig. 1d reappears. Three-dimensional biplots of concurrence models with noise (not shown) displayed perfect coplanarity for row and column points under a non-orthogonal angle, irrespective of the type of scaling. Plotted in three dimensions, concurrence models with added noise would be diagnosed as concurrence-plus-one models.

4.3. *Row regression*

Typical results of the diagnosis of row regression plus noise data are given in Figs 4a and 4b (analyses of variance in Table 2). Figure 4a represents the two-dimensional biplot of an exact row regression model with 19.5% variation for the row main effect, 45.5% for the column main effect, and 35% for the interaction due to the heterogeneity in slopes. The column points are collinear, the row points scattered over the plane. In Fig. 4b the row regression interaction amounted to 26.25% and noise to 8.75. Even this small amount of noise causes a rather drastic disturbance of the diagnostic pattern. Three-dimensional plots of row regression models plus noise could only be plotted with





g. 4. *Plots for row regression.* (a) *Two-dimensional plot for exact row regression data with* 35% *interaction.* (b) *Two-dimensional plot with* .25% *heterogeneity and* 8.75% *noise* (c). *Three-dimensional plot with* 17.5% *heterogeneity and* 17.5% *noise. In* (4a) *and* (4b) *Gm is row-ints mean and Em column-points mean*

radu–Gabriel scaling. For the Gower scaling and the ird type of scaling, row points generally clustered indis-nguishably together near the origin. However, also the radu–Gabriel scaling created difficulties with regard to e diagnosis of models, as can be seen in Fig. 4c. The data come from a model with 19.5% row main effect, 45.5% column main effect, 17.5% heterogeneity of regression slopes, and 17.5% noise. Planes are non-orthogonal and residuals from both planes are large, making it difficult to decide on a specific model.

5.2. *Explicit equations*

Use of explicit equations in order to draw best-fitting row and column planes proved to be of vital importance. Assessment of coplanarity of points in space without the visual aid of best fitting planes can hardly be done. Projection directions for reconstituting main effects and interaction effects proved to be valuable and rather robust.

5.3. *Scaling and centring*

Use of different scaling methods showed that plots can be improved considerably if data are scaled to the model where they come from. However, the weakness of this procedure, of course, is that one first has to have a reasonable idea of what model to expect, before being able to choose an appropriate scaling. Inappropriate scaling soon leads to a mess. Nevertheless, the simultaneous use of different scaling methods may provide extra diagnostic keys, which partly alleviates the circularity of the procedure.

Concurrence (-plus-one) models can be distinguished from the other models in that they can be formulated in a constraints-free form (Denis and Gower, 1992, 1994). The constraints-free forms invite factorizations leading to diagnostic keys based on parallel and coinciding lines (planes). However, the proposed factorization cannot be obtained from the standard application of a singular value decomposition to the centred two-way table. As long as the intercept parameter, λ, is not equal to the general mean, μ, this standard procedure will always lead to a configuration of non-orthogonal lines (planes). When $\lambda = \mu$ the rank of the table will effectively be reduced by 1, with predictable consequences. Not centring leads to decompositions with heavily predominating first singular values, without clear patterns in the biplot. To obtain a factorization resulting in parallel lines (planes) an appropriate scaling matrix should be introduced. As concurrence (-plus-one) models are already quickly recognizable, this would seem to complicate matters unnecessarily.

5.4. *Combination of strategies*

A general problem is created by the choise between a rank-2 model and its associated rank-3 model. Discrimination between the pairs additivity – additivity-plus-one, concurrence – concurrence-plus-one, row regression – row regression-plus-one seems possible throughout. Within each pair problems are bigger. The suggestion of using the size of the singular (eigen)values (Bradu and Gabriel, 1978) proved to be rather useless in practice. A procedure in which various permutations of singular vectors are inspected, e.g. in three dimensions first, second and third; first, third and fourth; second, third and fourth, and so on, seems too laborious.

Problems also arise in discriminating models from slightly more general models. For example, the row-regression-plus-one model is not easily distinguishable from a general rank-3 model, or an additivity-plus-two model. Therefore, it seems almost inevitable to have recourse to explicit model fitting of series of models to supplement the diagnostic plots. The combination of these strategies might prove more viable than each of them on their own. The general class of models (additivity-plus-?, concurrence-plus-?, row regression-plus-?) can be assessed reasonably by means of diagnostic biplots. Explicit fitting within the selected class may help determining the dimensionality of the model. Then the diagnostic biplot may be inspected in greater detail to search for local additivity or other substructures of interest.

Also there is little reason not to utilize 'normal' biplots derived from singular value decompositions of residuals from additivity as diagnostic plots (Gabriel, 1971; Gabriel, 1978, Kempton, 1984).

Yet another tool is offered by the sizes of the correlations between main effects and singular vectors from the decomposition of the residual matrix from additivity. These may be supplemented by plots of row and column entries, y_{ij}, against row means, $y_{i.}$, or column means, $y_{.j}$ (Milliken and Johnson, 1989). Related is the use of plotting fitted row and column entries against estimates of biadditive parameters in constraints-free concurrence models or shifted multiplicative models (Cornelius *et al.*, 1992).

There seems to be insufficient ground for attributing to diagnostic biplots, in the classical sense, a special role in model selection for genotype by environment tables, but they certainly provide extra insight into the structure of the data, when used in combination with the other methods mentioned in this section.

Acknowledgements

The authors wish to thank John Gower, Aad van Eijnsbergen and Jean-Baptiste Denis for helpful discussions.

References

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

CA-DISSPLA (1988) Computer Associates International, Inc.

Cornelius, P. L., Seyedsadr, M., and Crossa, J. (1992) Using the shifted multiplicative model to search for 'separability' in crop cultivar trials. *Theoretical and Applied Genetics*, **84**, 161–172.

Denis, J.-B. (1991) Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Revue de Statistique Appliqueé*, **XXXIX**, 5–24.

Denis, J.-B. and Gower, J. C. (1992) Biadditive models. Technical Report. Laboratoire de Biométrie, INRA-Versailles, F-78026, France.

Denis, J.-B. and Gower, J. C. (1994) Biadditive models. Letter to the editor. *Biometrics*, **50**, 310–311.

Finlay, K. W. and Wilkinson, G. N. (1963) The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research*, **14**, 742–754.

Freeman, G. H. and Dowker, B. D. (1973) The analysis of variation between and within genotypes and environments. *Heredity*, **30**, 97–109.

Gabriel, K. R. (1971) Biplot display of multivariate matrices with application to principal component analysis. *Biometrika*, **58**, 453–467.

Gabriel, K. R. (1978) Least squares approximation of matrices by additive and multiplicative models. *Journal of the Royal Statistical Society, Series B*, **40**, 186–196.

Gabriel, K. R. and Odoroff, C. L. (1986) Algebraic aspects of scalar product and biplot analysis. In *Multidimensional data analysis*, ed. J. De Leeuw, W. Heiser and J. Meulman, pp. 91–111. DSWO Press, Leiden.

Gabriel, K. R. and Zamir, S. (1979) Lower rank approximations of matrices by least squares with any choice of weights. *Technometrics*, **21**, 489–498.

Gauch, H. G. Jr (1988) Model selection and validation for yield trials with interaction. *Biometrics*, **44**, 705–715.

Genstat 5 Committee (1993) *Genstat 5 reference manual*. Clarendon Press, Oxford.

Gilbert, N. (1963) Non-additive combining abilities. *Genetical Research*, **4**, 65–73.

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

Gower, J. C. (1990) Three-dimensional biplots. *Biometrika*, **77**, 773–785.

Kempton, R. A. (1984) The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science, Cambridge*, **103**, 123–135.

Mandel, J. (1961) Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association*, **56**, 878–888.

Mandel, J. (1969) The partitioning of interaction in analysis of variance. *Journal of Research of the National Bureau of Standards*, **73B**, 309–328.

Mandel, J. (1971) A new analysis of variance model for non-additive data. *Technometrics*, **13**, 1–18.

Milliken, G. A. and Johnson, D. E. (1989) *Analysis of messy data, Volume 2: nonreplicated experiments*. Van Nostrand Reinhold, New York.

Perkins, J. M. (1972) The principal components analysis of genotype environmental interactions and physical measures of the environment. *Heredity*, **29**, 51–70.

Seyedsadr, M. and Cornelius, P. L. (1992a) Computational procedures for estimation and hypothesis testing in analyzing two-way tables with interaction and no replication. In *Compstat 1992* ed. Y. Dodge and J. Whittaker, Physica Verlag, Heidelberg.

Seyedsadr, M. and Cornelius, P. L. (1992b) Shifted multiplicative models for nonadditive two-way tables. *Communications in Statistics B. Simulation and Computation*, **21**, 807–832.

Snijders, C. S. and van Eeuwijk, F. A. (1991) Genotype × strain interactions for resistance to Fusarium head blight caused by Fusarium culmorum in winter wheat. *Theoretical and Applied Genetics*, **81**, 239–244.

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

van Eeuwijk, F. A. (1995) Multiplicative interaction in generalized linear models. *Biometrics*, in press.

van Eeuwijk, F. A. and Elgersma, A. (1993) Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. *Heredity*, **70**, 447–457.

Yates, F. and Cochran, W. G. (1938) The analysis of groups of experiments. *Journal of Agricultural Science, Cambridge*, **28**, 556–580.

Euphytica 84: 1–7, 1995.

Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models

F.A. van Eeuwijk
DLO-Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 16, 6700 AA Wageningen, The Netherlands

Received 28 February 1994; accepted 2 November 1994

Key words: AMMI, best linear unbiassed prediction, factorial regression, genotype by environment interaction, multiplicative interaction, reduced rank regression, two-way table, variance components, variety trials

Summary

The multi-environment trial, in which a number of genotypes is evaluated over a range of environmental conditions, is a standard experiment in plant breeding in general, and variety testing in particular. Useful statistical models for the analysis of multi-environment trials, with emphasis on the analysis of genotype by environment interaction, can be found in the classes of linear and bilinear models. Statistical properties of the most important representatives of these model classes are shortly reviewed. Structural differences between the models stem from: (1) the inclusion of random model terms in addition to fixed model terms; (2) the representation of the interaction by additive or multiplicative parameters; (3) the incorporation of concomitant variables on the levels of the environmental factor. For models with bilinear multiplicative structure for the interaction it is described how the interaction can be visualized by biplots. An illustration of the application of the models and biplots is given in a companion paper.

Introduction

A classic experiment in plant breeding is the multi-environment experiment, which in the standard case involves the evaluation of a number of genotypes at a number of locations over a number of years. Inferences to be made from multi-environment trials concern genotypes and environments. For the genotypes, typically, predictions are wanted for performance over years, or over years and locations. For environments discriminatory power prevails. Multi-environment experiments form the core of varietal testing programmes in many countries. These programmes have to assess the agronomic value of new varieties. Eventually, decisions have to be made about admittance of new varieties to the Variety List. A characteristic feature of data collected within Variety List testing programmes is their unbalancedness. The variety sortment changes over the years and not all varieties are tested at all locations within each year. Prediction of performance is better not based on simple mean performance. Two popular methods providing adjusted means are fitting constants (Searle, 1971; Patterson, 1978) and best linear unbiassed prediction (Henderson, 1963; Robinson, 1991; Searle et al., 1992). Both these methods are based on linear models with, usually, only indicator variables as explanatory variables, i.e. classic analysis of variance, or ANOVA, models. Interaction is modelled by a separate, additive parameter for each combination of genotype by environment, coarsely and unparsimoniously. These models are used primarily for arriving at good predictions over a range of environments, thereby in some sense averaging (weighted) over the interaction present. No attempt is made at interpretation of the interaction, thus leaving the causes of interaction for what they are.

As an alternative to linear formulations of interaction, multiplicative formulations can be chosen that do permit interpretation of interaction, as differential genotypic sensitivity to environmental variable(s). Three main classes may be distinguished. The first, and at the moment most popular, class consists of the Additive Main effects and Multiplicative Interaction effects (AMMI) models (Gollob, 1968; Mandel, 1969;

etry reveals that the interaction between a genotype i and an environment j can be obtained from a projection of either vector onto the other. The reason is that the interaction according to an AMMI model with two product terms for interaction, $\gamma_{1i}^{*}\delta_{1j}^{*} + \gamma_{2i}^{*}\delta_{2j}^{*}$, is equal to the inner product between the vectors $(\gamma_{1i}^{*},\gamma_{2i}^{*})$ and $(\delta_{1j}^{*},\delta_{2j}^{*})$, or the projection of either vector onto the other, times the length of the vector on which projection takes place. In case of an obtuse angle between genotypic and environmental vector, an additional minus sign is necessary. It is easy to read from a biplot the relative interactions that genotypes exhibit in a particular environment. One only needs to look at the ranking of the projections of the genotypic vectors on the particular environmental vector. Cosines of the angles between genotypic vectors approximate correlations between genotypes with respect to their interactions. The same holds true for the environments.

For reduced rank factorial regression the story is slightly more complicated. We again assume that two multiplicative terms suffice and distribute the singular values over the scores. Interaction can then be described as

$$\gamma_{1i}^{*}\left(\sum_{h=1}^{H}\rho_{1h}x_{hj}\right)^{*} + \gamma_{2i}^{*}\left(\sum_{h=1}^{H}\rho_{2h}x_{hj}\right)^{*} = \gamma_{1i}^{*}\left(\sum_{h=1}^{H}\rho_{1h}^{*}x_{hj}\right) + \gamma_{2i}^{*}\left(\sum_{h=1}^{H}\rho_{2h}^{*}x_{hj}\right).$$

In the reduced rank regression biplot we plot three types of vectors whose coordinates are determined by: (1) the genotypic sensitivities, $(\gamma_{1i}^{*},\gamma_{2i}^{*})$; (2) the environmental characterisations, $\left(\sum_{h=1}^{H}\rho_{1h}^{*}x_{hj}\,,\,\sum_{h=1}^{H}\rho_{2h}^{*}x_{hj}\right)$; and (3) the coefficients for the environmental variables within the reduced rank factorial regression axes, $(\rho_{1h}^{*},\rho_{2h}^{*})$. As in the AMMI biplot the inner product of the genotypic vector i with the environmental vector j gives the interaction (non-additivity) for genotype i in environment j. In addition, inner products between the genotypic sensitivity vectors, $(\gamma_{1i}^{*},\gamma_{2i}^{*})$, and the coefficient vectors, $(\rho_{1h}^{*},\rho_{2h}^{*})$, approximate the (full rank) factorial regression coefficients, $\xi_{hi} = \rho_{1h}^{*}\gamma_{1i}^{*} + \rho_{2h}^{*}\gamma_{2i}^{*}$. For illustrations of the use of biplots in reduced rank regression models see Ter Braak (1990) and Ter Braak & Looman (1994).

Information on measured environmental variables can also be added to AMMI biplots, although these variables had no influence on the determination of the environmental axes. We can indicate directions of greatest change with respect to a particular environmental variable, by depicting the variable by the coefficients of its regression on the axes. When the scaling constant c is chosen equal to one, this is equivalent to using the correlations of the environmental variable with the axes. The sum of the squared correlations over the axes gives a measure for the quality of the representation. Reduced rank regression biplots can be supplemented with environmental information not used in the determination of the axes in the same way.

Epilogue

In this paper linear and bilinear models for the analysis of genotype by environment interaction have been described in a somewhat theoretical context. The best appreciation of what the models may add to the insights of the practical plant breeder is obtained from their application to real life data. In the sequel to this paper (van Eeuwijk et al., 1995), data on dry matter content from the official Dutch Maize Variety Trials will be analyzed and it will be shown how the joint application of the models can lead to an interpretation of genotype by environment interaction in terms of differential sensitivity to external environmental variables.

Acknowledgement

The author wishes to thank Jean-Baptiste Denis (INRA, Versailles) for comments and discussion.

References

Cornelius, P.L., 1993. Statistical tests and retention of terms in additive main effects and multiplicative interaction model for cultivar trials. Crop Sci. 33: 1186–1193.

Davies, P.T. & M.K.S. Tso, 1982. Procedures for reduced-rank regression. Appl. Stat. 31: 244–255.

Denis, J.B., 1980. Analyse de régression factorielle. Biom. Praxim. 20: 1–34.

Denis, J.B., 1988. Two way analysis using covariates. Statistics 19: 123–132.

Denis, J.B., 1991. Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. Rev. Stat. Appl. XXXIX: 5–24.

Denis, J.B. & J.C. Gower, 1992. Biadditive models. Technical report. Laboratoire de Biométrie, INRA-Versailles.

Denis, J.B. & J.C. Gower, 1994. Biadditive models. Letter to the editor. Biometrics 50: 310–311.

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

Gabriel, K.R., 1971. Biplot display of multivariate matrices with application to principal component analysis. Biometrika 50: 453–467.

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

Gabriel, K.R. & S. Zamir, 1979. Lower rank approximations of matrices by least squares with any choice of weights. Technometrics 21: 489–498.

Gauch, H.G., 1988. Model selection and validation for yield trials with interaction. Biometrics 44: 705–715.

Gauch, H.G., 1992. Statistical analysis of regional yield trials. Elsevier, Amsterdam.

Genstat 5 Committee, 1993. Genstat 5 release 3 reference manual. Clarendon Press, Oxford.

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

Henderson, C.R., 1963. Selection index and expected genetic advance. In: W.D. Hanson & H.F. Robinson (Eds). Statistical genetics and plant breeding. pp. 141–163. National Academy of Sciences, National Research Council, Washington.

Johnson, D.E. & F.A. Graybill, 1972. An analysis of a two-way model with interaction and no replication. J. Am. Stat. Ass. 67: 862–868.

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

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

Mandel, J., 1969. The partitioning of interaction in analysis of variance. J. Res. NBS B. 73B: 309–328.

Patterson, H.D., 1978. Routine least squares estimation of variety means in incomplete tables. J. Natn. Inst. Agric. Bot. 14: 401–412.

Patterson, H.D. & R. Thompson, 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58: 545–554.

Perkins, J.M., 1972. The principal components analysis of genotype environmental interactions and physical measures of the environment. Heredity 29: 51–70.

Robinson, D.L., 1987. Estimation and use of variance components. Statistician 36: 3–14.

Robinson, G.K., 1991. That BLUP is a good thing: the estimation of random effects. Stat. Sci. 6: 15–51.

Searle, S.R., 1971. Linear Models. Wiley, New York.

Searle, S.R., G. Casella & C.E. McCulloch, 1992. Variance components. Wiley, New York.

Snedecor, G.W. & W.G. Cochran, 1980. Statistical Methods, 7th edn. Iowa State University Press, Ames.

Ter Braak, C.J.F., 1990. Interpreting canonical correlation analysis through biplots of structure correlations and weights. Psychometrika 55: 519–531.

Ter Braak, C.J.F. & C.W.N. Looman, 1994. Biplots in reduced-rank regression. Biom. J. 36: 983–1003.

van Eeuwijk, F.A., 1992. Interpreting genotype-by-environment interaction using redundancy analysis. Theor. Appl. Genet. 85: 89–100.

van Eeuwijk, F.A. & A. Elgersma, 1993. Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. Heredity 70: 447–457.

van Eeuwijk, F.A., 1995. Multiplicative interaction in generalized linear models. Biometrics: in press.

van Eeuwijk, F.A., L.C.P. Keizer & J.J. Bakker, 1995. Linear and bilinear models for the analysis of multi-environment trials: II. An application to data from the Dutch Maize Variety Trials. Euphytica: this issue.

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

Chapter XI

Linear and bilinear models for the analysis of multi-environment trials: II. An application to data from the Dutch Maize Variety Trials

F.A. van Eeuwijk, L.C.P. Keizer & J.J. Bakker
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

Euphytica **84:** 9–22, 1995.

Linear and bilinear models for the analysis of multi-environment trials: II. An application to data from the Dutch Maize Variety Trials

F.A. van Eeuwijk, L.C.P. Keizer & J.J. Bakker
DLO-Center for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 16, 6700 AA Wageningen, The Netherlands

Received 28 February 1994; accepted 2 November 1994

Key words: AMMI, best linear unbiassed prediction, factorial regression, genotype by environment interaction, maize, missing values, multiplicative interaction, reduced rank regression, two-way table

Summary

As the sequel to a paper that dealt with the theoretical aspects of linear and bilinear models for the analysis of genotype by environment interaction in multi-environment trials, this paper presents an illustration of the application of these models to real life data. The data come from maize trials that were conducted within the ongoing evaluation programme for the Dutch Descriptive Variety List of Field Crops. The variable that is analyzed is dry matter content. It is shown how linear and bilinear models can be used supplementary to each other within a general strategy for dealing with genotype by environment interaction.

Introduction

In the companion paper to this paper (van Eeuwijk, 1995a) various models suitable for dealing with genotype by environment problems were reviewed. In this paper their application to a real life data set will be described. The data concern the variable dry matter content in maize, and they were made available from the evaluation programme for the Dutch Descriptive List of Field Crops. We will start with a description of the data and then continue with sections on the application of the various models. Finally, the application of models will be considered in a general perspective in an attempt to provide a strategy for analysing data from multi-environment trials containing genotype by environment interaction. Special attention is given to the way in which models with multiplicative formulations for interaction can be used in addition to analysis of variance models to facilitate the interpretation of genotype by environment interaction. The statistical analyses presented can all be run with standard statistical software, no specialized software is needed. In an Appendix a brief description is given of the most important statements needed in Genstat (1993).

Material

From the Dutch Maize Variety Trials for the Descriptive List of Field Crops a selection of data was made for the variable dry matter content (DMC). Individual trials consisted of incomplete block experiments with three replicates at a particular location in a particular year. The experimental plots had a width of six rows and were 14 meter long. The basic data that were at our disposal were the fitting constants (adjusted) means for the variety means of the individual trials, expressed as a percentage. Table 1 shows the selection that was made. Included were 18 varieties, evaluated from 1980 to 1990 at four locations in the Netherlands; Southern Sand (Ss), Central Sand (Cs), Northern Sand (Ns), and River Clay (Rc). Unbalancedness was caused by shifts in the variety sortment over the years, and the absence of the location Rc in 83 and 89. The data in Table 1 were first logit transformed to achieve better homogeneity of variance and normality, and subsequently multiplied by 100 for readability. For interpretation of Variety by Year interaction the following information characterizing the years was used;

Table 1. Selected dry matter content data (rounded percentages) from the Dutch Maize Variety Trials. For explanation of the codes see Table 5 and text

Variety Year	Soil	Br	Sp	Fr	Pr	L1	An	Ma	Vi	Do	Ir	Cl	Gr	Sn	As	L2	Sc	Pr	Sg
80	Ss	30	29	29	28	30	31	28	29	28	29	*	*	*	*	*	*	*	*
	Cs	27	29	26	27	27	28	27	26	26	27	*	*	*	*	*	*	*	*
	Ns	30	35	28	31	29	28	28	30	28	31	*	*	*	*	*	*	*	*
	Rc	30	32	30	30	30	31	29	29	30	29	*	*	*	*	*	*	*	*
81	Ss	30	32	29	28	30	31	28	31	23	30	*	*	*	*	*	*	*	*
	Cs	26	26	25	26	28	26	25	27	27	27	*	*	*	*	*	*	*	*
	Ns	27	30	27	28	28	29	28	29	27	29	*	*	*	*	*	*	*	*
	Rc	35	36	34	35	35	34	34	35	34	35	*	*	*	*	*	*	*	*
82	Ss	34	32	32	33	35	34	33	32	33	32	34	35	*	*	*	*	*	*
	Cs	32	33	33	30	32	32	31	33	32	32	37	36	*	*	*	*	*	*
	Ns	27	29	28	27	29	28	28	27	27	26	31	30	*	*	*	*	*	*
	Rc	35	34	34	36	37	35	34	34	34	36	36	37	*	*	*	*	*	*
83	Ss	33	28	31	28	32	32	28	30	31	31	34	32	34	*	*	*	*	*
	Cs	29	28	29	29	31	29	27	28	29	30	31	29	32	*	*	*	*	*
	Ns	32	30	32	29	32	33	27	30	30	29	35	30	36	*	*	*	*	*
	Rc	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
84	Ss	29	30	29	30	31	31	28	29	28	30	36	29	32	28	*	*	*	*
	Cs	22	21	21	20	21	23	19	21	20	21	23	21	21	21	*	*	*	*
	Ns	20	20	20	19	20	21	19	20	19	19	21	19	21	19	*	*	*	*
	Rc	28	28	27	28	30	32	26	29	28	30	32	27	32	27	*	*	*	*
85	Ss	34	38	*	*	37	35	34	34	33	35	36	36	38	38	45	36	35	41
	Cs	26	26	*	*	26	27	24	26	25	26	27	24	27	24	32	26	27	27
	Ns	20	20	*	*	22	22	20	21	21	21	23	19	21	19	25	21	21	21
	Rc	28	29	*	*	30	29	27	28	28	29	30	27	30	28	35	28	28	29
86	Ss	29	28	*	*	*	*	26	30	28	29	31	27	31	27	35	30	28	30
	Cs	27	27	*	*	*	*	25	29	26	28	29	26	32	27	35	32	27	31
	Ns	22	20	*	*	*	*	20	23	22	22	24	19	22	20	27	23	21	23
	Rc	27	27	*	*	*	*	26	28	26	28	29	26	30	26	38	31	27	29
87	Ss	27	26	*	*	*	*	23	27	25	28	27	25	32	27	34	27	27	28
	Cs	27	26	*	*	*	*	23	27	25	26	27	24	30	26	29	26	25	28
	Ns	20	20	*	*	*	*	20	22	20	21	22	19	22	19	25	21	21	21
	Rc	23	24	*	*	*	*	21	25	23	24	25	22	27	24	30	23	23	25
88	Ss	24	23	*	*	*	*	*	*	24	24	24	22	24	23	28	23	23	24
	Cs	25	24	*	*	*	*	*	*	23	25	25	22	24	24	30	24	24	24
	Ns	28	26	*	*	*	*	*	*	27	29	29	25	27	28	37	28	27	28
	Rc	31	30	*	*	*	*	*	*	29	30	31	30	33	33	38	32	27	32
89	Ss	32	31	*	*	*	*	*	*	*	*	36	31	28	31	36	32	30	34
	Cs	37	35	*	*	*	*	*	*	*	*	37	34	34	33	39	36	33	37
	Ns	29	27	*	*	*	*	*	*	*	*	30	23	26	25	32	28	27	28
	Rc	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
90	Ss	35	33	*	*	*	*	*	*	*	*	*	*	29	29	37	29	30	33
	Cs	32	33	*	*	*	*	*	*	*	*	*	*	32	31	33	30	30	30
	Ns	30	30	*	*	*	*	*	*	*	*	*	*	30	28	33	26	28	31
	Rc	33	35	*	*	*	*	*	*	*	*	*	*	33	34	36	33	29	34

Table 2. Values of concomitant variables on levels of the factor Year, and contributions to interaction sum of squars for Variety by Year table of best linear unbiased predictions; year, number of days in May with temperature below 10°C, mean temperature over the period May-June (°C), mean temperature over the period July-August (°C), radiation during growing season (10^4 Joule/cm^2), precipitation in July (mm), mean dry matter content (%), absolute contribution to sum of squares for interaction, relative contribution to sum of squares interaction (%)

Year	L10-May	MTMJ	MTJA	Rad	Prec	DMC (Year)	Contr. SS Int.	Perc. SS Int.
80	6	10.8	14.7	10.7	146.7	29.6	63.00	5.6
81	6	13.3	18.1	11.6	109.4	30.2	69.82	6.2
82	9	16.6	11.4	12.9	24.4	32.8	142.26	12.6
83	12	10.5	13.5	12.5	20.3	31.3	137.30	12.2
84	11	13.6	19.1	10.7	69.1	25.0	41.62	3.7
85	7	15.4	11.0	11.3	83.9	27.7	38.59	3.4
86	1	10.5	12.2	12.3	42.3	26.8	146.44	13.0
87	17	13.2	17.0	11.0	164.4	24.5	108.40	9.6
88	4	15.7	11.0	10.6	140.4	26.5	75.12	6.7
89	3	10.9	13.4	12.4	34.5	31.4	128.95	11.5
90	1	13.7	16.6	12.0	43.3	30.7	173.69	15.4

- Number of days with mean temperature below 10 °C in May (L10-May)
- Mean temperature (°C) over the period May-June (MTMJ)
- Mean temperature (°C) over the period July-August (MTJA)
- Total Radiation (10^4 Joule/cm^2) during the growing season (May-August) corrected for light interception (Rad)
- Precipitation (mm) in July (Prec)

In addition to these clearly external variables we used the best linear unbiassed predictions for the Year main effect of the variable DMC itself, DMC(Year). The variables were standardized to give them equal variation. For reference the raw values of the variables are given in Table 2. DMC(Year) is given as a percentage, but the analysis was done on logit scale.

Methods

With one exception the models used for analysis are all described in van Eeuwijk (1995a), and seemingly obscure remarks in this section will hopefully become clear from reading that paper. We list the models here together with the main purpose they served for the analysis of our DMC data.

Linear model with all terms fixed/Fitting constants

A fixed analysis of variance model was fitted by the method of fitting constants. Terms in fitting order were: Year, Location, Year by Location, Variety, Variety by Year, Variety by Location. This analysis served to obtain a rough idea of the distribution of variation over the various terms. To that purpose the analysis of variance table was inspected.

Mixed model/REML

The following mixed model was fitted using REML: Fixed terms = Variety, Location, Variety by Location; Random terms = Year, Variety by Year, Year by Location. Variance components were inspected. Variety by Year interaction was chosen for further investigation. A table of best linear unbiassed predictions was made on basis of the fitted mixed model. Because Variety by Year interaction was taken random, no problems occurred with respect to predictions for cells which were empty in the Variety by Year table of (observed) means. The complete Variety by Year table of predictions contained the basic data set for the analyses with AMMI models, and full and reduced rank factorial regression models. The contributions of individual varieties and years to the interaction sum of squares in the Variety by Year table were calculated as descrip-

Table 5. Variety name, variety code, flowering date (1 = very early, 2 = early, 3 = mid early, 4 = mid late), absolute and relative contribution (%) to interaction sum of squares in the Variety by Year table of best linear unbiassed predictions, sensitivity to selected environmental variables in the factorial regression model (- or + = light, - - or ++ = moderate, - - - or +++ = heavy)

Variety	Code	Flo-wer-ing	Contr. SS int.	Perc. % SS int.	Sensitivity to			
					DMC-(year)	Rad	L10-May	MTIA
Brutus	Br	2	92.95	8.3			-	
Splenda	Sp	3	184.00	16.4	+++	- - -	-	++
Fronica	Fr	3	13.79	1.2				
Protasil	Pt	3	12.94	1.2				
LG 11	L1	2	11.31	1.0				
Anko	An	2	31.48	2.8	-			
Markant	Ma	4	66.04	5.9	++	-		
Vivia	Vi	2	65.03	5.8	- -	+		
Dorina	Do	3	35.90	3.2				-
Iria	Ir	2	26.90	2.4	-			
Clipper	Cl	2	55.72	5.0				+
Gracia	Gr	4	128.00	11.4	++		+	
Sonia	Sn	2	181.73	16.2	-	+	++	
Ascot	As	4	38.11	3.4	+	-		
LG 2080	L2	1	76.77	6.8	-			-
Scana	Sc	2	71.59	6.4	-	+		
Presta	Pr	2	17.46	1.6				
Sogetta	Sg	3	15.46	1.4				

Table 6. Analysis of variance table for AMMI analysis of Variety by Year table of best linear unbiassed predictions

Source	Degrees of freedom	Sum of squares	Mean square	Variance ratio
Variety·Year	170	1125.91	6.62	
AMMI axis 1	26	405.90	15.61	3.73
AMMI axis 2	24	217.15	9.05	2.16
Error	120	502.86	4.19	

A visual representation of the results is given in Fig. 1 (scaling: c = 1, see van Eeuwijk, 1995a, section on biplots). In accordance with expectation the varieties with the highest non-additivity (Table 5), Splenda, Gracia, and Sonia were the furthest away from the origin, which represents a hypothetical additively behaving variety. Years with high non-additivity, were also located far from the origin; 82, 83, 86, 87, 89, and 90 (Table 2). Approximations of the amount of interaction for a particular variety in a particular year can be obtained by projecting the year points (squares) on the variety lines. Splenda and Brutus had positive interaction in 89 and 90 (projections of the years on the same side of the origin as the lines for the varieties), and negative in 83. Sonia behaved opposite to Splenda and Brutus. Gracia and Markant had positive interactions in 82 and negative ones for 86 and 87. Vivia was the mirror image of Gracia and Markant.

Enrichment of the AMMI biplot by including the directions of greatest change for environmental variables made only sense for DMC(Year), as this was the only variable that could be represented sufficiently well (59%, obtained from the regression of DMC(Year) on the environmental values of the two axes). Projections

Table 7. Analysis of variance table for factorial regression of the interaction in the Variety by Year table of best linear unbiassed predictions. Interactions of varieties with individual environmental variables are presented in the order of inclusion and corrected for the other three variables in the model

Source	Degrees of freedom	Sum of squares	Mean square	Variance ratio
Variety·Year	170	1125.91	6.62	
Variety·(DMC(Year) + Rad + L10-May + MTJA)	68	583.07	8.57	1.61
Variety·DMC(Year)	17	209.95	12.35	2.32
Variety·Rad	17	119.13	7.01	1.32
Variety·L10-May	17	146.72	8.60	1.62
Variety·MTJA	17	113.38	6.67	1.25
Error	102	542.84	5.32	

of the year points on this direction approximate relative DMC(Year) values for the years. Roughly, years below the line y = - x had higher than average DMC(Year), with 82 being the most extreme. Above the line y = - x years with less than average DMC(Year) were found, with 86 and 87 being extreme.

Simultaneously, extra information on the varieties may be considered. Flowering date was decided to be the most interesting extra information (Table 5). Varieties were projected on the line determined by the direction of DMC(Year) (projections not shown). These projections ordered the varieties from very early (LG 20 80) and early flowering varieties on the upper right (Scana, Sonia, Vivia) to mid early (Splenda) and mid late flowering varieties on the lower left (Gracia, Markant). Years with high DMC(Year) are in general years with climatological circumstances that are deemed beneficial to the growth of maize, i.e. many sun hours and high temperatures. Later varieties seemed to profit relatively more from such weather than earlier varieties. Later flowering varieties had high positive interactions in good years, e.g. Gracia in 82 (the projection of the point for 82 on the line for Gracia is long and on the same side of the origin), whereas earlier varieties then had high negative interaction, e.g. Vivia in 82. Later varieties can compensate for their slower development in good years, and under these circumstances do comparably well, because in these years the gain of being early is not very important. Because all varieties were harvested at the same time within a year, it may even be expected that in good years the earlier varieties were harvested after having reached their optimum, and so were declining already.

In bad years, i.e. relatively little irradiation (few sun hours) and/or much rain, earlier varieties had a relative advantage, e.g. Vivia in 87, whereas later varieties had a relative disadvantage, e.g. Gracia in 87. The later varieties cannot get to their optimum in bad years, whereas the earlier varieties manage to get just enough. The tendency in West-European maize breeding is towards earlier varieties. These perform relatively better in less favourable circumstances. This can be interpreted as a form of adaptibility that is desirable in a climate like that of the Netherlands.

Factorial regression

After a general interpretation of the Variety by Year interaction following from the AMMI analysis, a more detailed interpretation of the interaction for individual varieties in relation to specific environmental variables was obtained by means of factorial regression. Four environmental variables were selected by the stepwise variable selection procedure described above. The variables are given in order of inclusion in Table 7. Each individual variable is presented with its fit corrected for the other three variables. Also, the joint fit is given. The four variables together accounted for 51.8% of the Variety by Year interaction.

Table 8. Regression coefficients for individual varieties in full (FR) and reduced rank factorial regression (RRR)

Variety	DMC(Year)		Rad		L10-May		MTJA	
	FR	RRR	FR	RRR	FR	RRR	FR	RRR
Br	0.92	1.22	- 0.43	- 0.97	- 1.03	- 0.71	0.64	0.59
Sp	3.32	3.08	- 3.03	- 2.70	- 1.79	- 2.21	2.04	1.74
Fr	- 0.20	0.02	0.86	0.45	0.54	0.76	- 0.38	- 0.47
Pt	0.42	0.18	- 0.89	- 0.38	- 0.36	- 0.49	0.00	- 0.19
Ll	- 0.69	- 0.53	0.80	0.50	0.27	0.43	- 0.29	- 0.33
An	- 1.14	- 1.29	0.12	0.27	- 0.16	- 0.49	0.50	0.14
Ma	2.28	2.13	- 1.67	- 1.24	- 0.55	- 0.48	0.03	0.56
Vi	- 2.61	- 2.26	1.41	0.57	- 0.79	- 0.71	0.90	0.16
Do	- 0.94	- 1.04	0.41	0.82	0.36	0.59	- 1.26	- 0.49
Ir	- 1.14	- 1.19	- 0.08	0.07	- 0.77	- 0.76	0.15	0.32
Cl	0.63	1.06	0.43	- 0.62	- 0.31	- 0.24	1.24	0.28
Gr	2.76	2.87	0.04	- 0.15	1.73	1.86	- 0.80	- 0.79
Sn	- 1.72	- 2.06	1.67	1.93	2.57	1.68	- 0.19	- 1.30
As	1.37	0.79	- 1.63	- 0.44	0.29	- 0.14	- 0.37	0.19
L2	- 1.62	- 1.99	- 0.03	1.04	0.08	0.26	- 1.77	- 0.41
Sc	- 1.81	- 1.13	1.75	0.62	- 0.68	0.19	- 0.46	- 0.26
Pr	0.20	- 0.25	- 0.47	0.38	0.91	0.44	- 0.44	- 0.30
Sg	- 0.04	0.39	0.74	- 0.16	- 0.31	0.02	0.44	0.04

Table 9. Analysis of variance for reduced rank factorial regression (RRR) of Variety by Year table of best linear unbiassed predictions, difference between full (FR) and reduced rank factorial regression, and difference between AMMI and reduced rank factorial regression

Source	Degrees of freedom	Sum ofsquares	Mean square	Variance ratio
Variety·Year	170	1125.9	6.62	
RRR axis 1	20	299.9	15.00	2.97
RRR axis 2	18	158.7	8.82	1.75
RRR error	132	666.6	5.05	
FR-RRR	30	123.8	4.13	
FR Error	102	542.8	5.32	
AMMI-RRR	12	163.7	13.64	
AMMI Error	120	502.9	4.19	

Table 8 shows the regression coefficients for the individual varieties. Table 5 contains a schematic representation of the main features of Table 8. Positive coefficients between 1.00 and 2.00 in Table 8 are depicted by a + in Table 5, coefficients between 2.00 and 3.00 by ++, and above 3.00 by +++. Negative coefficients were analogously represented using minus signs. The varieties with the largest non-additivity, i.e. highest ecovalence (Table 5), had in general the highest sensitivities to the various explanatory variables for interaction, and especially to the variables that were included first in the stepwise factorial regression; DMC(Year) and Rad. Also there was a relation between sensitivity to DMC(Year) and Rad on the one hand and flowering date on the other hand. Splenda, Markant, Gracia, and Ascott, all rather late varieties were positively sensitive to DMC(Year). Anko, Vivia, Irla, Sonia, Lg 20 80, and Scana, all rather early, were negatively sensitive to DMC(Year). High DMC(Year) values generally indicate conditions with enough sun hours and high

temperatures, without drought stress. Later varieties utilized these conditions relatively better than earlier varieties. In contrast, later varieties suffered relatively more from the less beneficial circumstances which lead to low DMC(Year) values. With respect to Rad, the later varieties Splenda, Markant, and Ascot were negatively sensitive, whereas the earlier varieties Vivia, Sonia, and Scana were positively sensitive. Taking into account the conditional nature of the regression coefficients, this means that for given values of DMC(Year), higher Rad was detrimental to later varieties, whereas lower Rad worked beneficially. The cause for this observation may be a foto-inhibition effect. No relationship seemed to exist between sensitivity to the variables L10-May or MTJA, and flowering date.

Reduced rank factorial regression

A reduced rank regression model was fitted to the interaction table using the explanatory variables from the eventual factorial regression model. To determine the number of (significant) axes necessary for adequate description, the same method was used as for AMMI; testing successive axes against a residual derived from the sums of squares and degrees of freedom not yet in the model. Table 9 shows that two axes were retained. The reduced rank regression model with two axes gave an adequate approximation to the corresponding (full rank) factorial regression model with 30 degrees of freedom more. This was concluded from the size of the mean square for the difference between factorial regression and reduced rank regression (4.13), which was comparable to the error for the full rank factorial regression model (5.32). On the other hand, a difference with respect to AMMI was unequivocally present. The AMMI model with two axes, though only containing 12 degrees of freedom more than the reduced rank regression model, clearly described interaction which was not covered by the latter.

The coefficients for the explanatory variables in the reduced rank regression axes are given in Table 10. The first axis was a contrast between DMC(Year) and MTJA on the one hand and Rad and L10-May on the other hand. The second axis is a sum of DMC(Year) and L10-May. The correlations between the explanatory variables and the reduced rank regression axes are all greater than the corresponding correlations between the variables and the AMMI axes. This is a consequence of the restriction imposed on the axes in reduced rank regression of having to be linear combinations of explanatory variables.

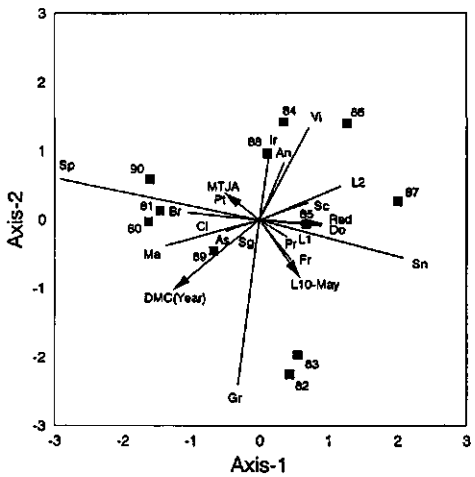


Fig. 2. Reduced rank regression biplot. Lines refer to varieties (for codes see Table 5), and squares to years. The arrows represent the coefficients for the environmental variables within the reduced rank regression axes (for codes see text).

The biplot of the reduced rank regression analysis (Fig. 2) resembled the AMMI biplot. Splenda, Gracia, Sonia seem as highly interactive as in the AMMI biplot, because of their distance from the origin. LG 20 80 and Markant exhibited interaction to a lesser degree. The relative positions of the years have shifted a little bit in comparison to the AMMI biplot. The highly differentiating years 82, 83, and 90 (Table 2) are recognizable by their large distance from the origin. Another differentiating year, 89, however, is placed relatively close to the origin, whereas the not very differentiating years 80 and 81 are located at great distance. It must be remarked that the scaling chosen for the AMMI and reduced rank regression biplots, with the scaling constant c equal to one (see van Eeuwijk, 1995a), tends to emphasize the quality of the representation of the varieties, somewhat at the cost of that of the years. Important interactions between varieties and years according to the reduced rank regression biplot are the following: Splenda did relatively well in 80, 81, and 90, relatively bad in 82, 83, 85, 86, and 87. Sonia behaved exactly contrary to Splenda. Gracia did relatively well in 82 and 83, bad in 84 and 86. Vivia formed the counterpart of Gracia.

The correlations of DMC(Year) with the reduced rank regression axes were comparable to those with the

Table 10. Correlations between environmental variables and AMMI axes, coefficients of the environmental variables in the reduced rank regression axes, and correlations between the environmental variables and the reduced rank regression axes

Variable	AMMI axes		Reduced Rank Regression axes			
	Correlations		Coefficients		Correlations	
	Axis 1	Axis 2	Axis 1	Axis 2	Axis 1	Axis 2
DMC(Year)	- 0.45	- 0.62	- 0.33	- 0.27	- 0.56	- 0.74
Rad	0.05	- 0.45	0.24	- 0.02	0.05	- 0.63
L10-May	0.47	- 0.20	0.15	- 0.22	0.55	- 0.32
MTJA	- 0.22	0.29	- 0.14	0.10	- 0.26	0.34

AMMI axes (Table 10). Therefore, years are diagonally arranged from lower left to upper right from generally *favorable with high* DMC(Year) *to less favorable with* low DMC(Year). As in the AMMI biplot, varieties are arranged in the same direction from mid late and mid *early flowering to early and very early flowering.* The Variety by Year interaction in the reduced rank regression biplot thus receives a similar interpretation as in the AMMI biplot. An extra feature of the reduced rank regression biplot is that in the same biplot approximations can be found to the coefficients for the factorial regression, thereby revealing more specifically which variables may be held responsible for the observed interactions per variety. Projecting the variety lines on the concomitant variable arrows, it can be seen that Splenda and Gracia should have the highest positive coefficients on DMC(Year), whereas Sonia, Vivia, and Lg 20 80 should have the largest negative coefficients on this variable. For the variable Rad, Sonia seems to have a high positive coefficient, whereas Markant is clearly negative and Splenda highly negative. With respect to L10-May, Gracia and Sonia have positive coefficients, and Splenda negative. Lastly, for MTJA, Splenda should have a highly positive coefficient.

Reduced rank regression combines features of AMMI with those of factorial regression. Technically it can be interpreted as a lower rank approximation to the matrix of factorial regression coefficients. Basically, the coefficients are approximated using fewer degrees of freedom. Table 8 gives the reduced rank regression approximations besides the factorial regression coefficients. The approximations were quite reasonable. Correlations between full rank factorial regression coefficients and reduced rank approximations, over the 18 varieties, were for DMC(Year), 0.98; Rad, 0.82; L10-May, 0.92; and MTJA, 0.73. The approximations were best for the most important differentiating variable, DMC(Year).

Mixed model including indicator and continuous variables

Instead of using the completely fixed approach embodied in full and reduced rank factorial regression, for further investigation of the genotype by environment interaction, we could have extended the mixed model approach by introducing analogous multiplicative terms for interactions, thus hoping to combine the best of both worlds. The mixed model formulation, however, does not lend itself very well to variable selection procedures, because of its heavy computational demands. Nevertheless, the mixed model approach certainly can be used as an extra verification of the results obtained from full and reduced rank factorial regression. We fitted the following mixed model: fixed terms = Variety + Location + Variety·Location + DMC(Year) + Variety·DMC(Year) + Variety·Rad + Variety·MTJA + Variety·L10-May; random terms = Year·Location + Variety·Year + Error. The variable DMC(Year) in the fixed part of the model replaces the random factor Year used earlier. We expected the (fixed) interaction of Variety with the explanatory environmental variables to take up part of the (random) Year·Location interaction of the original mixed model. Also we expected the coefficients of the varieties with respect to the environmental variables to be similar to those of the factorial regression. The estimates for the interaction variance components for this mixed model with concomitant variables were Variety by Year 7.56 (3.40) (was 15.40 (3.10)), and Year by Location 187.1 (46.2) (was 222.3 (60.0)). The change in the Year by Location component is relatively small. The Variety

by Year component has, however, been reduced by a factor two. This is in accordance with the fact that factorial regression with the same variables in the interaction explained half of the Variety by Year interaction (Table 7). The correlation between the factorial regression coefficients and the mixed model interaction coefficients was for DMC(Year) 0.82, Rad 0.74, L10-May 0.60, and MTJA 0.67. The results of the mixed model analysis with multiplicative interaction thus were in good agreement with those of the factorial regression analysis.

Discussion

The analyses described in the previous section illustrate how various methods can supplement and complement each other in an analysis of genotype by environment interaction, thereby eventually leading to a reasonably coherent picture of the interaction as differential genotypic sensitivity to a number of identified environmental variables. Each method has its own merits and weaknesses, and each method represents a specific way of looking at the phenomenon of genotype by environment interaction.

The fitting constants analysis may be seen as a starting point for getting a first, rough idea of how the variation is spread over the various terms. For our DMC data variation due to environments (Year times Location) was far more important than that due to varieties. However, this is to be expected in Variety List trials. More interesting are variety by environment interactions. Of these only Variety by Year interaction was judged to be present. The strong point of fitting constants is that it can always be applied, whatever the level of unbalancedness. A condition for its use in case of severe unbalancedness is, that it must be reasonably clear what the order of importance of the terms to be fitted is. Fitting constants is often used to find adjusted means for incomplete genotype by environment tables, with the environments consisting of years times locations. Because parameters for interaction are then undefined for empty cells, these adjusted means have to be based on an additive model (Patterson, 1978; Patterson & Nabugoomu, 1992), which may be an oversimplification.

Mixed model analysis is a natural alternative to fitting constants as soon as there are good reasons for distinguishing more than one random term. With the fitted mixed model, tables of best linear unbiassed predictions can be calculated. Mixed models with interaction can still be fitted to incomplete variety by environment tables, provided that the interaction is taken random (see van Eeuwijk, 1995a). Thus for our DMC data a complete Variety by Year table of best linear unbiassed predictions could be calculated, although a number of Variety by Year cells was empty.

The problem of missing values may be solved in a number of ways (Freeman, 1975; Gabriel & Zamir, 1979; Gauch & Zobel, 1990; Denis, 1991; Denis & Baril, 1992; van Eeuwijk, 1995b), but the mixed model approach described above has two attractive properties. Firstly, it is implemented in major statistical software packages and is thus easily applicable. Secondly, mixed models underlying the estimates for missing cell values are probably more realistic than some of the models underlying other missing value procedures.

Having available a complete genotype by environment table the interaction can straightforwardly be further investigated by multiplicative models for interaction. A natural starting point is the AMMI model, because it does not require external information for fitting the model. Interpretation of the AMMI biplot, however, can be enhanced considerably by the introduction of extra information on genotypes and environments. For our Variety by Year table of DMC a reasonable interpretation of the interaction was already possible by adding flowering date information on the varieties and DMC(Year) values for the environments. One should be careful not to focus exclusively on interpretation of individual axes in AMMI (and reduced rank regression). It is the joint picture that emerges from the axes that should be interpreted. This joint picture was implicitly addressed by the importance given to the direction of DMC(Year) in the biplot.

Factorial regression can be used to investigate the interaction with respect to a relevant set of environmental variables. Four variables were picked up as having a link with the Variety by Year interaction. The factorial regression elucidates another part of the interaction than the AMMI model. Both viewpoints are synthesized in the reduced rank regression, which provides a powerful means of visualization of the interaction in all its facets; interaction due to specific genotypes and environments; individual genotypic sensitivities to environmental variables; and descriptions of environments in terms of environmental variables. AMMI analyses will reveal similar patterns as reduced rank regression analyses when the major interactions in the data are clearly related to the environmental variables chosen. If that is not the case, AMMI and reduced rank regression solutions will differ and one should recon-

For the broader class of GAMMI models, it is still possible to visualize the interactions by means of biplots, but their interpretation depends on the particular link function.

For RC association models it is the form of non-independence rather than non-additivity, that is relevant. Goodman (1991) defines two forms of non-independence.

First, where

$$\lambda_{ij} = \log\left(\frac{P_{ij}}{\alpha_i\beta_j}\right) = \sum_{k=1}^{K} \sigma_k \gamma_{ki}\delta_{kj},$$

and second, the log-odds ratio,

$$\pi_{ij} = \log\left(\frac{P_{ij}P_{st}}{P_{it}P_{sj}}\right) = \sum_{k=1}^{K} \sigma_k(\gamma_{ki} - \gamma_{ks})(\delta_{kj} - \delta_{kt}),$$

defined for the cells in the rows i and s, and the columns j and t. The scaled row parameter $\gamma'_{ki} = \gamma_{ki}\sigma_k$, can be interpreted as the slope of a weighted linear regression of the non-independence measure λ_{ij} on the column scores, δ_{kj}: $\Sigma_{j=1}^{J} \lambda_{ij}\delta_{kj} = \gamma'_{ki}$. When the γ'_{ki} are used as coordinates for the row points in a biplot, the squared distance between two row points approximates the non-independence between the two rows, because

$$\sum_{k=1}^{K} (\gamma'_{ki} - \gamma'_{ks})^2 = \sum_{j=1}^{J} (\lambda_{ij} - \lambda_{sj})^2.$$

Similar relations can be deduced for δ'_{kj} and γ_{ki}. Therefore, Goodman (1986, p. 269; 1991, p. 1107) recommends for displays of row points alone to use $\gamma'_{ki} = \gamma_{ki}\sigma_k$, and for column points alone $\delta'_{kj} = \delta_{kj}\sigma_k$.

For simultaneous displays, the recommendation is to use $\gamma^*_{ki} = \gamma_{ki}\sigma_k^{1-c}$ and $\delta^*_{kj} = \delta_{kj}\sigma_k^c$ ($0 \leq c \leq 1$), where choice of c depends on whether the emphasis is on rows or columns. The inner product of row and column points in a simultaneous biplot approximates the non-independence measure λ_{ij} when γ and δ are scaled as γ^* and δ^*, as can be seen from

$$\lambda_{ij} = \log\left(\frac{P_{ij}}{\alpha_i\beta_j}\right) =$$

$$\sum_{k=1}^{K} \sigma_k\gamma_{ki}\delta_{kj} = \sum_{k=1}^{K} \gamma^*_{ki}\delta^*_{kj} = |\boldsymbol{\gamma}^*_i|\,|\boldsymbol{\delta}^*_j|\cos(\boldsymbol{\gamma}^*_i, \boldsymbol{\delta}^*_j), \tag{5}$$

where $\boldsymbol{\gamma}^*_i$ and $\boldsymbol{\delta}^*_j$ denote vectors of length K. In the same biplot, the inner product of a difference of row points with a difference of column points approximates the log-odds ratio

$$\pi_{ij} = \log\left(\frac{P_{ij}P_{st}}{P_{it}P_{sj}}\right) = \sum_{k=1}^{K} \sigma_k(\gamma_{ki} - \gamma_{ks})(\delta_{kj} - \delta_{kt}) \tag{6}$$

$$= \sum_{k=1}^{K} (\gamma^*_{ki} - \gamma^*_{ks})(\delta^*_{kj} - \delta^*_{kt}) = |\boldsymbol{\gamma}^*_i - \boldsymbol{\gamma}^*_s|\,|\boldsymbol{\delta}^*_j - \boldsymbol{\delta}^*_t|\cos(\boldsymbol{\gamma}^*_i - \boldsymbol{\gamma}^*_s, \boldsymbol{\delta}^*_j - \boldsymbol{\delta}^*_t),$$

with $\boldsymbol{\gamma}^*_i$, $\boldsymbol{\gamma}^*_s$, $\boldsymbol{\delta}^*_j$, and $\boldsymbol{\delta}^*_t$ vectors of length K. Simultaneous biplots thus provide a powerful tool for visualizing non-independence in two-way tables of counts analyzed by means of RC association models.

For other GAMMI models the interpretation of the biplot relations has still to be investigated. Nevertheless, distances between points of either rows or columns will always indicate some form of non-additivity or non-independence. Simultaneous displays should be interpreted with more care, but here the inner products of row and column points will still approximate non-additivity on the linear predictor scale.

6. An Application of a Log-Bilinear Model to Counts of Potato Cyst Nematodes on Potatoes

Table 1 gives the number of newly formed cysts on 11 potato genotypes for five potato cyst nematode populations belonging to the species *Globodera pallida* (part of a larger table in Arntzen and van Eeuwijk (1992)). The numbers are means over four or five replicates. An assessment of the genetic similarity of the potato cyst nematode populations was required. We will concentrate mainly on illustration of the possibilities for modeling and visualizing the interaction with an appropriate GAMMI model.

Table 1
Mean number of cysts of five nematode populations on eleven potato genotypes

	Nematode population				
Potato genotype	P2-22	Rookmaker	75-884-4	74-768-20	1077
12380	2.8	18.6	5.6	136.8	8.2
Vantage	2.6	31.8	5.4	107.2	17.0
AM78-3778	4.5	5.25	2.0	9.8	6.2
$(VT^{n})^2$ 62-33-3	19.6	96.2	85.4	307.0	29.4
Desiree	333.0	402.6	339.0	549.6	374.0
Maritta	421.0	426.8	221.8	630.4	260.8
Astarte	118.4	122.4	145.4	102.4	106.2
Elles	3.6	58.0	43.8	123.4	9.6
Pansta	23.4	199.0	208.4	83.8	104.6
Mara	134.2	216.0	257.2	277.4	204.4
Saturna	269.0	384.6	361.6	477.8	356.8

Using the alternating GLM algorithm described above, a sequence of nested models was fitted. The link function was chosen to be the natural log and the distribution to be the Poisson distribution. The differences in deviance corresponding to the inclusion of individual terms are given in Table 2.

Table 2
Deviance table for nematode data

Source	Degrees of freedom	Deviance	Mean deviance
PCN population	4	690.57	127.64
Potato genotype	10	7111.38	711.14
Axis 1	13	715.96	55.07
Axis 2	11	351.10	31.92
Axis 3	9	43.72	4.86
Residual	7	34.62	4.94
Total	54	8947.35	

The criterion for inclusion of a multiplicative interaction term (axis) was an F-test for the ratio of the mean deviance for the particular interaction term to a current estimate for the over-dispersion, the latter being either the mean deviance of the rest or Pearson's chi-square divided by the residual degrees of freedom:

$$\left\{\sum_i \sum_j \frac{(y_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}\right\} / df_{\text{residual}}.$$

In the following examples, the difference between these estimators for over-/under-dispersion was negligible. The estimate for the over-dispersion for the model with three multiplicative axes on the basis of the residual deviance was 4.94; on the basis of Pearson's chi-square it was 5.05. Table 2 shows that a model with two axes (GAMMI-2 model) was satisfactory, because the ratio of the mean deviance for the third axis to the residual mean deviance for a model with three axes was 4.86/4.94 = .98, which is not significant when compared with the $P \leq .05$ point of an $F_{[9;7]}$ distribution. The residual deviance estimate for over-dispersion for the GAMMI-2 model was (43.72 + 34.62)/(9 + 7) = 4.90; the Pearson chi-square estimate was 4.91. Plots of Pearson and deviance residuals against fitted values and linear predictor values revealed no severe anomalies. A plot of the link-adjusted dependent variate,

Table 4
Analysis of variance for AMMI-3 model fitted to two-way table of logit transformed Fusarium head blight incidences in winter wheat

Source	Degrees of freedom	Sum of squares	Mean square
Year-Strain combination	11	224.09	20.37
Wheat genotype	16	171.60	10.73
Axis 1	26	28.17	1.08
Axis 2	24	23.20	0.97
Axis 3	22	9.11	0.41
Residual	104	30.60	0.29
Total	203	486.77	

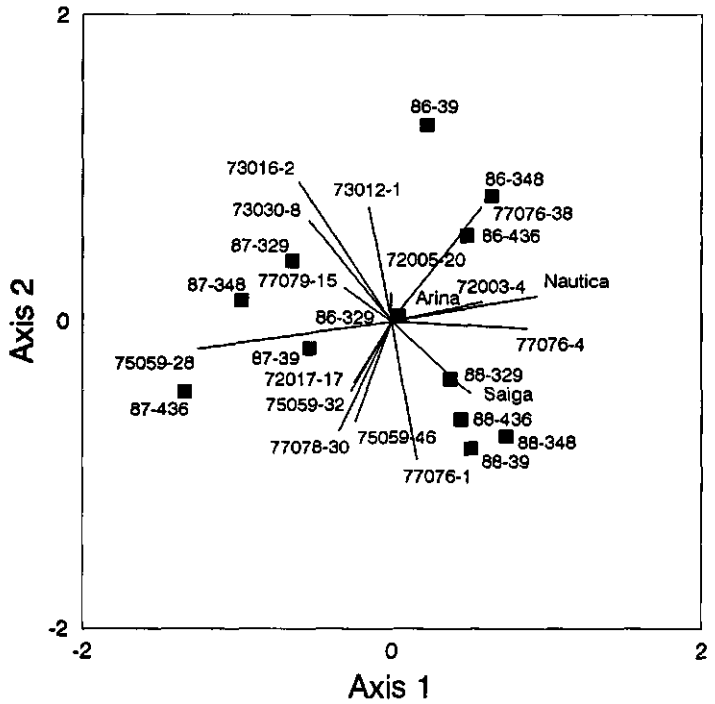


Figure 2. Biplot of the interaction in the *Fusarium* head blight data following from AMMI-2 model on logit transformed data. Year by strain combinations are represented by squares; wheat genotypes by lines.

interactions can be tested somewhat more rigorously by performing a three-way analysis of variance with the factors genotype, strain, and year on the table of mean incidence proportions (logit transformed). As an estimate for the error we used the three-factor interaction. From Table 5 we see that the mean square for the three factor interaction, .33, was almost equal to our estimate for the error derived from the AMMI-2 analysis, (9.11 + 30.60)/(22 + 104) = .32. Hence there are good reasons to assume that there was indeed no "real" three-factor interaction. Genotype by strain interaction was not found significant, and the resistance may therefore be classified as horizontal. Furthermore, the conjecture of significant genotype by year interaction is also confirmed by Table 5. The AMMI analysis combined with the three-way ANOVA gives a reasonably clear picture of what is happening. A plot of standardized residuals versus fitted values shows no notable anomalies (Figure 3a).

Next, we analyzed the same data with a GAMMI model with a logit link and a binomial

Table 5
Three-way analysis of variance for logit transformed Fusarium head blight incidences

Source	Degrees of freedom	Sum of squares	Mean square
Year	2	33.89	16.49
Fusarium strain	3	130.25	43.42
Strain-Year	6	59.96	9.99
Wheat genotype	16	171.60	10.73
Genotype-Year	32	43.86	1.37
Genotype-Strain	48	15.31	.32
Error	96	31.91	.33
Total	203	486.77	

distribution. Interaction was again describable by two multiplicative axes. However, a plot of (quasi-) deviance residuals against linear predictor was far from satisfactory (Figure 3b). The situation encountered here is clearly reminiscent of Wedderburn's leaf-blotch data on barley as treated in McCullagh and Nelder (1989, pp. 328–332) (see also Wedderburn, 1974). Therefore, a GAMMI model was fitted with logit link and with variance function $V(\mu) = \mu^2(1-\mu)^2$, where μ is a proportion. The quasi-likelihood function corresponding to this combination of link and variance function is

$$Q(\mu;y) = \left[(2y-1)\log\left(\frac{\mu}{1-\mu}\right) - \frac{y}{\mu} - \frac{1-y}{1-\mu}\right] - \left[(2y-1)\log\left(\frac{y}{1-y}\right) - 2\right]/\phi. \tag{7}$$

The function is not defined for $\mu = 0$ or 1, or $y = 0$ or 1. We had, however, no observations being equal to 0 or 1, so will ignore this complication. Because the iterative weights are exactly unity, this type of GLM can be implemented particularly easily in our alternating GLM scheme. For an ordinary, i.e., non-alternating, GLM, an unweighted, but, still iterative, regression can be performed with as dependent variate $z = \hat{\eta} + (y - \hat{\mu})/\hat{\mu}(1-\hat{\mu})$, in which $\hat{\eta}$ and $\hat{\mu}$ are the last updates for η and μ. The unity weights preserve the orthogonality properties in the iterative regression. These results carry over to our alternating GLM algorithm.

The results in Table 6 of the quasi-likelihood analysis were in close agreement with those of the AMMI analysis on the logit transformed data as given in Table 4. This similarity is also obvious from the comparison of the biplots (Figure 4, $c = .5$, and Figure 2) and the residual plots (Figures 3c and 3a). Table 6 indicates that two axes suffice. The association parameters for the GAMMI-2 model were $\sigma_1 = 5.20$ and $\sigma_2 = 4.57$.

The agreement between the quasi-likelihood analysis and the AMMI analysis on the logit transformed data is not surprising. The logit transform applied to the data themselves worked quite well when judged by its residual plot (Figure 3a), which exhibited good constancy of variance. Besides, a half-normal plot of the residuals (not shown) also was satisfactory, indicating approximate normality. On these grounds it might have been expected that the quasi-likelihood analysis used here would also do rather well. Because for both methods the weights are unity, the only difference between them is the use of a logit transform for the expected values (quasi-likelihood) instead of a logit transform for the observations (AMMI). In other situations the analyses may differ. For example, Wedderburn (1974) found differences, which were caused by observations either very close to one or zero.

8. Discussion

Estimation of main and interaction effects for the AMMI model is simple in case of a complete table, i.e., first fit a two-way ANOVA model to estimate the main effects and subsequently perform a singular value decomposition on the residuals (Gabriel, 1978). For incomplete tables Gabriel and Zamir (1979) proposed an iterative procedure of alternating row and column regressions with unit weights for presence and zero weights for absence. They also presented the idea of using weights inversely related to Poisson variance in the weighted singular value decomposition of a log transformed matrix of counts. This is very close to the GAMMI estimation procedure described in Section 3. In contrast to the fixed weights chosen by Gabriel and Zamir (1979), Gabriel and Odoroff (1984) proposed the use of iterative reweighing in a procedure for a more robust form of singular

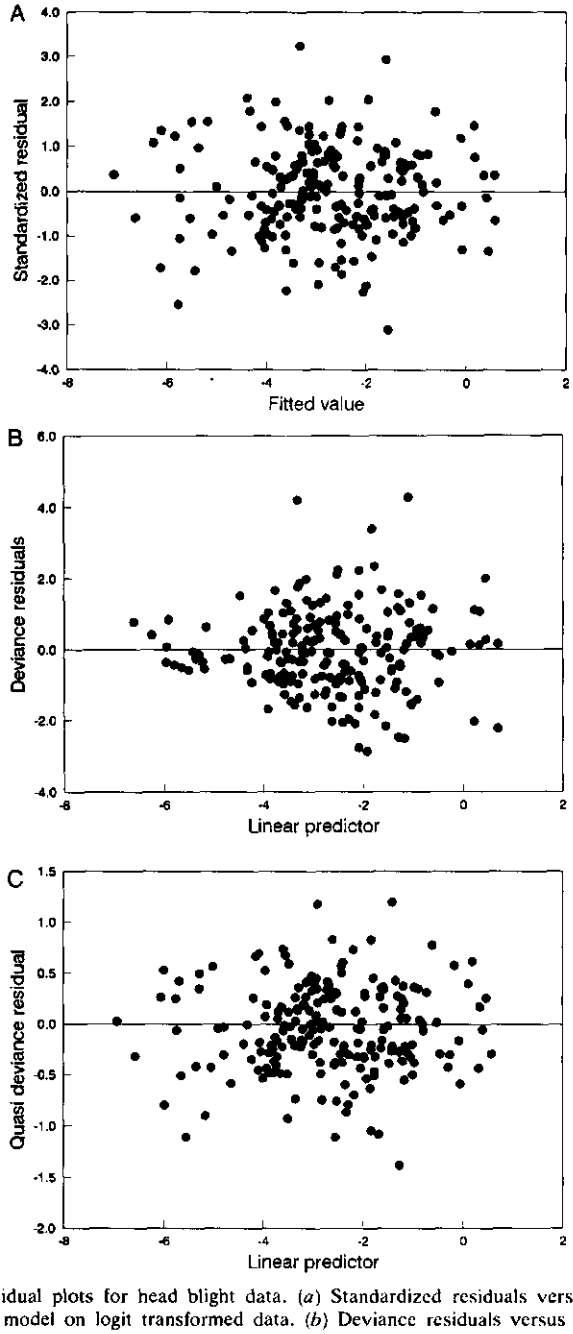


Figure 3. Residual plots for head blight data. (*a*) Standardized residuals versus fitted values from AMMI-2 model on logit transformed data. (*b*) Deviance residuals versus linear predictor from GAMMI-2 model with logit link and binomial distribution. (*c*) Quasi-deviance residuals versus linear predictor for GAMMI-2 model with logit link and variance function $V(\mu) = \mu^2(1 - \mu)^2$ (μ is expressed as a proportion).

Table 6
Quasi-deviance table for Fusarium head blight incidences

Source	Degrees of freedom	Quasi-deviance	Mean quasi-deviance
Year-strain combination	11	150.85	13.71
Wheat genotype	16	145.26	9.08
Axis 1	26	26.13	1.01
Axis 2	24	19.49	.81
Axis 3	22	8.77	.40
Residual	104	29.36	.28
Total	203	379.86	

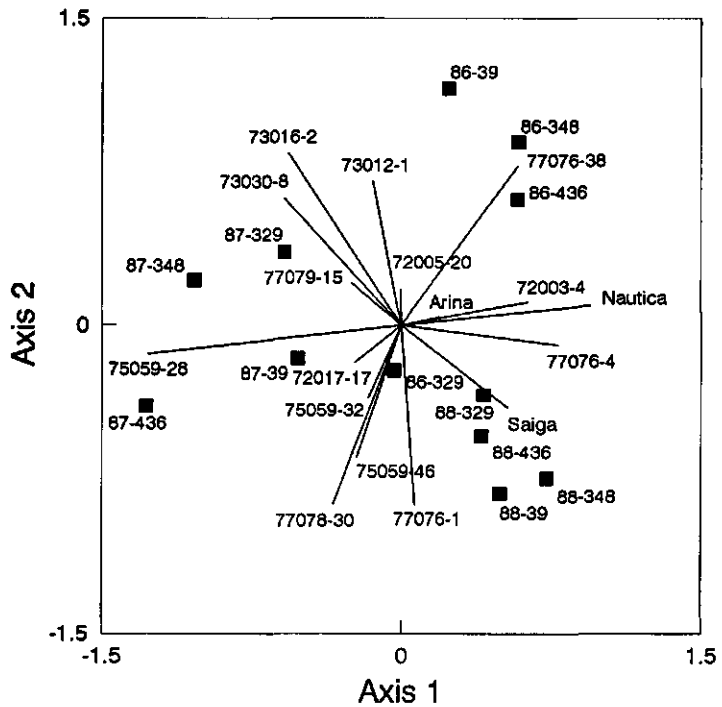


Figure 4. Biplot of the interaction in the *Fusarium* head blight data following from GAMMI-2 model with logit link and variance function $V(\mu) = \mu^2(1 - \mu)^2$ (μ is expressed as a proportion). Year by strain combinations are represented by squares; wheat genotypes by lines.

value decomposition. McNeil and Tukey (1975) elaborated a similar idea in the context of the diagnosis of models for two-way tables.

Though Gabriel and Zamir (1979) were already close to the use of alternating GLMs, Pettitt (1989) seems the first to explicitly propose iteration of alternating GLMs in a generalization of Tukey's one-degree-of-freedom test for non-additivity, which can be interpreted as a test for the presence of a restricted form of multiplicative interaction, namely the form for which there is only one multiplicative term and for which the row and column scores are proportional to the row and column main effects. In contrast to Pettitt, who left the possibilities for generalization to GAMMI unexplored, De Falguerolles and Francis (1992) acknowledged the possibility of using alternating GLMs for estimation of row and column parameters for a multiplicative interaction model. Their brief description, however, ignored the necessity of using offsets.

The use of offsets is essential for arriving at the maximum quasi-likelihood solution. Otherwise row and column regressions may reach different maxima, and the iterative scheme will continue

cycling between these two maxima. For weighted AMMI estimation (Gabriel and Zamir, 1979) the use of offsets reduces to subtraction of the column effects before the row regression and vice versa. Failure to use offsets will inevitably lead to convergence problems as observed by Gabriel and Zamir. Denis (1991) described an alternating least squares algorithm for weighted AMMI estimation including offsets.

The alternating GLM scheme can be extended to cover another generalization of the AMMI model; the reduced rank regression model (Davies and Tso, 1982; van Eeuwijk, 1992b). Reduced rank regression models can be derived from AMMI models by imposing the restriction on the scores that they have to be linear combinations of explicitly measured variables (Rao, 1964). AMMI model interaction scores can be forced to be linear combinations of explicitly measured variables simply by regressing them on the relevant variables. Thus row scores, γ_{ki}, are replaced by their fitted values from a regression on the appropriate explanatory variables between steps 2 and 3 (Sections 3.2 and 3.3) of the estimation scheme. Column scores, δ_{kj}, are likewise replaced by their fitted values between steps 4 and 5 (Sections 3.4 and 3.5). The generalization of reduced rank regression to generalized reduced rank regression will be obvious. With this generalization a wide range of techniques from AMMI analysis to the RC association model analog of canonical correspondence analysis (Ter Braak, 1986) can be covered by the same estimation procedure.

Acknowledgements

The author thanks Bertus Keen, Bas Engel, Jean-Baptiste Denis, Ritsert Jansen, and two referees for comments on drafts and discussion, and L. C. Paul Keizer for the graphics.

Résumé

Les modèles multiplicatifs bilinéaires ou biadditifs d'interaction dans des tableaux à deux facteurs fournissent les moyens essentiels d'étude des problèmes d'interaction genotype environnement. Dans les applications, les suppositions classiques sont celles d'une erreur Normalement distribuée et d'une fonction d'identite de lien. Elles sont inutilement restrictives. L'introduction de termes multiplicatifs pour l'interaction dans GLMs les supprime. On obtient les estimations des paramètres par un processus itératif alternant des régressions généralisées sur lignes et colonnes dans le cadre d'une quasi-vraisemblance. Les exemples les plus connus de cette classe de modèles généralisés d'effets principaux additifs et d'effets d'interaction multiplicatifs (GAMMI) sont les modèles AMMI (Gauch, 1988, *Biometrics* **44,** 705–715) et les modèles d'association RC de Goodman (Goodman, 1981, *Journal of the American Statistical Association* **76,** 320–334). On peut visualiser la partie multiplicative de l'interaction à l'aide de "biplots". On présente deux applications des modèles GAMMI sur des données provenant d'expériences de croisement de plantes. La première illustration traite d'un modèle log-bilinéaire pour des données de dénombrement avec variation poissonienne. La seconde concerne un modèle logit-bilinéaire pour des données d'incidence d'une maladie avec un type particulier de fonction de variance, extension d'un modèle présenté par Wedderburn (1974, *Biometrika* **61,** 439–447).

References

Arntzen, F. K. and van Eeuwijk, F. A. (1992). Variation in resistance level of potato genotypes and virulence of potato cyst nematode populations. *Euphytica* **62,** 135–143.

Cornelius, P. L. (1993). Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. *Crop Science* **33,** 1186–1193.

Davies, P. T. and Tso, M. K.-S. (1982). Procedures for reduced-rank regression. *Applied Statistics* **31,** 244–255.

De Falguerolles, A. and Francis, B. (1992). Algorithmic approaches for fitting bilinear models. In *Proceedings Compstat Neuchatel*, pp. 77–82.

Denis, J.-B. (1991). Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Rev. Statistique Appliquée* **XXXIX,** 5–24.

Denis, J.-B. and Gower, J. C. (1992). *Biadditive models.* Technical Report. Laboratoire de Biométrie, INRA-Versailles.

Denis, J.-B. and Gower, J. C. (1994). Biadditive models. Letter to the editor. *Biometrics* **50,** 310–311.

Finlay, K. W. and Wilkinson, G. N. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research* **14,** 742–754.

Freeman, G. H. and Dowker, B. D. (1973). The analysis of variation between and within genotypes and environments. *Heredity* **30,** 97–109.

Gabriel, K. R. (1971). Biplot display of multivariate matrices with application to principal component analysis. *Biometrika* **58,** 453–467.

Gabriel, K. R. (1978). Least squares approximation of matrices by additive and multiplicative models. *Journal of the Royal Statistical Society, Series B* **40,** 186–196.

Gabriel, K. R. and Odoroff, C. L. (1984). Resistant lower rank approximations of matrices. In *Data Analysis and Informatics, III*, E. Diday et al. (ed). North-Holland: Elsevier Science Publishers B. V.

Gabriel, K. R. and Zamir, S. (1979). Lower rank approximations of matrices by least squares with any choice of weights. *Technometrics* **21,** 489–498.

Gauch, H. G., Jr. (1988). Model selection and validation for yield trials with interaction. *Biometrics* **44,** 705–715.

Gauch, H. G., Jr. (1992). *Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs*. Amsterdam: Elsevier.

Gauch, H. G., Jr. and Zobel, R. W. (1988). Predictive and postdictive success of statistical analyses of yield trials. *Theoretical and Applied Genetics* **76,** 1–10.

Gauch, H. G., Jr. and Zobel, R. W. (1989). Accuracy and selection success in yield trial analyses. *Theoretical and Applied Genetics* **77,** 473–481.

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

Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association* **74,** 537–552.

Goodman, L. A. (1981). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association* **76,** 320–334.

Goodman, L. A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. *International Statistical Review* **54,** 243–309.

Goodman, L. A. (1991). Measures, models and graphical displays in the analysis of cross-classified data. *Journal of the American Statistical Association* **86,** 1085–1138.

Goodman, L. A. and Haberman, S. J. (1990). The analysis of nonadditivity in two-way analysis of variance. *Journal of the American Statistical Association* **85,** 139–145.

Hegemann, V. J. and Johnson, D. E. (1976). On analyzing two-way AOV data with interaction. *Technometrics* **18,** 273–281.

Hinkley, D. (1985). Transformation diagnostics for linear models. *Biometrika* **72,** 487–496.

Johnson, D. E. and Graybill, F. A. (1972). An analysis of a two-way model with interaction and no replication. *Journal of the American Statistical Society* **67,** 862–868.

Kempton, R. A. (1984). The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science, Cambridge* **103,** 123–135.

Mandel, J. (1961). Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **56,** 878–888.

Mandel, J. (1969). The partitioning of interaction in analysis of variance. *Journal of Research of the National Bureau of Standards—B. Mathematical Sciences* **73B,** 309–328.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.

McNeil, D. R. and Tukey, J. W. (1975). Higher-order diagnosis of two-way tables, illustrated on two sets of demographic empirical distributions. *Biometrics* **31,** 487–510.

Milliken, G. A. and Johnson, D. E. (1989). *Analysis of Messy Data. Volume 2: Nonreplicated Experiments*. New York: Van Nostrand Reinhold.

Pettitt, A. N. (1989). One degree of freedom for nonadditivity: Applications with generalized linear models and link functions. *Biometrics* **45,** 1153–1162.

Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhya A* **26,** 329–358.

Snedecor, G. W. and Cochran, G. (1980). *Statistical Methods*. 7th ed. Ames, Iowa: The Iowa State University Press.

Snee, R. D. (1982). Nonadditivity in a two-way classification: is it interaction or nonhomogeneous variance? *Journal of the American Statistical Association* **77,** 515–519.

Snijders, C. S. and van Eeuwijk, F. A. (1991). Genotype x strain interactions for resistance to *Fusarium* head blight caused by *Fusarium culmorum* in winter wheat. *Theoretical and Applied Genetics* **81,** 239–244.

Ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67,** 1167–1179.

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

Vanderplank, J. E. (1984). *Disease Resistance in Plants*. 2nd ed. New York: Academic Press.

van Eeuwijk, F. A. (1992a). Multiplicative models for genotype-environment interaction in plant breeding. *Statistica Applicata* **4,** 393–406.

van Eeuwijk, F. A. (1992b). Interpreting genotype-by-environment interaction using redundancy analysis. *Theoretical and Applied Genetics* **85,** 89–100.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. *Biometrika* **61,** 439–447.

Williams, E. R. and Wood, J. T. (1993). Testing the significance of genotype-environment interaction. *Australian Journal of Statistics* **35,** 359–362.

Yates, F. and Cochran, W. G. (1938). The analysis of groups of experiments. *Journal of Agricultural Science, Cambridge* **28,** 556–580.

Yochmowitz, M. G. and Cornell, R. G. (1978). Stepwise tests for multiplicative components of interaction. *Technometrics* **20,** 79–84.

Zobel, R. W., Wright, M. J., and Gauch, H. G., Jr. (1988). Statistical analysis of a yield trial. *Agronomy Journal* **80,** 388–393.

Received August 1993; *revised February and July* 1994; *accepted August* 1994.

Chapter XIII

Assessing non-specificity of resistance in wheat to head blight caused by inoculation with European strains of *Fusarium culmorum*, *F. graminearum* and *F. nivale* using a multiplicative model for interaction

F.A. van Eeuwijk[1], *A. Mesterhazy*[2], *Ch. I. Kling*[3], *P. Ruckenbauer*[4], *L. Saur*[5], *H. Bürstmayr*[4], *M. Lemmens*[4], *L.C.P. Keizer*[1], *N. Maurin*[5] & *C.H.A. Snijders*[1]

[1] *DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)*
P.O.Box 16, 6700 AA Wageningen, The Netherlands

[2] *Cereal Research Institute (CRI)*
H-6701 Szeged, pF. 391, Hungary

[3] *University Hohenheim, State Plant breeding Institute*
D-70593 Stuttgart, Germany

[4] *Universität für Pflanzenbau und Pflanzenzüchtung*
Gregor-Mendel-Strasse 33, A-1180 Vienna, Austria

[5] *INRA Centre de Recherches de Rennes, Station de Pathologie Végétale*
BP 29, F-35650 Le Rheu, France

Theor Appl Genet (1995) 90: 221–228

F. A. van Eeuwijk · A. Mesterhazy · Ch. I. Kling
P. Ruckenbauer · L. Saur · H. Bürstmayr · M. Lemmens
L. C. P. Keizer · N. Maurin · C. H. A. Snijders

Assessing non-specificity of resistance in wheat to head blight caused by inoculation with European strains of *Fusarium culmorum*, *F. graminearum* and *F. nivale* using a multiplicative model for interaction

Received: 2 May 1994 / Accepted: 10 June 1994

Abstract To determine whether resistance to Fusarium head blight in winter wheat is horizontal and non-species specific, 25 genotypes from five European countries were tested at six locations across Europe in the years 1990, 1991, and 1992. The five genotypes from each country had to cover the range from resistant to susceptible. The locations involved were Wageningen, Vienna, Rennes, Hohenheim, Oberer Lindenhof, and Szeged. In total, 17 local strains of *Fusarium culmorum*, *F. graminearum*, and *F. nivale* were used for experimental inoculation. One strain, *F. culmorum* IPO 39-01, was used at all locations. Best linear unbiased predictions (BLUPs) for the head blight ratings of the genotypes were formed within each particular location for each combination of year and strain. The BLUPs over all locations were collected in a genotype-by-environment table in which the genotypic dimension consisted of the 25 genotypes, while the environmental dimension was made up of 59 year-by-strain-by-location combinations. A multiplicative model was fitted to the genotype-by-environment interaction in this table. The inverses of the variances of the genotype-by-environment BLUPs were used as weights. Interactions between genotypes and environments were written as sums of products between genotypic scores and environmental scores. After correction for year-by-location influence very little variation in environmental scores could be ascribed to differences between strains. This provided the basis for the conclusion that the resistance to Fusarium head blight in winter wheat was of the horizontal and non-species specific type. There was no indication for any geographical pattern in virulence genes. Any reasonable aggressive strain, a *F. culmorum* strain for the cool climates and a *F. graminearum* strain for the warmer humid areas, should be satisfactory for screening purposes.

Key words Head blight · Resistance breeding
Genotype-by-environment interaction · Multiplicative interaction · Host-specificity

Introduction

Fusarium head blight, a fungal disease of wheat and other small cereals is found in both temperate and semi-tropical regions. A number of species of *Fusarium* may be responsible, but generally *F. graminearum* Schwabe, with perfect stage *Gibberella zeae* (Schw.), and *F. culmorum* (W.G. Smith) Sacc., perfect stage unknown, predominate as the causal factor for *Fusarium* head blight (Lemmens et al. 1993, Mesterhazy 1977, Stack and McMullen 1985, Wilcoxson et al. 1988, Zadoks and Rijswijk 1984). Very exceptionally, *F. avenaceum* (Fr.) Sacc. has been reported to be highly pathogenic (Arseniuk et al. 1993) and *F. nivale* (Fr.) Ces. (Marasas et al. 1984) predominating (Daamen et al. 1991). Species frequencies are influenced by geography, climate and year. *Fusarium graminearum* and *F. culmorum* are non-host specific, i.e. they are pathogenic to wheat, maize and other cereals and grasses without showing specialization for any one crop. Nevertheless, some host prefences among *Fusarium* spp. have been observed

Communicated by J. MacKey

F. A. van Eeuwijk (✉) L. C. P. Keizer C. H. A. Snijders
DLO-Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16, NL-6700 AA Wageningen, The Netherlands

A. Mesterhazy
Cereal Research Institute (CRI),
H-6701 Szeged, pF. 391, Hungary

Ch. I. Kling
University Hohenheim, State Plant Breeding Institute,
D-70593 Stuttgart, Germany

P. Ruckenbauer H. Bürstmayr M. Lemmens
Universität für Bodenkultur Vienna (BOKU),
Institut für Pflanzenbau und Pflanzenzüchtung,
Gregor-Mendel-Strasse 33, A-1180 Vienna, Austria

L. Saur N. Maurin
INRA Centre de Recherches de Rennes,
Station de Pathologie Végétale,
BP 29, F-35650 Le Rheu, France

(Arseniuk et al. 1993). Significant interactions between strains of *F. graminearum* or *F. culmorum* and wheat genotypes have been reported (Mesterhazy 1984, 1988, Snijders and van Eeuwijk 1991). However, the interaction patterns were not stable over experiments and genotype ranking was only slightly influenced by the strains. No evidence has been found for the occurrence of races of *F. culmorum* or *F. graminearum* adapted to different wheat genotypes. To what extent resistance to *F. culmorum* is related to resistance to *F. graminearum* is not clear. Mesterhazy (1983, 1988) found correlation coefficients of up to 0.90 between the reaction of wheat genotypes to *F. culmorum* isolates and their reaction to *F. graminearum* isolates. Spring wheat genotypes which had been reported to be resistant to head blight caused by *F. graminearum* were also resistant to *F. culmorum* (Snijders 1990). Miedaner et al. (1993) concluded for 16 rye inbreds that the genetic basis of resistance to head blight caused by the two *Fusarium* species is very probably the same. Also, Arseniuk et al. (1993) concluded that the cereal resistance to a broad range of *Fusarium* spp. including the above two should be considered at the genus and not at the species level.

In this paper an international study using observations from six locations in five European countries will be presented. This study was initiated to assess the (non-)specificity of head blight resistance in wheat for *Fusarium* spp. and for strains within them, and was coupled to a geographical distribution of virulence genes. The variable selected to represent resistance was visually assessed Fusarium head blight rating. As the data were taken from ongoing research programmes that were not primarily developed towards the question addressed in this paper, the data had a rather complicated structure. Not all locations participated every year, at particular locations different genotypes and strains were used over the years, and experimental design differed between locations. Therefore, special statistical methodology had to be used in which a multiplicative model for the interaction of genotypes with environments (strain-by-location-by-year combinations), that took into account differences in precision, played a central role.

Materials and methods

Years and locations

During the years 1990, 1991 and 1992, 25 wheat genotypes were tested for resistance to Fusarium head blight (FHB). The genotypes were tested at six locations across Europe, namely Wageningen in the Netherlands, Gross-Enzersdorf near Vienna in Austria, Le Rheu near Rennes in France, Hohenheim (350 m altitude) and Oberer Lindenhof (600 m altitude) near Stuttgart in Germany and Szeged in Hungary (Fig. 1).

Genotypes

All genotypes were winter type wheats. From each participating country 5 genotypes covering the range from resistant to susceptible were tested. Variety names, line codes and origins are given in Table 1. The 5 Austrian lines were tested only in 1991 and 1992.

Fig. 1 The six locations and participating countries across Europe at which 25 winter wheat genotypes were tested for Fusarium head blight resistance. Wageningen in The Netherlands, Gross-Enzersdorf near Vienna in Austria, Le Rheu near Rennes in France, Hohenheim and Oberer Lindenhof near Stuttgart in Germany and Szeged in Hungary

Strains

Seventeen local strains of *F. culmorum*, *F. graminearum* and *F. nivale* were used for experimental inoculation (Table 2). The strain *F. culmorum* IPO 39–01 was used at all locations. In Wageningen, each year inoculum production was started up from an ampoule of lyophilized monospored spores. Vienna started up inoculum production from monospores and stored the cultures in earth medium culture (Lemmens et al. 1993). Rennes and Szeged stored the strains on Potato Dextrose Agar (PDA). The strains used in Oberer Lindenhof and Hohenheim were stored in earth medium culture. Wageningen produced inoculum consisting of purely conidiospores on a wheat/oat seed mixture (Snijders and van Eeuwijk 1991). In Rennes, conidiospores of *F. culmorum* were produced on autoclaved, soaked barley. *F. nivale* was produced on PDA at 10°C under near-ultra-violet light. Szeged and Vienna produced an inoculum suspension containing conidiospores and mycelium by continuous aeration of an inoculated liquid Czapek-Dox medium (Mesterhazy 1978). The German locations produced a suspension of conidiospores and mycelium by continuous aeration of an inoculated SNA medium (Nirenberg 1981).

Each participant used its own familiar inoculation method and assessment scale, listed in Table 3. Basically, the inoculation methods applied can be divided into two types. Type 1 method uses the method published by Mesterhazy (1978, 1983). Wheat lines are inoculated at anthesis by spraying 20 ml inoculum suspension containing spores and mycelium on separate bunches of 20–25 heads. Controls are treated with distilled water. The bouquets are then covered with a plastic bag for 24 h. Type 2 method is described by Snijders and van Eeuwijk (1991) and Saur (1991). Whole field plots are inoculated with conidiospores when 30% of the genotypes is flowering. This is repeated two or three times with intervals of 3–4 days until all genotypes are flowering. During the 2 weeks after inoculation a sprinkler irrigation guarantees a high relative humidity.

Table 1 Name, donor, origin, and mean head blight rating of the wheat varieties and lines

Wheat genotype	Donor	Origin	Mean (in %)
SVP 75059-28	CPRO-DLO, Wageningen, the Netherlands	CPRO-DLO	40
Arina	CPRO-DLO, Wageningen, the Netherlands	EFAP Zürich, Switzerland	31
SVP 72005-20-30-1	CPRO-DLO, Wageningen, the Netherlands	CPRO-DLO	67
SVP 72017-17-5-10	CPRO-DLO, Wageningen, the Netherlands	CPRO-DLO	31
SVP 75059-32	CPRO-DLO, Wageningen, the Netherlands	CPRO-DLO	56
NR-172/90	BOKU, Vienna, Austria	Saatzucht Neuhof/Rohrau, Austria	51
P 4371.88	BOKU, Vienna, Austria	Probstdorfer Saatzucht, Austria	64
P 2118.89	BOKU, Vienna, Austria	Probstdorfer Saatzucht, Austria	58
SL 8/80–28	BOKU, Vienna, Austria	Saatbau Linz, Austria	54
SL 34/81–12	BOKU, Vienna, Austria	Saatbau Linz, Austria	51
Copain	INRA, Rennes, France	Ets. C. Benoist, France	44
Rescler	INRA, Rennes, France	INRA	55
RC 103	INRA, Rennes, France	INRA	41
82 F3 28	INRA, Rennes, France	INRA	34
81 F3 79	INRA, Rennes, France	INRA	34
25/83/02	LSA, Hohenheim, Germany	LSA	64
47/83/02	LSA, Hohenheim, Germany	LSA	53
77/82/01	LSA, Hohenheim, Germany	LSA	62
163/81/03	LSA, Hohenheim, Germany	LSA	48
204/81/03	LSA, Hohenheim, Germany	LSA	58
Zombor GKI	CRI, Szeged, Hungary	CRI	65
Szoke GKI	CRI, Szeged, Hungary	CRI	48
Bence GKI	CRI, Szeged, Hungary	CRI	46
85-92 GKI	CRI, Szeged, Hungary	CRI	49
Sgv/GT-Pdj*UhrGK	CRI, Szeged, Hungary	CRI	47

Table 2 Name, species, donor and origin of *Fusarium* strains used in the experiments and presence of the strains at the locations. The code refers to the strain as represented in Fig. 2D

Code	Strain	*Fusarium* spp	Donor	Isolated from	Presence at location					
					Wageningen	Vienna	Rennes	Hohenheim	Ob Lindenhof	Szeged
1	IPO 39–01	*F. culmorum*	Wageningen	bread wheat (Flevina) kernel	+	+	+	+	+	+
2	SVP 8901	*F. culmorum*	Wageningen	bread wheat (Kanzler) head	+					
3	SVP 8904	*F. culmorum*	Wageningen	bread wheat (Beatrix) head	+					
4	91031	*F. graminearum*	Vienna	durum wheat kernel		+				
5	91047	*F. graminearum*	Vienna	durum wheat kernel		+				
6	91015	*F. culmorum*	Vienna	durum wheat kernel		+				
7	Le Rheu 89–4	*F. culmorum*	Rennes	bread wheat kernel			+			+
8	Le Rheu mix	*F. culmorum*	Rennes	bread wheat kernel			+			
9	F. nivale Fn002	*F. nivale*	Rennes	durum wheat (Durelle) kernel			+			
10	HOH 200/207 mix	*F. graminearum*	Hohenheim	bread wheat/durum wheat kernel				+	+	
11	HOH 214/223 mix	*F. culmorum*	Hohenheim	durum wheat head				+	+	
12	F.g. 12216	*F. graminearum*	Szeged	bread wheat kernel (6947/7)						+
13	F.g. 13377	*F. graminearum*	Szeged	maize crown						+
14	F.c. 12375	*F. culmorum*	Szeged	wheat crown						+
15	F.c. 12551	*F. culmorum*	Szeged	wheat lower stem						+
16	F.c. D 223	*F. culmorum*	Szeged (strain from Hoh.)	durum wheat head						+
17	F.g. D 207	*F. graminearum*	Szeged (strain from Hoh.)	durum wheat kernel						+

Data

Head blight symptoms were observed on different scales (Table 3). For analysis all Fusarium head blight (FHB) ratings were first expressed on a 0–100 scale of which subsequently the logit (=log(FHB/(100-FHB))) was taken. At Wageningen, in addition to the FHB rating, yield and thousand kernel weight reduction were measured as described in Snijders (1990). Also flowering date was observed. In Vienna the extra measurement concerned yield reduction and ear weight reduction based on ten heads. In Rennes, the extra measurements besides the FHB rating consisted of yield and thousand kernel weight loss, determined by comparison with the control: in 1990 and 1991, on a whole hill plot basis (Saur and Trottet 1992); in 1992, on basis of a sample of 40 heads. Also, the percentage damaged (pink) kernels was assessed based on a sample of 500 seeds. At the German locations, besides the FHB rating, yield components were determined based on ten heads and expressed as a percentage of the non-infected control (Mesterhazy 1978, 1983). In Szeged, the

Table 3 Inoculation method and Fusarium head blight (FHB) assessment scale used at each testing location

Location	Inoculation method	FHB-scale[a]	Observation after
Wageningen	type 2/field plots	0–100%	28 days
Vienna	type 1/bunch of heads	0.0–4.0	22 days
Rennes	type 2/field plots	1.0–9.0	450°days in 1990 350°days in '91/'92
Hohenheim	type 1/bunch of heads	1–9	31 days
Oberer Lindenhof	type 1/bunch of heads	1–9	22 days
Szeged	type 1/bunch of heads	0.0–4.0	18, 22, 26 and 30 days For analysis the means were used

[a] Lowest value indicates no symptoms; highest value indicates 100% infection

extra measurement concerned yield reduction based on ten heads (Mesterhazy 1978, 1983) and percentage grain infection estimated as percentage white/pink kernels.

FHB was chosen as the variable whose analysis had to elucidate the type of resistance. The factors whose effect had to be quantified before being able to answer that question were genotype, year, location and strain. FHB data were unbalanced with respect to these four factors taken together. The 5 Austrian genotypes were absent in all the trials of 1990. Also in 1990, the Austrian strains 91015, 91031 and 91047, and the French *F. nivale* were not used at any location. In 1991, strain 91015 was again absent, in this case together with the Le Rheu mix from France, the latter also being absent in 1992. Year-by-location combinations that were not available, were Vienna in 1990 and Oberer Lindenhof in 1992. For the presence of strain-by-location combinations, see Table 2.

General strategy of analysis

The general methodology chosen to answer the question on the type of resistance is an extension of an approach developed earlier towards the same problem in Snijders and van Eeuwijk (1991). Resistance will be defined as horizontal if no genotype-by-strain interactions can be found over years and locations. Firstly, analyses per location were done to determine whether genotype-by-strain interactions were stable over years within the individual locations. Mixed models, models with fixed and random terms (Searle 1971), were fitted per location. Parameters were estimated by residual maximum likelihood (Patterson and Thompson 1971) using Genstat (1993). Interest focussed on the genotype-by-strain interaction and the genotype-by-strain-by-year interaction. The presence of a genotype-by-strain interaction in combination with the absence of a genotype-by-strain-by-year interaction should indicate resistance of the strain-specific type.

For each location, genotype-by-environment two-way tables of best linear unbiased predictions, or BLUPs (Robinson 1991, Verdooren 1992), were calculated based on the fitted mixed model. The environments in these tables consisted of the combinations of strains and years present at a particular location. For some environments no BLUPs were available for the 5 Austrian genotypes. The genotype by environment tables per location were then combined over all six locations to give a 25-(genotypes) by-59 (environments=location-by-year-by-strain combinations; see Table 4) two-way table that served as the basis for the overall analysis. For the answer to our research question on the type of resistance only the interaction in this table was relevant. Let x_{ij} be the residual from additivity for the i-th genotype in the j-th environment, which is what is left of the BLUP after correction for the main effects of genotype and environment. The residual from additivity, x_{ij}, can be separated in structure and noise. One way to do that is by means of a singular value decomposition of the matrix of residuals from additivity. The singular value decomposition of the matrix with the entries x_{ij} writes each x_{ij} as the sum of a number of product terms:

$$x_{ij} = c_{1i} \times d_{1j} + c_{2i} \times d_{2j} + \ldots + c_{Ki} \times d_{Kj},$$

where c_{1i} to c_{Ki} represent the genotypic scores for genotype i, and d_{1j} to d_{Kj} the corresponding environmental scores for environment j. In principle, K is equal to the minimum of I-1 and J-1, with I the number of rows and J the number of columns. That is, for a 25×59 table, 24 product terms can be estimated. However, usually the first few product terms suffice for an adequate description of the interaction structure. The rest of the product terms are then collected in a residual representing noise. The residual from additivity is thus decomposed as

$$x_{ij} = c_{1i} \times d_{1j} + c_{2i} \times d_{2j} + \ldots + c_{Ni} \times d_{Nj} + e_{ij},$$

where the product terms express the structure, e_{ij} the noise, and N the number of products necessary for adequate description. Application of this method to complete two-way tables is rather straightforward (Gabriel 1978, Gauch 1988, Snijders and van Eeuwijk 1991). Because the Austrian genotypes were not present in 18 of the 59 environments and because it was deemed better to weigh the BLUPs by the inverse of their variances, an adapted method was used to estimate main effects and product terms simultaneously, taking into account incompleteness of the table and differential weighting of the entries (Denis 1991, van Eeuwijk 1995). Product terms that were retained as structure were those whose relative contribution to the interaction sum of squares exceeded the average per term of 4.2% (100% divided by 24, the latter number giving the number of product terms available for a table of 25 by 59; Jolliffe 1986).

For determining the type of resistance, only the environmental scores are of importance. If for the environmental scores the effect of years, locations, or their joint (interaction) effect dominates the effect of strains, this is an indication for non-specificity of the resistance. To be sure, also the joint (interaction) effect of strains and years, strains and locations, and strains, years and locations must be considered and proven negligible. The assessment of the effect of the different environmental factors on the environmental scores can take place in several ways. Very informally, one can plot the environmental scores (Kempton 1984), d_{2j} against d_{1j} to start with, and inspect the resulting plot on clustering of environments due to shared strains or other factors. If the environments cluster mainly on the basis of strains this is a strong argument for the existence of strain specific resistances. More formally, one can perform an analysis of variance on the environmental scores, treat d_{1j} to d_{Nj} as individual variables, and assess the importance of the strains after correction for years and locations. Finally, one can look at the environmental scores d_{1j} to d_{Nj} simultaneously, again interpreting them as variables, and carry out two discriminant analyses on them: firstly, using strains as a grouping factor and secondly, using year-by-location combinations. The 'variables' d_{1j} to d_{Nj} are used to construct discriminant functions. The number of discriminant functions that can be formed is equal to the minimum of the number of groups in the grouping factor minus 1 and the number of variables. These discriminant functions are linear combinations of the original variables, e.g. the first discriminant function can be calculated as $b_{11}d_{1j} + b_{12}d_{2j} + \ldots + b_{1N}d_{Nj}$. The weights b_{pq} for the q-th variable in the p-th discriminant function are chosen so that the between-groups variation is maximized with respect to the within variation. Roughly said, this means that the first discriminant function is that linear combination of the original variables that has the highest F value possible in an analysis of variance on basis of the grouping factor used for construction of the discriminant function. The second discriminant function is the linear combination giving the second highest F value under the re-

Table 4 Environmental means for head blight rating in percent. Code A and B refer to the location-year-strain combinations as represented in Fig. 2A and B

Plot code		Environment	Mean
A	B		
1	1	Wageningen 1990 IPO39-01	18
2	1	Wageningen 1990 SVP8901	4
3	1	Wageningen 1990 SVP8904	13
4	2	Wageningen 1991 IPO39-01	54
5	2	Wageningen 1991 SVP8901	22
6	2	Wageningen 1991 SVP8904	24
7	3	Wageningen 1992 IPO39-01	79
8	3	Wageningen 1992 SVP8901	13
9	3	Wageningen 1992 SVP8904	16
10	4	Vienna 1991 IPO39-01	86
11	4	Vienna 1991 91031	64
12	4	Vienna 1991 91047	89
13	5	Vienna 1992 IPO39-01	77
14	5	Vienna 1992 91015	48
15	5	Vienna 1992 91031	56
16	5	Vienna 1992 91047	66
17	6	LeRheu 1991 IPO39-01	68
18	6	LeRheu 1991 LeRheu 89–4	57
19	7	LeRheu 1992 IPO39-01	83
20	7	LeRheu 1992 LeRheu 89–4	68
21	8	LeRheu 1990 LeRheu mix	80
22	6	LeRheu 1991 F.nivale	51
23	7	LeRheu 1992 F.nivale	60
24	9	Hohenheim 1990 HOH 200/207 mix	71
25	9	Hohenheim 1990 HOH 214/223 mix	88
26	10	Hohenheim 1991 IPO39-01	97
27	10	Hohenheim 1991 HOH 200/207 mix	94
28	10	Hohenheim 1991 HOH 214/223 mix	96
29	11	Hohenheim 1992 IPO39-01	89
30	11	Hohenheim 1992 HOH 200/207 mix	79
31	11	Hohenheim 1992 HOH 214/223 mix	87
32	12	Oberer Lindenhof 1990 HOH 200/207 mix	58
33	12	Oberer Lindenhof 1990 HOH 214/223 mix	74
34	13	Oberer Lindenhof 1991 IPO 39-01	94
35	13	Oberer Lindenhof 1991 HOH 200/207 mix	82
36	13	Oberer Lindenhof 1991 HOH 214/223 mix	94
37	14	Szeged 1990 LeRheu 89-4	62
38	14	Szeged 1990 F.g.216	11
39	14	Szeged 1990 F.g.377	61
40	14	Szeged 1990 F.c.375	50
41	14	Szeged 1990 F.c.551	52
42	14	Szeged 1990 F.c. D 223	61
43	14	Szeged 1990 F.g. D 207	49
44	15	Szeged 1991 IPO39-01	51
45	15	Szeged 1991 LeRheu 89-4	31
46	15	Szeged 1991 F.g.216	17
47	15	Szeged 1991 F.g.377	16
48	15	Szeged 1991 F.c.375	46
49	15	Szeged 1991 F.c.551	24
50	15	Szeged 1991 F.c. D 223	20
51	15	Szeged 1991 F.g. D 207	43
52	16	Szeged 1992 IPO39-0101	4
53	16	Szeged 1992 LeRheu 89-4	5
54	16	Szeged 1992 F.g.216	3
55	16	Szeged 1992 F.g.377	24
56	16	Szeged 1992 F.c.375	43
57	16	Szeged 1992 F.c.551	10
58	16	Szeged 1992 F.c. D 223	6
59	16	Szeged 1992 F.g. D 207	4

striction that it is orthogonal to the first discriminant function, etc. After calculation, discriminant functions can be treated as ordinary variables. A mean for every group of environments can be calculated. Then one can check each of the environments for the group mean it is closest to and allocate the environment to that group. This is what we call the a posteriori classification. The original grouping is called the a priori classification. The principle does not change when we consider classification on more than one discriminant function. Large differences between group membership in a priori and a posteriori classifications shows that the grouping used for a priori classification does not make much sense with respect to the variation found in the variables used for construction of the discriminant functions. So, if the a priori classification according to strains differs considerably from the a posteriori classification there is not much reason to assume strain specificity of the resistance. An extra argument for that conclusion would be close correspondence between a priori and a posteriori classification on year-by-location basis.

Designs and analyses per location

Experimental design differed between locations and sometimes between strains per location. As mentioned above, mixed models were fitted to the data collected for a particular location or for the data collected for a particular location by strain combination, as for Rennes.

Wageningen

In each of the 3 years a split plot design was used with three replicates (blocks), strains as main plots and genotypes as sub plots (2.00×0.75 m) (Snijders and Van Eeuwijk 1991). The nine combinations of location, year and strain involved correspond to the environments 1–9 in Table 4. A mixed model was fitted to the complete set of data collected in Wageningen; that is, a model was fitted for all of the genotypes over the environments 1–9. BLUPs were calculated for each of the environments, i.e. for environments 1, 2 and 3 for 20 genotypes (Austrian genotypes absent in 1990), for environments 4–9 for all 25 genotypes. In the mixed model used the fixed terms were g + s + y + gs + gy + sy and the random terms gsy + yb + ybs + ybsg, where g stands for genotype, s for strain, y for year and b for block. Letters alone represent main effects, their combinations interactions.

Vienna

Vienna participated in 1991 and 1992, with 3 and 4 strains, respectively, thus defining the environments 10–16 in Table 4. In both years the wheat genotypes were sown in 10 m^2 plots. Each genotypic plot was then split in three parts or repeats (r). Within each repeat plants of the particular genotype were inoculated with every one of the strains to be evaluated in that year. Fixed terms for the analysis of environments 10–16 were g + s + y + gs + gy + sy, random terms gsy + ygr + ygrs.

Rennes

Evaluations from the environments 17–20 of Table 4 were analysed together. The combinations involved were the years 1991 and 1992 and the *F. culmorum* strains IPO 39–01 and Le Rheu 89–4. Within each year the strains were evaluated independently of each other on hill plots (20 seeds, 50 cm apart) in a randomized complete-block design with three replicates. That is, each block contained all of the genotypes within the particular *F. culmorum* strain. The model used for analysis was: fixed terms g + s + y + gs + gy + sy; random terms gsy + ysb.

In 1990 there were also evaluations of all genotypes except the Austrian ones inoculated by a Le Rheu mix. This is environment 21 in Table 4. The evaluations took place in a randomized complete blocks design with six replicates. The model fitted was simple: fixed g; random bg. (Random components with negative estimates for the

corresponding variance component were removed from models. In this case b was removed.)

Rennes was the only location where *F. nivale* was used for inoculation. In 1991 and 1992, environments 22 and 23 in Table 4, randomized complete-block designs were used in two and three replicates, respectively. The model included the fixed terms g + y + gy, and the random terms gyb.

Hohenheim

The design was comparable to that used in Vienna, except that field plots were 7.5 m^2, and two repeats were used within each genotype. The model used to analyse environments 24–31 of Table 4 was the same as to that used for Vienna.

Oberer Lindenhof

This was a location in 1990 and 1991. The design and model were the same as those of Hohenheim. The environments are 32–36 in Table 4.

Szeged

The design was the same as that of Vienna. Field plots were 5 m^2. Plots were subdivided in three repeats (Mesterhazy 1988). The model contained for the fixed terms g + s + y + gs + gy + sy + gsy and random term ygrs. The environments in question were 37–59 in Table 4.

Results

Table 5 shows that in general FHB ratings were highly correlated with yield reduction and weight reduction. Anthesis dates in Wageningen showed that flowering time did not influence infection levels.

The most important result from the analyses of the FHB ratings per location was that within each location significant genotype-by-strain-by-year interaction was present. (For Rennes, of course, no three way interaction could be determined for the evaluations with the Le Rheu mix and *F. nivale*.) Consequently, no straightforward determination of the type of resistance as strain specific or vertical seemed possible. However, the environmental scores per location on average were clustered more by year than by strain (not shown). This finding did not support an interpretation of the resistance as vertical. Because the pattern of the environmental scores that emerged from the overall analysis over the six locations was almost a superposition of the six analyses per location, the patterns per location will not be dealt with individually.

The overall analysis on the two-way table of BLUPs indexed by the 25 genotypes on one side and the 59 environments on the other side showed that the environmental main effect accounted for 81.9% of the variation in the table and the genotypic main effect for 5.0%. Tables 1 and 4 contain the genotypic and environmental means, after back transformation to percentages. One should be cautious with the interpretation of differences between the genotypic means, as the non-additivity, comprising 13.1% of the total variation, is considerable, and in fact precludes unconditional interpretation. For the structural part of the non-additivity five product terms, -together accounting for

Table 5 Pearson correlation coefficients based on means over replicates per location between Fusarium head blight infection level (FHB), yield reduction, weight reduction and visible kernel infection

	Yield reduction	Weight reduction	Kernel infection
FHB			
Wageningen	0.82	0.84	–
Vienna	0.77	0.73	–
Rennes *F. culmorum* 1990/1991	0.03	0.30	0.08
Rennes *F. culmorum* 1992	0.86	0.62	0.44
Rennes *F. nivale* 1991	–	0.71	–
Rennes *F. nivale* 1992	0.47	0.29	–
Hohenheim	0.51	0.79	–
Oberer Lindenhof	0.58	0.83	–
Szeged	0.78	–	0.71
kernel infection			
Rennes *F. culmorum* 1990/1991	0.91	0.80	
Rennes *F. culmorum* 1992	0.61	0.52	
Szeged	0.71	–	

77.8% of that non-additivity-, were judged to be relevant. These five product terms exceeded the critical value of 4.2%, and accounted for respectively 44.1, 11.5, 9.0, 6.7, and 6.5% of the interaction sum of squares.

In Fig. 2A–D the environmental scores corresponding to the two most explaining product terms are plotted. Fig. 2A just gives an overview of the 59 environments from Table 4. To facilitate closer inspection different plotting symbols are used: in Fig. 2B, year-by-location (Table 4), in Fig. 2C, *Fusarium* species, and in Fig. 2D, strains. Figure 2B shows that environments belonging to the same year-by-location combination tend to cluster; Fig. 2C illustrates that *Fusarium* species cannot be distinguished by their position; and Fig. 2D illustrates that environments cannot be distinguished on the basis of the strains. Special attention must be given to strain 1, IPO 39–01, which was used at all locations. The points for IPO 39–01 are scattered all over the plot. If there had been a case for strain specific resistance, the points for IPO 39–01 should have been close together. Therefore, Fig. 2 shows that year-by-location effects dominate *Fusarium* species and strain effects in the interaction with the genotypes, and supports the hypothesis of non-species specificity of the resistance. The plot for the two most explaining product terms may be seen as characteristic for all other possible plots that could have been made, like the third against the first term, the third against the second, etc. All plots more-or-less revealed the same pattern, that of dominating year-by-location effects.

More formally, this was also found in the analyses of variance on the environmental scores of the product terms one to five. First the amount of variation in the environmental scores of a particular term due to differences in year-by-location groups was calculated. Next, the amount of variation due to strains after correction for year-by-location effects was calculated. What was then left, represented variation due to the interaction between year-by-location effects and strains. For environmental scores corresponding to the first product term 85.7% was due to year-by-location effects, 6.7 to strain effects after correction for year-

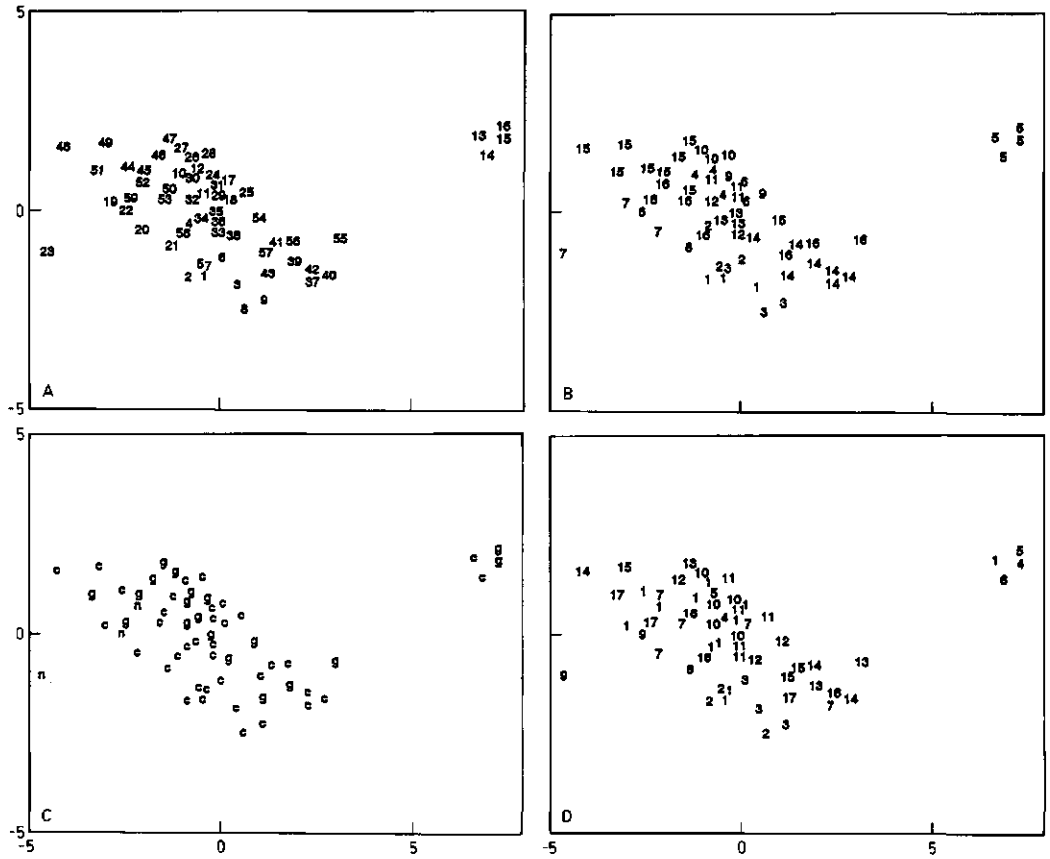


Fig. 2 Plots of environmental scores indexed by **A** environmental codes of Table 4, **B** year by location codes of Table 4, **C** *Fusarium* species (*c culmorum, g graminearum, n nivale*), **D** strain numbers of Table 2

by-location effects, and then 7.6% was left for the interaction between both grouping factors. The same quantities for environmental scores of the second product term were; 89.9, 4.8, and 5.2; for the third; 77.2, 9.4, and 13.4; for the fourth; 77.1, 12.4 and 10.5; and for the fifth; 72.1, 9.8 and 18.1. It is clear that after correction for year by location effects, strain effects contribute very little to the variation in environmental scores. Interaction between genotypes and environments can be described to a major extent as interaction between genotypes and the environmental circumstances due to the combination of year and location. Consequently, there seems no reason to assume species specific resistance.

A last argument for that thesis is provided by the discriminant analyses using either a year-by-location grouping factor or a strain grouping factor and using as variables the environmental scores belonging to the first five product terms of the interaction. For the year by location case only 6 out of the 59 environments were not allocated to the year by location group they came from. In contrast, using the strains as grouping factor only 19 environments were correctly allocated; 40 were wrongly allocated. This once again shows the strong prevalence of year-by-location effects over strain effects.

Discussion

Inoculation method

The data do not allow a firm conclusion to be made about which of both inoculation methods is preferable. Still, a major interaction occurred in Vienna in 1992 due to a change in the weather during inoculation. An extremely hot and dry period with low infection pressure was succeeded by a humid and cool period with higher infection pressure. However, even for these extreme environmental circumstances no differentiation between the strains could be observed with respect to their interaction with the wheat

genotypes. Within all locations the same dominance of year-by-location effects occurred, irrespective of the inoculation method used. Nevertheless, it is recommended to avoid unnecessary interactions of whichever type. The inoculation method giving rise to the smallest interactions should therefore be used.

FHB rating as selection criterion

The size of the correlations of FHB with yield reduction and percentage *Fusarium*-infected kernels showed that the FHB rating is a reliable selection criterion for resistance. This was concluded earlier by Mesterhazy (1990), Snijders (1990), and Miedaner et al. (1993).

Specificity of FHB resistance

There seems to be no reason for believing that the resistance to Fusarium head blight as caused by *F. culmorum* is specific. The same is true for *F. graminearum*. Neither is there any indication for a geographical pattern in virulence genes. Furthermore, the resistance to *F. graminearum* and *F. nivale* seems to be of the same type as that to *F. culmorum*. Any reasonable aggressive strain, a *F. culmorum* strain for the cool climates, a *F. graminearum* strain for the warmer humid areas, should be satisfactory for screening purposes. This confirms the results of Snijders and van Eeuwijk (1991) for *F. culmorum* strains from The Netherlands and the results of Mesterhazy (1983, 1988) for Hungary. Shuttle programmes for selection for *Fusarium* head blight resistance are unnecessary.

The dependence of aggressiveness of strains on the environmental circumstances, which to a large extent are unpredictable, complicates the choice of strain. Screening programmes can be safeguarded by the inclusion of a number of strains, whether pure isolates or mixtures, having varying sensitivities to the environment (Lemmens et al. 1993, Mesterhazy (1984, 1987), Snijders and van Eeuwijk 1991).

References

Arseniuk E, Goral T, Czembor HJ (1993) Reaction of triticale, wheat and rye accessions to graminaceous *Fusarium* spp. infection at the seedling and adult plant growth stages. Euphytica 70:175–183

Daamen R, Langerak CJ, Stol W (1991) Surveys of diseases and pests in the Netherlands. 3. *Monographella nivalis* and *Fusarium* spp. in winter wheat fields and seed lots. Neth J Plant Pathol 97:105–114

Denis J-B (1991) Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. Rev Stat Appl XXXIX:5–24

Gabriel KR (1978) Least squares approximation of matrices by additive and multiplicative models. J R Stat Soc B 40:186–196

Gauch HG (1988) Model selection and validation for yield trials with interaction. Biometrics 44:705–715

Genstat 5 Committee (1993) Genstat 5 release 3 reference manual. Clarendon Press, Oxford

Jolliffe IT (1986) Principal component analysis. Springer, Berlin Heidelberg, New York

Kempton RA (1984) The use of biplots in interpreting variety by environment interactions. J Agric Sci 103:123–135

Lemmens M, Bürstmayr H, Ruckenbauer P (1993) Variation in *Fusarium* head blight susceptibility of international and Austrian wheat breeding material. Bodenkultur 44:65–78

Marasas WFO, Nelson PE, Toussoun TA (1984) Toxigenic *Fusarium* species, identity and mycotoxicology. The Pennsylvania State University Press, University Park and London

Mesterhazy A (1977) Reaction of winter wheat varieties to four *Fusarium* species. J Phytopathol 90:104–112

Mesterhazy A (1978) Comparative analysis of artificial inoculation methods with *Fusarium* spp. on winter wheat varieties. Z Phytopathol Z 93:12–25

Mesterhazy A (1983) Breeding wheat for resistance to *Fusarium graminearum* and *Fusarium culmorum*. Plant Breeding 91:295–311

Mesterhazy A (1984) A laboratory method to predict pathogenicity of *Fusarium graminearum* in field and resistance of wheat to scab. Acta Phytopathol Acad Sci Hung 19:205–218

Mesterhazy A (1987) Selection of head blight resistant wheats through improved seedling resistance. Plant Breed 98:25–36

Mesterhazy A (1988) Expression of resistance of wheat to *Fusarium graminearum* and *F. culmorum* under various experimental conditions. J Phytopathol 123:304–310

Mesterhazy A (1990) Über die Möglichkeiten und Praktik der Züchtung widerstandsfähiger Weizensorten gegen Ährenfusariose. Bericht über 41. Arbeitstagung der Vereinigten Österreichischen Pflanzenzüchter, Gumpenstein, pp 253–234

Miedaner T, Borchardt DC, Geiger HH (1993) Genetic analysis of inbred lines and their crosses for resistance to head blight (*Fusarium culmorum*, *F. graminearum*) in winter rye. Euphytica 65:123–133

Nirenberg HI (1981) A simplified method for identifying *Fusarium* spp. occurring on wheat. Can J Bot 59:1599–1609

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554

Robinson GK (1991) That BLUP is a good thing: the estimation of random effects. Stat Sci 6:15–51

Saur L (1991) Recherche de géniteurs de résistance à la fusariose de l'épi causée par *Fusarium culmorum* chez le blé et les espèces voisines. Agronomie 11:535–541

Saur L, Trottet M (1992) Héritabilité de la résistance à la fusariose de l'épi et sélection récurrente dans une population de blé tendre. Agronomie 12:297–302

Searle SR (1971) Linear models. Wiley, New York

Snijders CHA (1990) Genetic variation for resistance to *Fusarium* head blight in bread wheat. Euphytica 50:171–179

Snijders CHA (1994) Breeding for resistance to *Fusarium* in wheat and maize. In: Miller JD, Trenholm HL (eds), Mycotoxins in grain: compounds other than aflatoxin. Eagan Press, St. Paul, Minnesota, USA, pp 37–58

Snijders CHA, Eeuwijk FA van (1991) Genotype×strain interactions for resistance to *Fusarium* head blight caused by *Fusarium culmorum* winter wheat. Theor Appl Genet 81:239–244

Snijders CHA, Krechting CF (1992) Inhibition of deoxynivalenol translocation and fungal colonization in *Fusarium* head blight resistant wheat. Can J Bot 70:1570–1576

Stack RW, McMullen MP (1985) Head blighting potential of *Fusarium* species associated with spring wheat heads. Can J Plant Path 7:79–82

van Eeuwijk FA (1995) Multiplicative interaction in generalized linear models. Biometrics (in press)

Verdooren LR (1992) Practical aspects of the estimation of varietal contrasts from variety trials. Biuletyn oceny odmian 24–25:105–137

Wilcoxson RD, Kommedahl T, Ozmon EA, Windels WE (1988) Occurrence of *Fusarium* species in scabby wheat from Minnesota and their pathogenicity to wheat. Phytopathology 78:586–589

Zadoks JC, Rijswijk FH (1984) Agro-ecological atlas of cereal growing in Europe, vol.3: atlas of cereal diseases and pests in Europe. PUDOC, Wageningen

Chapter XIV

Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables

F.A. van Eeuwijk[1], *J.-B Denis*[2] & *M.S. Kang*[3]

[1] *DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)*
P.O.Box 16, 6700 AA Wageningen, The Netherlands

[2] *Laboratoire de Biométrie, I.N.R.A.*
route de Saint-Cyr, F-78026 Versailles, France

[3] *Department of Agronomy, 104 Madison Sturgis Hall, Louisiana State University,*
Baton Rouge, Louisiana 70803-2110, United States of America

Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables

I. INTRODUCTION

For genotypes and environments making up the factor levels in two-way genotype by environment tables, often substantial additional information is either available or easily obtainable. For genotypes, additional quantitative information may be present from laboratory and greenhouse tests bearing on the physiology of the plants, while additional qualitative information may be present from various categorizations, like those on basis of genealogy. For environments, quantitative information can consist in edaphic and climatological data, whereas a minimum in qualitative information consists in year and location groupings. The additional information on genotypes and environments includes more than direct measurements. More remotely, statistics calculated from previous or comparable trials, concerning the variable under study as well as other variables, may be used.

For the additional information, more or less clear-cut hypotheses may be entertained regarding its relation with the structure of the genotype by environment interaction in the variable to be analyzed. To test these hypotheses statistically, models are necessary that allow the incorporation of that information. In plant breeding, the emphasis has long been on models not offering this opportunity. We feel that for a broad spectrum of hypotheses, between exploratory and inferential, it is imperative to pay more attention to regression based statistical methods. As a consequence, more parsimonious models may be built, providing more accurate tools to decide and act on. Similar ideas have been expressed by Hinkelmann (1974), Denis and Vincourt (1982), Tai (1990), van Eeuwijk (1993), and Federer and Scully (1993).

The main point of this paper is to give a survey of the most important regression based models for two-way tables and to illustrate the interpretation of their interaction parameters. Three families of models will be presented. After some details on notation (Section II), fixed factorial regression is introduced first (Section III), and an account is given of how quantitative as well as qualitative covariates may be included. Secondly, reduced rank factorial regression, based on bilinear descriptions of the interaction, will be dealt with (Section IV). Lastly, mixed factorial regression is presented, in which either the genotypes or the environments are supposed to represent a random sample from a population (Section V). Estimation and testing, together with software availability are briefly discussed in Section VI. Considerations playing a role in model choice form the subject of Section VII. Finally, Section VIII presents alternatives and further extensions to the models presented in the sections III, IV and V.

II. DATA AND NOTATION

A. THE DATA TABLE TO BE INTERPRETED

The topic of the paper will be restricted to the interpretation of the joint effects of the factors 'genotype' and 'environment' on a continuous variable Y. The subscript i $(1,...,I)$ will be used to indicate the genotype, and j $(1,...,J)$ to indicate the environment. Typically, Y_{ij} represents yield, but many other quantitative variables are equally substitutable. Averaging

Table 1
Hypothetical yields for 17 genotypes (rows) in 12 environments (columns).

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
G1	10	6	6	7	9	7	2	7	8	6	5	11
G2	9	6	5	7	8	8	2	6	7	7	5	9
G3	11	7	6	7	9	6	4	6	7	7	6	8
G4	8	7	4	6	8	8	5	7	7	9	5	7
G5	9	7	6	7	9	6	3	6	8	9	6	9
G6	8	8	4	7	8	7	3	8	7	8	6	8
G7	11	8	5	7	9	7	4	8	9	9	7	9
G8	10	8	5	8	11	9	4	9	9	11	8	10
G9	8	8	6	7	9	7	3	7	9	9	7	9
G10	8	6	6	7	9	6	3	7	6	10	6	10
G11	8	7	4	5	9	7	3	5	7	9	4	9
G12	8	6	4	6	8	7	3	8	6	8	5	9
G13	8	8	4	6	8	9	4	7	7	8	5	8
G14	9	6	6	6	10	7	3	7	8	12	5	9
G15	9	7	5	9	11	9	5	8	6	10	4	8
G16	7	7	4	6	8	6	4	6	7	8	6	7
G17	8	6	3	7	9	5	3	6	6	8	4	8

over replicates makes Y_{ij} to comply closer with distributional assumptions.

For a number of statistical tests with regard to the structure of interaction an estimate of error is required. This estimate may be obtained from the mean intra-block error, or from the part of the interaction not modeled. We will not consider the question of which estimate to use, but assume that a non-controversial estimate of error is available.

B. ADDITIONAL INFORMATION.

Besides the data, Y_{ij}, additional information is assumed to be present at the levels of the genotypes and/or environments. The value of the k-th covariate ($k = 1,...,K$) for the i-th genotype will be denoted by x_{ik}. Covariates are either quantitative, as in multiple regression, or qualitative, as in analysis of variance. Examples of quantitative genotypic covariates are physiological characterizations, such as earliness, and disease susceptibilities. Examples of qualitative genotypic covariates are genetic and geographic origin. When only one covariate is considered, the subscript in question is dropped, *i.e.* x_i instead of x_{i1}. Quantitative covariates are throughout supposed to be centered.

Similarly, we will denote the h-th covariate ($h = 1,...,H$) for the j-th environment by z_{jh}. For environments, we can think of humidity and soil pH as quantitative covariates, and location (region, country) and cultivation regimes as qualitative covariates.

To indicate that the covariate values are considered known, they are written with lower case letters. This does not mean that the corresponding factor may not be random, but only that the analyses are done conditionally on the values of the covariates.

A popular environmental covariate is $y_{.j}$, the mean over genotypes (Finlay and Wilkinson, 1963). $y_{.j}$ can be treated as any other covariate (Mandel, 1961), because of its statistical independence of the interaction estimates. Also the genotypic main effect, $y_{i.}$ can be used in

that way. A further possibility is to use the product of y_j and y_l, as a covariate for the entire table. A simple extension uses covariates of the type o_l and o_j, *i.e.* main effects of another variable on the same set of genotypes and environments. For even further extensions, see Baril (1992). Though mostly covariates are used in a linear fashion only, higher order terms as squares, cubes, and cross-products can be considered equally well.

In some models, pseudo covariates are estimated. When they are defined as linear combinations of measured covariates, they will be designated as synthetic covariates. When they are only subject to statistical/numerical construction rules, they will be called artificial covariates.

C. DESCRIPTION OF THE MODELS.

The models presented below consist of sums of model terms. Terms are related to the expectation or the variance. Fixed parameters are represented by lower case Greek letters, random parameters and variates by standard upper case letters, and observations on (co)variates by standard lower case letters. The error term is written E_{ij}. Unless stated otherwise, E_{ij}'s are assumed to have zero mean, constant variance, and to be uncorrelated. For fixed effects models, the decomposition of the degrees of freedom (parametric dimension) corresponding to the different model terms is displayed via recapitulative tables. These tables are two-dimensional depictions showing the composition of the models, in which each model term corresponds with a zone in the table, and where the area of this zone is proportional to the associated degrees of freedom (Denis, 1991).

Table 2

Genotypic covariates. Rstc represents resistance, a centred measurement. The covariates $\rho_{(2)}$, $\rho_{(12)}$ and $\xi_{(16)}$ are artificial covariates corresponding to models (2), (12) and (16).

	Rstc	$\rho_{(2)}$	$\rho_{(12)}$	$\xi_{(16)}$
G1	-41	-0.0082	0.5525	0.5820
G2	-13	-0.0038	0.2614	0.2320
G3	-37	-0.0053	0.3335	0.3073
G4	34	0.0041	-0.2899	-0.3620
G5	-5	-0.0023	0.1635	0.1991
G6	25	0.0002	-0.0297	-0.0792
G7	-32	-0.0025	0.1780	0.1522
G8	-7	0.0017	-0.1037	-0.1038
G9	-24	-0.0024	0.1538	0.1169
G10	2	0.0014	-0.0795	0.0564
G11	16	0.0010	-0.0347	-0.0748
G12	8	0.0011	-0.0737	-0.0438
G13	-13	0.0012	-0.0900	-0.2335
G14	11	0.0043	-0.2351	-0.1182
G15	39	0.0069	-0.4850	-0.4306
G16	14	0.0008	-0.0840	-0.1441
G17	23	0.0019	-0.1373	-0.0558

D. EXAMPLE DATA SET.

To enhance understanding, for some of the presented models a numerical example will be used. The data set used is a modified and rounded version of the data used by Kang and Gorman (1989), including yield figures for 17 genotypes in 12 environments (Table 1). For each genotype, an associated fictitious resistance measure, Rstc, is available (Table 2). The 12 environments are characterized by four climatological variables (Table 3). An independent estimate for the error variance is also present.

Table 3

Environmental covariates. Maximum temperature (MaxT), minimum temperature (MinT), Rain, and relative humidity (RH) are measurements that were averaged over the growing season. The covariates $\tau_{(3)}$ and $\zeta_{(16)}$ are artificial covariates corresponding to models (3) and (16).

	MaxT	MinT	Rain	RH	$\tau_{(3)}$	$\zeta_{(16)}$
E1	5	-3	-119	-63	-0.0228	0.3628
E2	11	7	-39	-10	0.0028	-0.1212
E3	15	1	-121	61	-0.0164	0.3370
E4	16	12	41	65	0.0055	0.0145
E5	-2	3	21	20	0.0092	-0.0839
E6	-10	-10	41	18	0.0112	-0.3009
E7	0	14	111	13	0.0202	-0.3695
E8	-8	1	91	4	0.0083	-0.1516
E9	-7	-5	-99	-37	-0.0174	0.2259
E10	-18	-18	291	-6	0.0308	-0.4894
E11	13	13	-119	-24	-0.0152	0.1739
E12	-15	-15	-99	-41	-0.0160	0.4025

III. FIXED FACTORIAL REGRESSION

A general formal treatment of the models presented in this section is given by Denis (1980, 1988).

A. THE ADDITIVE MODEL AS BASE LINE

It is common to define interaction in two-way tables relative to the two-way additive model,

$$Y_{ij}=\mu+\alpha_i+\beta_j+E_{ij}. \qquad (1)$$

The additive model provides a first, rough approximation to the data. Analysis of variance (ANOVA) on our example data showed that 82% of the total sum of squares could be explained by only 13% of the degrees of freedom (Table 4). Nevertheless, interaction was highly significant and could not be omitted.

For many purposes it is useful to express (1) as a double regression model with the constant covariates $1_i=1_j=1$:

$$Y_{ij}=\mu+\alpha_i 1_j+1_i\beta_j+E_{ij}. \qquad (1')$$

The main effects (α_i,β_j) thus are the regression coefficients for these non-informative constant covariates. The structure of the model is displayed in Table 5.

Table 4

Two-way ANOVA, and interaction as described by various models. Numbers for models correspond to numbers for model formulations in text.

Model	Source	Df	SS	MS
	Genotypic main effect	16	67.3	4.208
	Environmental main effect	11	550.6	50.053
	Interaction (non-additivity)	176	136.7	0.853
(6)	Rstc.v.MaxT	1	4.2	4.185
(6)	Rstc.v.MinT	1	0.05	0.049
(6)	Rstc.v.Rain	1	28.0	27.954
(6)	Rstc.v.RH	1	4.9	4.913
(8)	Rstc.($\Sigma v_{ih}.z_{jh}$)	4	29.2	7.294
(3)	Rstc.τ_j	11	31.0	2.816
(2)	ρ_i.Rain	16	38.3	2.395
(12)	ρ_i.($\Sigma\lambda_h.z_{jh}$)	19	38.8	2.042
(16)	$\xi_i\zeta_j$	26	42.6	1.640
	Error	561	252.4	0.450

B. INCLUDING ONE QUANTITATIVE ENVIRONMENTAL COVARIATE.

Perhaps the simplest way of introducing a covariate associated with the environments, is to write the interaction as a regression on this covariate with the coefficient depending on the genotype. Early applications of this type of model include Knight (1970), and Freeman and Perkins (1971). More recent applications are Fakorede and Opeke (1986), and McGraw *et al.* (1986). The model can be written as

$$Y_{ij} = \mu + \alpha_i + \beta_j + \rho_i z_j + E_{ij}. \qquad (2)$$

For illustration we choose the rainfall data from Table 3 as environmental covariate. The estimates for the genotypic regression coefficients, ρ_i, are given in Table 2. These coefficients can be interpreted as underlying a differential genotypic response to rainfall. For example, for *G15* yield increases with rainfall relative to what might have been expected on the basis of an additive model. Under dry circumstances, this yield decreases. Recall that the covariates were all centered.

The partitioning of the interaction sum of squares according to model (2) is given in Table 4. The sum of squares due to heterogeneity of genotypic slopes amounted to 38.3, with 16 degrees of freedom. The corresponding mean square, 2.395, is clearly greater than the mean square for the total interaction, 0.853. Therefore, rainfall can be considered to be a good explanatory covariate.

Table 6 reveals the structure of the model by showing its recapitulative table. Because the environmental main effect was fitted before the regression on z_j, the $\rho_i z_j$ term of the interaction corresponds to I-1 degrees of freedom for I parameters.

Table 5 Recapitulative table associated with model (1).

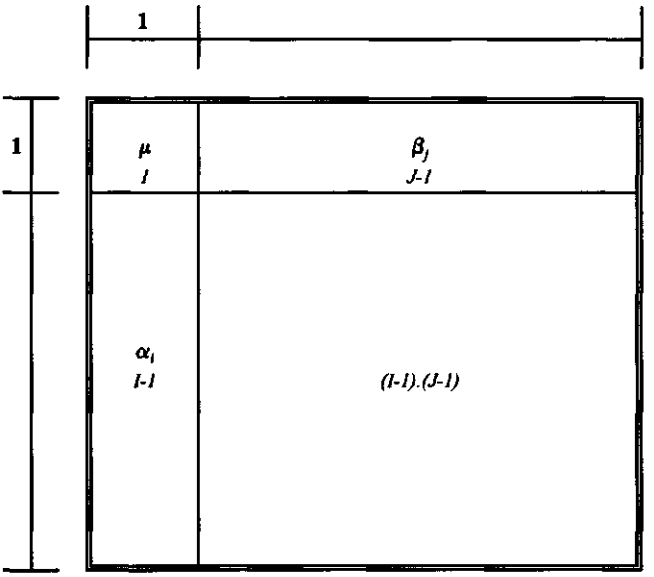
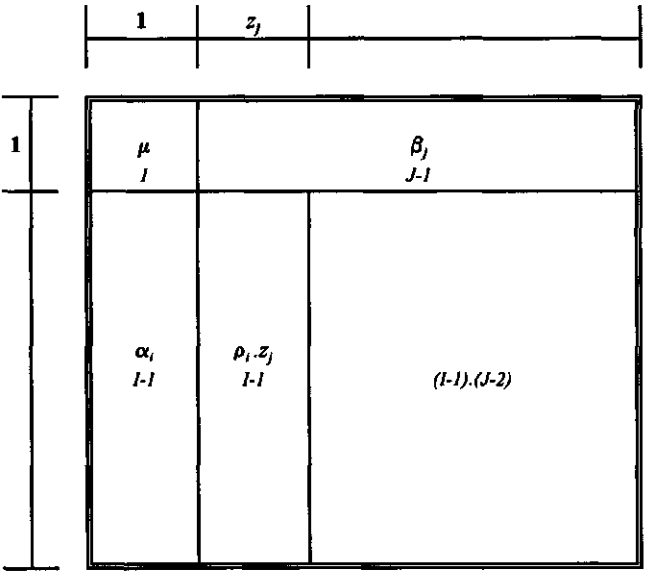


Table 6 Recapitulative table associated with model (2).



C. INCLUDING ONE QUANTITATIVE GENOTYPIC COVARIATE.

The counterpart of model (2), including one genotypic covariate is

$$Y_{ij}=\mu+\alpha_i+\beta_j+x_i\tau_j+E_{ij}. \quad (3)$$

An untimely use of this model can be found in Freeman and Crisp (1979). For our example, x_i is a resistance measure for genotype i, as given in Table 2. Now, τ_j can be interpreted as the potential of environment j to favor the spread of the disease. If the environment is beneficial to the spread of the disease, *i.e.* τ_j is large and positive, and if the genotype is susceptible, *i.e.* x_i is large and negative, then the correction term $x_i\tau_j$ will be large and negative, implying a decrease in yield. Estimates for τ_j are given in Table 3, and the explained sum of squares is given in Table 4. The mean square amounted to 2.816, again much greater than the total interaction mean square. The number of degrees of freedom attributed to a term is not determined by the factor with which the covariate is associated, but by the opposite factor (Table 7).

D. INCLUDING SEVERAL QUANTITATIVE ENVIRONMENTAL COVARIATES

A generalization of (2), including two environmental covariates leads to

$$Y_{ij}=\mu+\alpha_i+\beta_j+\rho_{i1}z_{j1}+\rho_{i2}z_{j2}+E_{ij}, \quad (4)$$

where z_{j1} and z_{j2} can be rainfall and average maximum temperature over the growing season in environment j. The structure of the model is given in Table 8. When covariates are correlated, inclusion of more than one covariate complicates interpretation of the coefficients, just as for multiple regression. Coefficients are conditional upon the values of the other included covariates, so one should be cautious in interpretations. Examples of the application of model (4) can be found in Hardwick (1972), Hardwick and Wood (1972), Rameau and Denis (1992), and van Eeuwijk and Elgersma (1993).

E. INCLUDING ONE QUALITATIVE GENOTYPIC COVARIATE

Qualitative genotypic covariates attribute group membership to genotypes. Let x_i be a qualitative variable that indicates to which of three groups with a common ancestor a genotype belongs. For example, x_i=3 would mean that genotype i belongs to the third group of genotypes. Including this variable x_i in model (3) does not result in anything sensible, because the numbering of the groups is arbitrary and does not refer to something inherent to that group of genotypes. What we can do is replace the qualitative variable x_i by indicator variables (valued 0 or 1), just as when ANOVA models are presented in multiple regression form. Now x_{i1}, x_{i2}, and x_{i3} attribute membership when they have value 1, *e.g.* x_{i3}=1 means genotype i belongs to group 3. Of course, if x_{i3}=1, then x_{i1}=x_{i2}=0, therefore $x_{i1}+x_{i2}+x_{i3}$ is always one. This redundancy can be removed by leaving out one of the indicator variables, or imposing an additional constraint. Another possibility is to remove the environmental main effect, as is done in

$$Y_{ij}=\mu+\alpha_i+x_{i1}\tau_{j1}+x_{i2}\tau_{j2}+x_{i3}\tau_{j3}+E_{ij}. \quad (5)$$

The parameters τ_{jk} represent the environmental 'main' effects for each of the three groups of genotypes separately. Table 9 displays the structure of the model.

F. INCLUDING GENOTYPIC AND ENVIRONMENTAL COVARIATES

1. Quantitative-quantitative

The simplest extension of the additive model including one genotypic and one

Table 7 Recapitulative table associated with model (3).

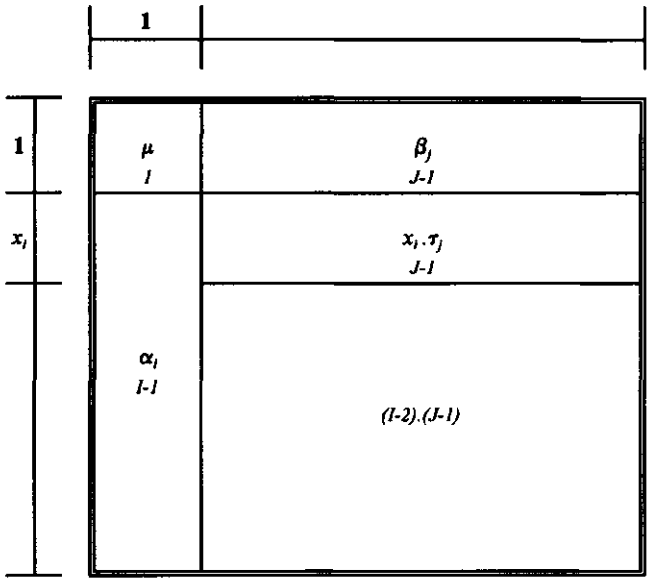


Table 8 Recapitulative table associated with model (4).

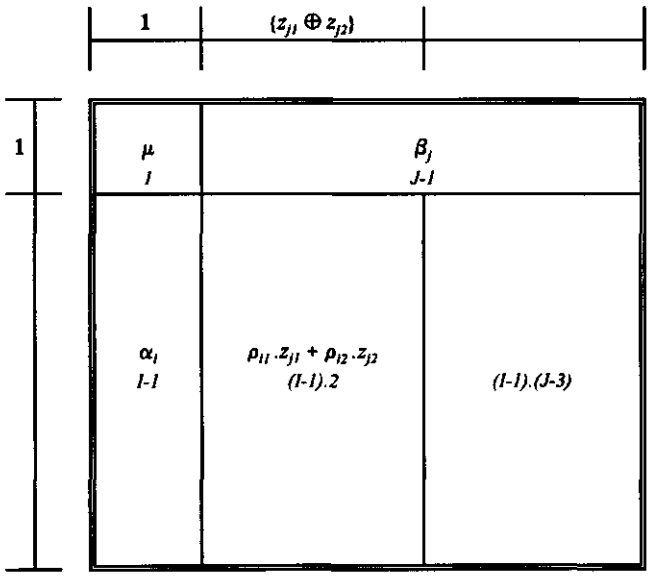


Table 9 Recapitulative table associated with model (5).

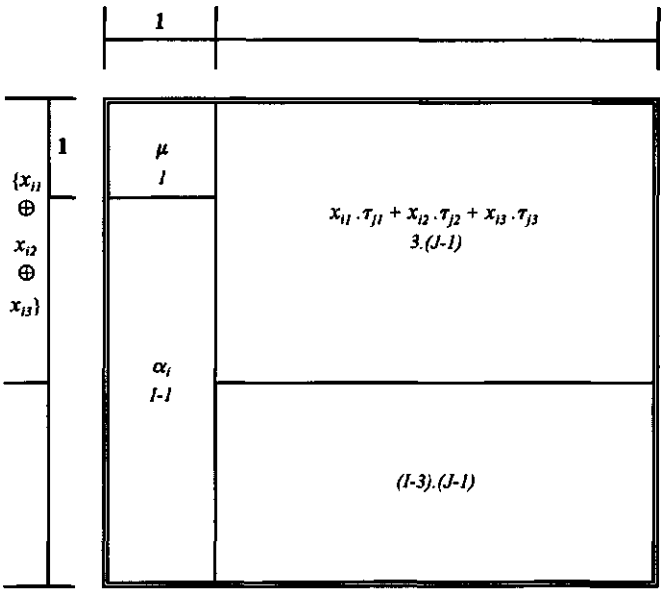


Table 10 Recapitulative table associated with model (6).

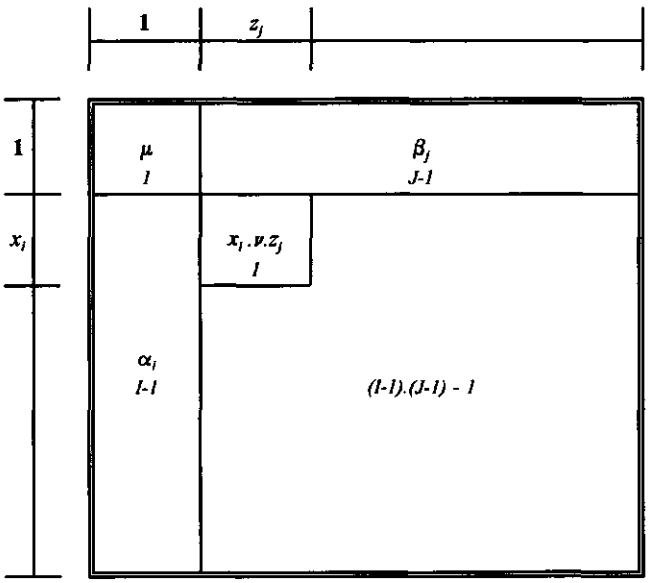
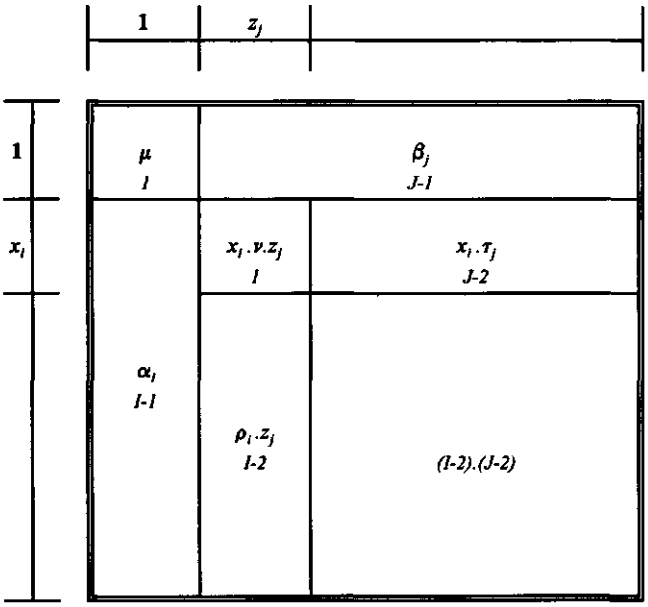


Table 11
Recapitulative table associated with model (7).



environmental covariate is

$$Y_{ij}=\mu+\alpha_i+\beta_j+x_i\nu\, z_j+E_{ij}. \tag{6}$$

It can be derived from models (2) or (3) by imposing the restriction of $\rho_i=x_i\nu$, or $\tau_j=\nu z_j$, respectively. In Table 10 it is shown how the single parameter ν represents one degree of freedom in the interaction space.

The model has been fitted for all combinations of genotypic and environmental covariates at our disposal (Table 4). The combination of genotypic resistance and environmental rainfall produced the highest mean square, as might have been expected from the previous results.

One step further than model (6), a simple combination of (2), (3), and (6) gives

$$Y_{ij}=\mu+\alpha_i+\beta_j+x_i\nu\, z_j+x_i\tau_j+\rho_i z_j+E_{ij}. \tag{7}$$

The recapitulative table (Table 11) for this model shows how $x_i\nu z_j$ is common to both $x_i\tau_j$ and $\rho_i z_j$. To estimate ν, supplementary constraints have to be imposed on τ_j and ρ_i. The ANOVA table (Table 12) shows that a significant amount of interaction was left unexplained by model (7). For a more telling example, see Paul *et al.* (1993).

Model (6) can straightforwardly be extended to include several genotypic as well as environmental covariates, to give

$$Y_{ij}=\mu+\alpha_i+\beta_j+\sum_{k=1}^{K}\sum_{h=1}^{H}x_{ik}\nu_{kh}z_{jh}+E_{ij}. \tag{8}$$

Good illustrations of applications of model (8) are given by Charmet *et al.* (1993), and Baril *et al.* (1995).

2. Qualitative-quantitative

Taking x_i qualitative in model (6) leads to the model

$$Y_{ij}=\mu+\alpha_i+\beta_j^*+x_{i1}\nu_1 z_j+x_{i2}\nu_2 z_j+x_{i3}\nu_3 z_j+E_{ij}, \tag{9}$$

and Table 13. The parameters β_j^* represent the environmental main effects after adjustment for the general mean and the regressor z_j. The β_j^*'s may be interpreted as a type of residuals. For applications see Saeed and Francis (1984), and Royo *et al.* (1993).

3. Qualitative-qualitative

When in model (9) both the genotypic covariate and the environmental covariate are qualitative, we arrive at

$$\begin{aligned} Y_{ij}=\alpha_i^*+\beta_j^*+x_{i1}\nu_{11}z_{j1}+x_{i2}\nu_{21}z_{j1}+x_{i3}\nu_{31}z_{j1}+ \\ x_{i1}\nu_{12}z_{j2}+x_{i2}\nu_{22}z_{j2}+x_{i3}\nu_{32}z_{j2}+E_{ij}, \end{aligned} \tag{10}$$

and Table 14. The environmental covariate z_j may indicate one of two regions, and is represented in the model by two indicator variables, z_{j1} and z_{j2}. In addition to the β_j^*'s of model (9), α_i^*'s appear, representing the genotypic main effects after adjustment for cross-product terms involving x_i. The parameters ν_{kh} represent the mean for the genotypes of the genotypic group k (descendance) in the environments of the environmental group h (region).

Model (10) can be reparametrized giving

$$Y_{ij}=\mu+\alpha_i+\beta_j+\nu^*_{[x_i],[z_j]}+E_{ij}, \tag{10'}$$

and Table 15. In (10') the usual main effects are included, and the $\nu^*_{[x_i],[z_j]}$'s have to sum to zero over genotypes (sum over i) and environments (sum over j). Being adjusted for the main effects, they are interaction parameters. Interaction is exclusively of the 'between by between' type. One might think of classifying the original data in the six groups following from the

Table 12

Two-way ANOVA with decomposition of the interaction according to model (7), with Rstc for x_i and Rain for z_j.

Source	Df	SS	MS
Genotype	16	67.3	4.208
Environment	11	550.6	50.053
Interaction	176	136.7	0.853
$\text{Rstc}_i.\nu.\text{Rain}_j$	1	28.0	27.954
$\text{Rstc}_i.\tau_j$	10	3.0	0.301
$\rho_i.\text{Rain}_j$	15	10.4	0.691
Remainder	150	108.7	0.725
Error	561	252.4	0.450

intersection of the three genotypic groups with the two environmental groups. Interaction is present only between these six groups.

Many authors have studied models of the type exemplified by (10'). Although the use of a priori groupings is inferentially superior over the use of a posteriori groupings, most references relate to the latter (Horner and Frey, 1957; Abou-el-Fittouh *et al.*, 1969; Lin and Thompson, 1975; Byth *et al.*, 1976; Denis, 1979; Seif *et al.*, 1979; Berbigier *et al.*, 1980; Brennan *et al.*, 1981; Brown *et al.*, 1983; Lefkovitch, 1985; Lin and Butler, 1988; Corsten and Denis, 1990; Crossa *et al.*, 1990; Arntzen and van Eeuwijk, 1992; Muir *et al.*, 1992; Oliveira and Charmet, 1992). With a priori grouping, the procedure is fully inferential, otherwise it is more exploratory. The inferential value when using a posteriori groupings remains a point of discussion. Certainly, the type of testing needs more consideration in these cases.

IV. REDUCED RANK FACTORIAL REGRESSION

Theory on general reduced rank regression models has been developed over time by a number of authors belonging to very different disciplines. Among the major contributions we list Rao (1964), Izenman (1975), van den Wollenberg (1977), Gabriel (1978), Obadia (1978), Tso (1981), Davies and Tso (1982), Sabatier *et al.* (1989), van der Leeden (1990), and Velu (1991). As a solution to genotype by environment interaction problems in plant breeding, reduced rank *factorial* regression models have been proposed. Important contributions are due to Wood (1976), Denis (1991), van Eeuwijk (1992a), and van Eeuwijk (1995a).

A. ONE-WAY REDUCED RANK REGRESSION WITH ONE TERM

Considering model (4) and Table 8, we see that up to J-1 covariates are conceivable. For the case of J-1 covariates, the interaction described would be equal to the total non-additivity remaining from the additive two-way model (1). A number of covariates would then very likely be modeling mere noise, as in most situations with large numbers of covariates. A method allowing the incorporation of substantial amounts of covariates, while using fewer degrees of freedom than a comparable factorial regression model, is reduced rank (factorial) regression. Basically, a so-called synthetic covariate is formed as a linear combination of the available covariates, *i.e.* the most explanatory linear combination that can be constructed according to a least squares criterion. A synthetic covariate can be incorporated in a model like (2) without further complications. Define the synthetic covariate

$$\zeta_j = \sum_{h=1}^{H} \lambda_h z_{jh}. \tag{11}$$

The coefficients λ_h are unknown parameters to be estimated from the data.
The model becomes:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \rho_i \left[\sum_{h=1}^{H} \lambda_h z_{jh} \right] + E_{ij}. \tag{12}$$

Table 16 shows the distribution of the degrees of freedom over the various terms. It is obvious that substantial amounts of degrees of freedom can be won. As an illustration, compare Table 16 with Table 8. Table 8 gives the degrees of freedom for the interaction in a factorial regression model with 2 environmental covariates (H=2), 2(I-1). The comparable reduced rank regression model (12) uses I degrees of freedom. In general, the difference between a reduced rank model as (12) and the corresponding full rank model amounts to (I-2)(H-1) degrees of

Table 13 Recapitulative table associated with model (9).

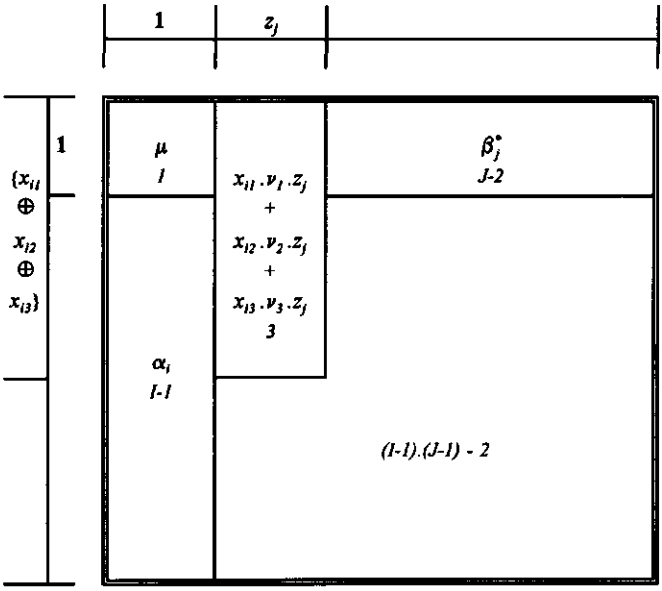


Table 14 Recapitulative table associated with model (10).

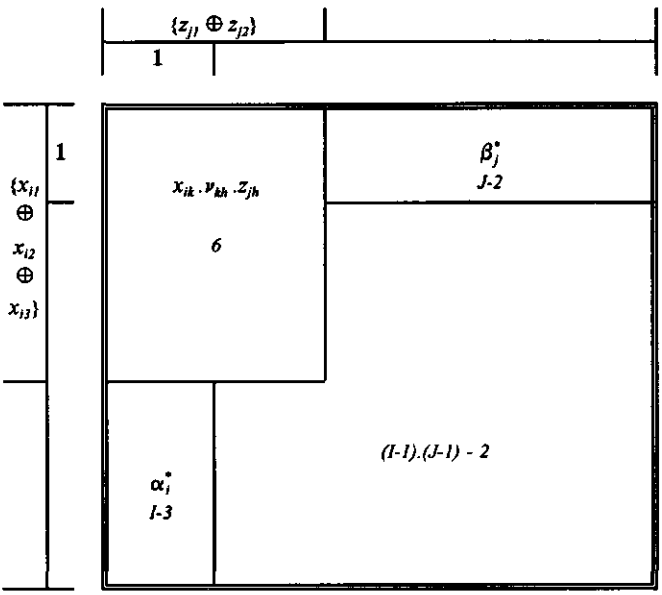


Table 15 Recapitulative table associated with model (10').

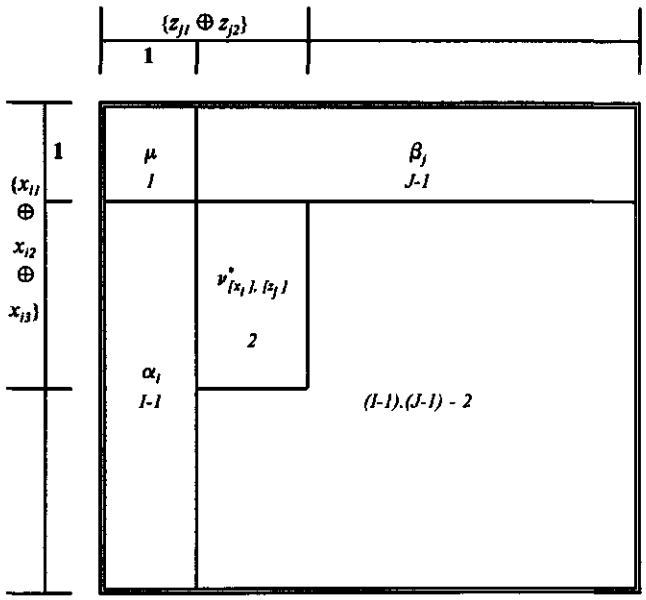
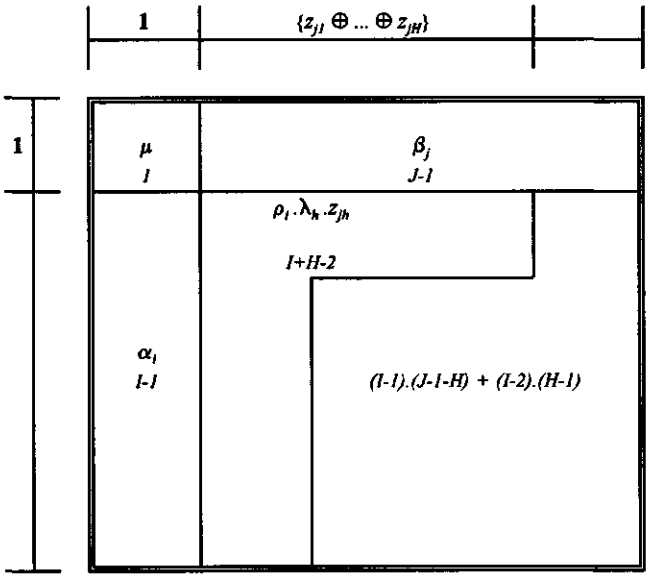


Table 16 Recapitulative table associated with model (12).



freedom. The difference increases with I and H. This increase in parsimony can express itself in greater accuracy and stability. For interpretational purposes, one should try to integrate the synthetic covariate in subject matter knowledge about the environments. For the genotypic sensitivities, ρ_i, a physiological basis should be sought.

Model (12) is not linear in its parameters, but bilinear. Least squares estimates are no longer linear combinations of the observations, but can come from a singular value decomposition of the fitted values matrix of the factorial regression model including the same set of covariates.

Wood (1976) contains an example of a reduced rank factorial regression model with one synthethic covariate.

B. ONE-WAY REDUCED RANK REGRESSION WITH SEVERAL TERMS

There is no need to restrict the number of synthetic covariates in reduced rank models to just one,

$$Y_{ij}=\mu+\alpha_i+\beta_j+\sum_{r=1}^{R}\rho_{ir}\left[\sum_{h=1}^{H}\lambda_{hr}z_{jh}\right]+E_{ij}. \tag{13}$$

Model (12) follows from (13) by taking R=1. Table 17 represents the recapitulative table.

Illustrations of the use of model (13) are presented in van Eeuwijk (1992a), and van Eeuwijk *et al.* (1995).

C. TWO-WAY REDUCED RANK REGRESSION

Synthetic covariates may be used on both the genotypic as well as the environmental dimension of the table. We define the genotypic synthetic covariate as

$$\xi_i=\sum_{k=1}^{K}\pi_k x_{ik}, \tag{14}$$

with the π_k as unknown parameters to be estimated. This leads to the model

$$Y_{ij}=\mu+\alpha_i+\beta_j+\left[\sum_{k=1}^{K}\pi_k x_{ik}\right]\left[\sum_{h=1}^{H}\lambda_h z_{jh}\right]+E_{ij}, \tag{15}$$

with the recapitulative table given in Table 18.

D. REDUCED RANK REGRESSION INDEPENDENT OF COVARIATES

When I-1 linearly independent genotypic covariates are used to create a synthetic covariate, there is no restriction on the ξ_i's of having to be a linear combination of the x_k's. The same holds true for the ζ_j's when there are J-1 environmental covariates. Model (15) can thus be defined without reference to covariates as

$$Y_{ij}=\mu+\alpha_i+\beta_j+\xi_i\zeta_j+E_{ij}. \tag{16}$$

Model (16) is known under various names, like AMMI model (Gauch, 1988) and bilinear model (Denis, 1991). Recently it was placed in the biadditive model family by Denis and Gower (1992, 1994a), in an attempt to create a more unified nomenclature for models for two-way tables. Table 19 gives the distribution of the degrees of freedom over the model

terms.

For completeness we give the extension of (16) to more than one term,

$$Y_{ij}=\mu+\alpha_i+\beta_j+\sum_{r=1}^{R}\xi_{ir}\zeta_{jr}+E_{ij}. \tag{17}$$

The recapitulative table of model (17) is shown in Table 20.

Models (16) and (17) are extensively used in plant breeding. A good review that emphasizes prediction can be found in Gauch (1992). A brief exposition emphasizing interpretation is given by van Eeuwijk (1992b). Generalized bilinear models are described in van Eeuwijk (1995b).

V. MIXED FACTORIAL REGRESSION

General presentations of the models proposed in the subsections A and B can be found in Goldstein and McDonald (1988), and Denis and Dhorne (1989).

A. GENOTYPES FIXED AND ENVIRONMENTS RANDOM

Sometimes, the environments included in an experiment can be assumed to represent a random sample from a population of environments, thereby fulfilling a sufficient condition for a mixed model approach. Model (2) can be changed into a mixed model by replacing the fixed environmental parameters indexed by j, by random parameters;

$$Y_{ij}=\mu+\alpha_i+B_j+\rho_i z_j+E_{ij}. \tag{18}$$

Although the environments are considered random, z_j is not considered to be random, as the analysis proceeds conditional on the value of z_j. The random parameters B_j can be obtained as best linear unbiased predictions, after estimation of the variance component $\sigma_{BB}=\text{var}(B_j)$ (Searle *et al.*, 1992).

B. ENVIRONMENTS FIXED AND GENOTYPES RANDOM

In the early phases of the selection process, plant breeders tend to work with groups of genotypes that are considered to be samples from larger populations, whose performance needs to be estimated in a number of well defined environments. By inserting a random genotype in model (2) we obtain

$$Y_{ij}=\mu+A_j+\beta_j+R_i z_j+E_{ij}. \tag{19}$$

Variance components to be estimated are $\sigma_{AA}=\text{Var}(A_i)$, $\sigma_{RR}=\text{Var}(R_i)$, and $\sigma_{AR}=\text{Cov}(A_i,R_i)$. For individual genotypic performances, again best linear unbiased predictions can be calculated. A noticeable feature of model (19) is that the variance of Y_{ij} depends on j, the environment,

$$Var(Y_{ij})=\sigma_{AA}+2\sigma_{AR}z_j+\sigma_{RR}(z_j)^2+\sigma_{EE}. \tag{20}$$

When one wants to use model (19) to predict future genetic gain, one should be aware that the gain depends on the particular environment j.

Table 21

Genotypic variances when no covariates are included, when rain has been included, and when maximum and minimum temperature, rain and relative humidity have been included. The asterisks indicate P<0.05 for the χ^2 test proposed by Shukla (1972b).

	No cov.	Rain	MaxT+MinT +Rain+RH
G1	1.9104 *	0.8121 *	0.6109
G2	0.6472	0.4469	0.3895
G3	1.2316 *	0.8261 *	0.6176
G4	0.8452 *	0.6169	0.7042
G5	0.4048	0.3559	0.3828
G6	0.5306	0.5990	0.6612
G7	0.5785	0.5249	0.1611
G8	0.5104	0.5193	0.5838
G9	0.5987	0.5660	0.8117
G10	0.9164 *	0.9870 *	1.2004 *
G11	0.6593	0.7235	0.6811
G12	0.4134	0.4481	0.4988
G13	0.8578 *	0.9308 *	1.2673 *
G14	1.4336 *	1.2328 *	0.8737
G15	1.8593 *	1.1239 *	1.0126 *
G16	0.5851	0.6479	0.4165
G17	0.5139	0.5131	0.6351

C. GENOTYPES FIXED, ENVIRONMENTS RANDOM, AND RANDOM INTERACTION DEPENDING ON GENOTYPE

Shukla (1972a,b) introduced a model that included fixed genotypes and random environments, besides a genotypic specific error component. Interesting applications are present in Kang and Miller (1984), Gorman *et al.* (1989), Kang and Gorman (1989), Gravois *et al.* (1990), Helms (1993), Magari and Kang (1993), and Kang (1993). The model formulation is very similar to (18);

$$Y_{ij}=\mu+\alpha_i+B_j+\rho_i z_j+E_{ij}, \tag{21}$$

with the variance of E_{ij} depending on the genotype;

$$Var(E_{ij})=\sigma_{EE}(i). \tag{22}$$

The variance $\sigma_{EE}(i)$ is usually interpreted as a stability associated with genotype i. Inclusion of more than one covariate is straightforward;

$$Y_{ij}=\mu+\alpha_i+B_j+\sum_{h=1}^{H}\rho_{ih}z_{jh}+E_{ij}. \tag{23}$$

Results of the application of models (21) and (23) to the example data are presented in Table 21.

Deleting the regression term $\rho_i z_j$ from model (21) produces a well-known, particularly

simple type of heteroscedastic model, often discussed in literature (Russel and Bradley, 1958; Shukla, 1982; Snee, 1982; Denis, 1983; Vincourt *et al.*, 1984; Longford, 1987; Searle *et al.*, 1992; Mudholkar and Sarkar, 1992).

VI. ESTIMATION AND TESTING

A. FIXED FACTORIAL REGRESSION

Fixed factorial regression models fall in the class of fixed linear models and therefore no special problems arise with regard to estimation of parameters and testing of hypotheses.

B. REDUCED RANK FACTORIAL REGRESSION

The inclusion of bilinear terms complicates estimation and testing. Closed form least squares estimators and asymptotic variances are only known for orthogonal cases, *i.e.* without missing values and with proportional numbers of replications (Denis and Gower, 1992, 1994b). In non-orthogonal cases, numerical approaches are inevitable, and tests and confidence intervals will be approximate.

C. MIXED FACTORIAL REGRESSION

With the exception of the models developed by Shukla, which seem to need a specific procedure, estimation for mixed factorial regressions can be done using restricted maximum likelihood.

D. SOFTWARE

Most of the models presented can be processed with the main statistical packages that include programming facilities, for example Genstat (1993), SAS (1992), and S-plus (1994). The most important Genstat statements for fixed full and reduced rank factorial regression have been added as an Appendix to van Eeuwijk *et al.* (1995). Special purpose packages also have been developed. We mention first MatModel (Gauch, 1990), which deals mainly with AMMI models. An attractive feature of this package is the cross-validation procedure for assessing the number of interaction terms, when replicates are present. INTERA (Decoux and Denis, 1991) offers facilities for a wide range of fixed factorial regression models and AMMI models, applicable to balanced and unbalanced data. Furthermore, INTERA can fit models combining features of both factorial regression and AMMI. Computer programs to calculate the ecovalence (Wricke, 1962) and Shukla's stability statistics, σ_i^2 and s_i^2 (Shukla, 1972a), are described in Kang (1988, 1989). Presently a new program is available that calculates, in addition to the above mentioned statistics, the YS_i statistic, which combines yield and stability into a single selection criterion (Kang, 1993).

VII. SOME CONSIDERATIONS WITH RESPECT TO MODEL CHOICE

The basic question for the experimenter is, which model to choose out of all the possibilities enumerated above? No definite answer is possible. The choice strongly depends on the desired goal. Various choices accompany the model selection process. We briefly address four issues.

A. CONSTRAINTS

All the models described are overparameterized. Supplementary constraints can be imposed to solve this indeterminacy. Various possibilities exist. Natural extensions of the sum-to-zero constraints were proposed by Denis (1991). These lead to orthogonal decompositions, that are convenient for the construction of recapitulative tables. However, we feel that mathematical convenience should always be made subordinate to biological knowledge, also in choosing identification constraints.

B. COVARIATE SELECTION

The most difficult point in the application of factorial regression models seems to be the choice of a good subset of covariates for genotypes as well as environments. It is a variable selection problem having the square of the complexity of that of variable selection in the standard 'one-way' multiple regression context. It is important to keep in mind that the size of the sample for factorial regressions is not IJ, but I-1 for regressions with genotypic covariates, and J-1 for regressions with environmental covariates.

Denis (1988) contains a discussion of variable selection strategies for factorial regression models. It is shown how nesting relationships between models can be used to test for the inclusion of covariates and the possibility of rank reduction.

In the absence of subject matter knowledge, exhaustive variable searches may be used as exploratory analyses. One should then be cautious against over-interpretation, and correct for selection bias by using an appropriate experiment-wise error rate. If possible, it is, however, always preferable to work inferentially, *i.e.* test specific hypotheses following from subject matter knowledge about the interaction of physiological processes in the plant with defined environmental factors. The relevancy of selected covariates can be further investigated in future trials, as a safeguard against conclusions based on chance correlations.

C. FIXED OR RANDOM

Another important question is the choice of terms as fixed or random. Two main types of arguments can be distinguished. A first type of argument is based on sampling considerations. Do the genotypes and/or environments in the experiment constitute a sample from a population to which the inference is directed? The second kind of arguments is more pragmatic, and involves the desirability of shrinkage and recovery of information, and the convenience of choosing a model term random when many parameters are associated with the term. With regard to shrinkage we may question whether it is reasonable to shrink estimates deviating from the mean of the sample back towards that mean? Or, should relatively good genotypes pay for being an element of a relatively bad sample, while relatively bad genotypes benefit from being an element of a relatively good sample? Considerations concerning recovery of information play a role when data are unbalanced. At all times it must be possible to assess whether the random effects indeed could have come from the assumed distribution. For example, for the estimation of a variance component, at least 10 degrees of freedom should be available, otherwise it is preferable to take the term fixed. The same remark applies to Shukla's approach, many environments are needed for accurate estimates of individual genotypic variances.

D. PARSIMONY

In model building and model choice, one should always take into account the parsimony principle (Gauch, 1988), *i.e.* avoid-over fitting. By including ever more covariates, the amount of interaction described will keep on increasing. However, as a consequence, more noise will

be fitted, leading to less robust models. From this perspective, reduced rank regression models are attractive as they allow more covariates for the same number of degrees of freedom. A word of caution should be given for too uncritically accepting the degrees of freedom attributed to synthetic and artificial covariates. When pattern does not clearly dominate noise, these degrees of freedom will be too low, thus declaring the influence of synthetic and artificial covariates significant, when it is not (Gauch, 1992; Williams and Wood, 1993; Cornelius, 1993).

VIII. ALTERNATIVES AND EXTENSIONS

Despite the long enumeration of models given above, the possibilities of modeling interaction using additional information are not exhausted. In this final section, we give some further ideas on the subject.

A. DECOMPOSING MAIN EFFECTS

Use of covariates for decomposition of variation need not be kept restricted to interaction effects. Main effects can be split into a part due to regression on a covariate and a residual

$$Y_{ij}=\mu+x_i\alpha_0+\alpha_i^*+\beta_0 z_j+\beta_j^*+x_i\nu z_j+E_{ij}. \qquad (6')$$

The residual from a main effects regression almost always is strongly significant because of the dominant role of the main effects in the description of the total variation.

B. ANALYSIS OF COVARIANCE

Some authors (Snedecor and Cochran, 1976; Searle, 1979) have used the following model under the name of analysis of covariance;

$$Y_{ij}=\mu+\alpha_i+\beta_j+\rho o_{ij}+E_{ij}, \qquad (24)$$

where o_{ij} is a covariate whose value depends specifically on the cell (i,j). As previously indicated, covariates defined on cell level can be subsumed under the factorial regression models by defining a genotypic covariate $x_i=o_{i.}-o_{..}$ and an environmental covariate $z_j=o_{.j}-o_{..}$, where a dot means averaging.

C. PARTIAL LEAST SQUARES REGRESSION

Multivariate partial least squares regression models have been proposed to model interaction in dependence on covariates (Aastveit and Martens, 1986; Talbot and Wheelwright, 1989). These models can be interpreted in a way reminiscent of reduced rank regression. Partial least squares can be viewed as a robust estimation procedure.

D. BIADDITIVE MIXED MODELS

An interesting conjunction of model classes is given by allowing the multiplicative covariates in biadditive models, to which the Finlay-Wilkinson and AMMI model belong, to be random. Some preliminary work has been done here by Oman (1991).

E. PIECEWISE REGRESSION

Genotypic responses to many environmental factors will reach an upper limit. A simple way to model this kind of response is;

$$Y_{ij}=\mu+\alpha_i+\beta_j+\rho_i Min(\phi_i,z_j)+E_{ij}. \quad (25)$$

In model (25) the covariate z_j is replaced by the minimum of a threshold ϕ_i and the covariate z_j. Each genotype has its own threshold after which the response cannot increase any more.

F. GENERALIZED LINEAR AND BILINEAR MODELS

All fixed factorial regression models dealt with so far assume that the expectation can be modeled linearly in the parameters and that the variance is constant. Deviations from these assumptions sometimes can be cured by transformation of the response. However, the optimal transformation for achieving linearity need not be the same as the optimal transformation for achieving homogeneity of variance. For the models in the class of generalized linear models it is not necessary to find a transformation of the response as a compromise between first (expectation linear in the parameters) and second order (homogeneous variance) requirements. In generalized linear models a suitable transformation of the expectation can be combined with a convenient choice for a variance function, expressing the dependence of the variance on the mean (McCullagh and Nelder, 1989). Generalized factorial regression models extend considerably the range of application for factorial regression models. A further elaboration in the form of generalized bilinear models is discussed in van Eeuwijk (1995b).

G. HIGHER WAY FACTORIAL REGRESSION

Genotype by environment problems often involve more than two factors. Environments are usually cross-classifiable by years and locations. This fact does not complicate the use of factorial regression models for the fixed and mixed cases. Somewhat harder to make are the extensions to the class of biadditive models, although progress is made also here. In van Eeuwijk and Kroonenberg (1995) quadri-additive models are introduced for three-way interaction.

ACKNOWLEDGEMENT

Thanks are due to L.C.P. (Paul) Keizer for his help in the preparation of the manuscript.

REFERENCES

Aastveit, A.H., and H. Martens. 1986. ANOVA interactions interpreted by partial least squares regression. *Biometrics* 42:826-844.

Abou-el-Fittouh, H.A., J.O. Rawlings and P.A. Miller. 1969. Classification of environments to control genotype by environment interactions with an application to cotton. *Crop Sci* 9:135-140.

Arntzen, F.K., and F.A. van Eeuwijk. 1992. Variation in resistance levels of potato genotypes and virulence level of potato cyst nematode populations. *Euphytica* 62:135-143.

Baril, C.P. 1992. Factor regression for interpreting genotype-environment interaction in bread-wheat trials. *Theor. Appl. Genet.* 83:1022-1026.

Baril, C.P., Denis, J.-B., Wustman, R. and van Eeuwijk, F.A. 1995. Analysing genotype by environment interaction in Dutch Potato Variety Trials using factorial regression. *Euphytica* 82:149-155.

Berbigier, A., J.-B. Denis and C. Dervin. 1980. Interaction variété-lieu: analyse du rendement d'orges de printemps. *Ann. Amélior. Plantes* 30:79-94.

Brennan, P.S., D.E. Byth, D.W. Drake, I.H. DeLacy and D.G. Butler. 1981. Determination of the location and number of test environments for a wheat cultivar evaluation program. *Aust. J. Agric. Res.* 32:189-201.

Brown, K.D., M.E. Sorrells and W.R. Coffman. 1983. Method for classification and evaluation of testing environments. *Crop Sci* 23:889-893.

Byth, D.E., R.L. Eisemann and I.H. DeLacy. 1976. Two-way pattern analysis of a large data set to evaluate genotypic adaptation. *Heredity* 37:215-230.
Charmet, G., F. Balfourier, C. Ravel and J.-B. Denis. 1993. Genotype x environment interactions in a core collection of French perennial rye grass populations. *Theor. Appl. Genet.* 86:731-736.
Cornelius, P.L. 1993. Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. *Crop Sci* 33:1186-1193.
Corsten, L.C.A., and J.-B. Denis. 1990. Structuring interaction in two-way tables by clustering. *Biometrics* 46:207-215.
Crossa, J., W.H. Pfeiffer, P.N. Fox and S. Rajaram. 1990. Multivariate analysis for classifying sites: application to an international wheat yield trial. p. 214-233. *In* M.S. Kang (ed.) *Genotype-by-environment interaction and plant breeding.* Louisiana State Univ., Baton Rouge.
Davies, P.T., and M.K.-S. Tso. 1982. Procedures for reduced-rank regression. *Appl. Statist.* 31:244-255.
Decoux, G., and J.-B. Denis. 1991. INTERA, version 3.3, Notice d'utilisation, logiciel pour l'interprétation statistique de l'interaction entre deux facteurs. Laboratoire de Biométrie, INRA, route de St-Cyr, F-78026 Versailles Cédex, 175pp.
Denis, J.-B. 1979. Structuration de l'interaction. *Biom. Praxim.* 19:15-34.
Denis, J.-B. 1980. Analyse de régression factorielle. *Biom.Praxim.* 20:1-34.
Denis, J.-B. 1983. Extension du modèle additif d'analyse de variance par modélisation multiplicative des variances. *Biometrics* 39:849-856.
Denis, J.-B. 1988. Two-way analysis using covariates. *Statistics* 19:123-132.
Denis, J.-B. 1991. Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Rev. Statist. Appl.* 39:5-24.
Denis, J.-B., and T. Dhorne. 1989. Modelling interaction by regression with random coefficients. *Biuletyn Oceny Odmian* 21-22:63-73.
Denis, J.-B., and J.C. Gower. 1992. Biadditive models. Technical report of the Laboratoire de Biométrie, INRA, F-78026 Versailles, 33 pp.
Denis, J.-B., and J.C. Gower. 1994a. Biadditive models. Letter to the editor. *Biometrics* 50:310-311.
Denis, J.-B., and J.C. Gower. 1994b. Asymptotic covariances for the parameters of biadditive Models. *Utilitas Mathematica* 46: 193-205.
Denis, J.-B., et P. Vincourt. 1982. Panorama des méthodes statistiques d'analyse des interactions génotype x milieu. *Agronomie* 2:219-230.
Fakorede, M.A.B., and B.O. Opeke. 1986. Environmental indices for the analysis of genotype x environment interaction in maize. *Maydica* 31:233-243.
Federer, W.T., and B.T. Scully. 1993. A parsimonious statistical design and breeding procedure for evaluating and selecting desirable characteristics over environments. *Theor. Appl. Genet.* 86:612-620.
Finlay, K.W., and G.N. Wilkinson. 1963. The analysis of adaptation in a plant-breeding program. *Aust. J. Agric. Res.* 14:742-754.
Freeman, G.H., and P. Crisp. 1979. The use of related variables in explaining genotype-environment interactions. *Heredity* 42:1-11.
Freeman, G.H., and J.M. Perkins. 1971. Environmental and genotype-environmental components of variability. VIII. Relations between genotypes grown in different environments and measures of these environments. *Heredity* 27:15-23.
Gabriel, K.R. 1978. Least squares approximation of matrices by additive and multiplicative models. *J. R. Statist. Soc.* B 40:186-196.
Gauch, H.G.Jr. 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44:705-715.
Gauch, H.G.Jr. 1990. *MATMODEL version 2.0, AMMI and related analyses for two-way data matrices.* Cornell University, Ithaca, New-York 14853, USA, 69pp.
Gauch, H.G.Jr. 1992. *Statistical analysis of regional yield trials: AMMI analysis of factorial designs.* Elsevier, Amsterdam, 278pp.
Genstat 5 Comittee. 1993. *Genstat 5 release 3, reference manual.* Clarendon Press, Oxford.
Goldstein, H., and R.P. McDonald. 1988. A general model for the analysis of multi level data. *Psychometrika* 53:455-467.
Gorman, D.P., M.S. Kang and M.R. Milam. 1989. Contribution of weather variables to genotype x environment interaction in grain sorghum. *Plant Breeding* 103:299-303.
Gravois, K.A., K.A.K. Moldenhauer and P.C. Rohman. 1990. Genotype-by-environment interaction for rice yield and identification of stable, high-yielding genotypes. p. 181-188. *In* M.S. Kang (ed.) *Genotype-by-environment interaction and plant breeding.* Louisiana State Univ., Baton Rouge.
Hardwick, R.C. 1972. Method of investigating genotype environment and other two factor interactions. *Nature New Biology* 236:191-192.

Hardwick, R.C., and J.T. Wood. 1972. Regression methods for studying genotype-environment interactions. *Heredity* 28:209-222.
Helms, T.C. 1993. Selection for yield and stability among oat lines. *Crop Sci* 33:423-426.
Hinkelmann, K. 1974. Genotype-environment interaction: Aspects on statistical design, analysis and interpretation. Invited paper to the 8th International Biometric Conference, Constanta, Romania.
Horner, T.W., and K.J. Frey. 1957. Methods for determining natural areas for oat varietal recommendations. *Agron. J.* 49:313-315.
Izenman, A.J. 1975. Reduced-rank regression models for the multivariate linear model. *J. Mult. Anal.* 5:248-264.
Kang, M.S. 1988. Interactive BASIC program for calculating stability-variance parameters. *Agron. J.* 80:153.
Kang, M.S. 1989. A new SAS program for calculating stability-variance parameters. *J. Hered.* 80:415.
Kang, M.S. 1993. Simultaneous selection for yield and stability in crop performance trials: consequences for growers. *Agron. J.* 85:754-757.
Kang, M.S., and D.P. Gorman. 1989. Genotype x environment interaction in maize. *Agron. J.* 81:662-664.
Kang, M.S., and J.D. Miller. 1984. Genotype x environment interactions for cane and sugar yield and their implications in sugar cane breeding. *Crop Sci* 24:435-440.
Knight, R. 1970. The measurement and interpretation of genotype-environment interactions. *Euphytica* 19:225-235.
Lefkovitch, L.P. 1985. Multi-criteria clustering in genotype-environment interaction problems. *Theor. Appl. Genet.* 70:585-589.
Lin, C.S., and G. Butler. 1988. A data-based approach for selecting locations for regional trials. *Can. J. Plant. Sci.* 68:651-659.
Lin, C.S., and B. Thompson. 1975. An empirical method of grouping genotypes based on a linear function of the genotype-environment interaction. *Heredity* 34:255-263.
Longford, N.T. 1987. A fast algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika* 74:817-827.
Magari, R., and M.S. Kang. 1993. Genotype selection via a new yield-stability statistic in maize yield trials. *Euphytica* 70:105-111.
Mandel, J. 1961. Non-additivity in two-way analysis of variance. *J. Am. Statist. Ass.* 56:878-888.
McCullagh, P., and J.A. Nelder. 1989. *Generalized linear models*. Chapman and Hall, London.
McGraw, R.L., P.R. Beuselinck and R.R. Smith. 1986. Effect of latitude on genotype x environment interactions for seed yield in birdsfoot trefoil. *Crop Sci* 26:603-605.
Mudholkar, G.S., and I.C. Sarkar. 1992. Testing homoscedasticity in a two-way table. *Biometrics* 48:883-888.
Muir, W., W.E. Nyquist and S. Xu. 1992. Alternative partitioning of the genotype-by-environment interaction. *Theor. Appl. Genet.* 84:193-200.
Obadia, J. 1978. L'analyse en composantes explicatives. *Rev. Statist. Appl.* 26:5-28.
Oliveira, J.A., and G. Charmet. 1992. Genotype by environment interaction in Lolium perenne: grouping wild populations by cluster analysis. *Invest. Agr.: Prod. Prot. Veg.* 7:117-128.
Oman, S.D. 1991. Multiplicative effects in mixed model analysis of variance. *Biometrika* 78:729-739.
Paul, H., F.A. van Eeuwijk and W. Heijbroek. 1993. Multiplicative models for cultivar by location interaction in testing sugar beets for resistance to beet necrotic yellow vein virus. *Euphytica* 71:63-74.
Rameau, C., and J.-B. Denis. 1992. Characterization of environments in long-term multi-site trials in Asparagus, through yield of standard varieties and use of environmental covariates. *Plant Breeding* 109:183-192.
Rao, C.R. 1964. The use and interpretation of principal component analysis in applied research. *Sankhya* B 26:329-358.
Royo, C., A. Rodriguez and I. Romogosa. 1993. Differential adaptation of complete and substituted triticale. *Plant Breeding* 111:113-119.
Russel, T.S., and R.A. Bradley. 1958. One-way variances in a two-way classification. *Biometrika* 45:111-129.
Sabatier, R., J.-D. Lebreton and D. Chessel. 1989. Principal component analysis with instrumental variables as a tool for modelling composition data. p. 341-352. *In* R. Coppi and S. Bolasco (eds) *Multiway Data Analysis*. Elsevier, North-Holland.
Saeed, M., and C.A. Francis. 1984. Association of weather variables with genotype x environment interactions in grain sorghum. *Crop Sci* 24:13-16.
SAS Institute Inc., SAS*. 1992. *Software: changes and enhancements, release 6.07*. Technical report P-229, SAS/STAT*, Cary, NC: SAS Institute Inc., 620 pp.
Searle, S.R. 1979. Alternative covariance models for the 2-way crossed classification. *Comm. St.* A 8:799-818.
Searle, S.R., G. Casella and C.E. McCulloch. 1992. *Variance Components*. Wiley, New York.
Seif, E., J.C. Evans and L.N. Balaam. 1979. A multivariate procedure for classifying environments according to their interaction with genotypes. *Aust. J. Agric. Res.* 30:1021-1026.

Shukla, G.K. 1972a. An invariant test for the homogeneity of variances in a two-way classification. *Biometrics* 28:1063-1072.
Shukla, G.K. 1972b. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29:237-245.
Shukla, G.K. 1982. Testing the homogeneity of variances in a two-way classification. *Biometrika* 69:411-416.
Snedecor, G.W., and W.G. Cochran. 1976. *Statistical Methods* (6th edn., 8th printing). Iowa State Univ. Press, Ames.
Snee, R.D. 1982. Nonadditivity in a two-way classification: is it interaction or nonhomogeneous variance?. *J. Am. Statist. Ass.* 77:515-519.
S-plus. 1994. *S-plus for Windows version 3.2 supplement.* Seattle: StatSci, a division of MathSoft, Inc.
Tai, G.C.C. 1990. Path analysis of genotype-environment interactions. p. 273-286. *In* M.S. Kang (ed.) *Genotype-by-environment interaction and plant breeding.* Louisiana State Univ., Baton Rouge.
Talbot, M., and A.V. Wheelwright. 1989. The analysis of genotype x environment interactions by partial least squares regression. *Biuletyn Oceny Odmian* 21-22:19-25.
Tso, M.K.-S. 1981. Reduced-rank regression and canonical analysis. *J. R. Statist. Soc.* B 43:183-189.
van den Wollenberg, A.L. 1977. Redundancy analysis an alternative for canonical correlation analysis. *Psychometrika* 42:207-219.
van der Leeden, R. 1990. Reduced rank regression with structured residuals. DSWO Press, Leiden.
van Eeuwijk, F.A. 1992a. Interpreting genotype-by-environment interaction using redundancy analysis. *Theor. Appl. Genet.* 85:89-100.
van Eeuwijk, F.A. 1992b. Multiplicative models for genotype-environment interaction in plant breeding. *Statistica Applicata* 4:393-406.
van Eeuwijk, F.A. 1993. Genotype by environment interaction; Basic ideas and selected topics. p. 91-104. *In* Johan H.L. Oud and Rian A.W. van Blokland-Vogelesang (eds.) *Advances in longitudinal and multivariate analysis in the behavioral sciences.* ITS, Nijmegen.
van Eeuwijk, F.A. 1995a. Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models. *Euphytica* 84:1-7.
van Eeuwijk, F.A. 1995b. Multiplicative interactions in generalized linear models. *Biometrics* 51:1017-1032.
van Eeuwijk, F.A., and A. Elgersma. 1993. Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. *Heredity* 70:447-457.
van Eeuwijk, F.A., L.C.P. Keizer and J.J. Bakker. 1995. Linear and bilinear models for the analysis of multi-environment trials : II. An application to data from the Dutch Maize Variety Trials. *Euphytica* 84:9-22.
van Eeuwijk, F.A., and P.M. Kroonenberg. 1995. Multiplicative decompositions of interactions in three-way analysis of variance, with applications to plant breeding. Submitted.
Velu, R.P. 1991. Reduced rank models with two sets of regressors. *Appl. Statist.* 40:159-170.
Vincourt, P., M. Derieux and A. Gallais. 1984. Quelques méthodes de choix des génotypes à partir d'essais multilocaux. *Agronomie* 4:843-848.
Williams, E.R., and J.T. Wood. 1993. Testing the significance of genotype-environment interaction. *Aust. J. Agric. Res.* 35:359-362.
Wood, J.T. 1976. The use of environmental variables in the interpretation of genotype-environment interaction. *Heredity* 37:1-7.
Wricke, G. 1962. Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. *Z. Pflanzenzücht.* 47:92:96.

Chapter XV

Multiplicative decompositions of interactions in three-way anova, with applications to plant breeding

Fred A. van Eeuwijk
DLO-Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O.Box 16, 6700 AA Wageningen

Pieter M. Kroonenberg
Department of Education, Leiden University, Wassenaarseweg 52, 2333 AK Leiden

Multiplicative decompositions of interactions in three-way anova, with applications to plant breeding

Summary

In plant breeding multiplicative models for two-way analysis of variance interaction have become a general means of describing genotype-by-environment interaction. Multiplicative models for genotype-by-environment interaction offer parsimonious descriptions and facilitate interpretations in biological terms. A disadvantage of the prevailing dominance of two-way multiplicative models is that data with more complicated environmental structure are often forced to fit the two-way framework. As a partial solution to this problem, three-way multiplicative models are presented that can be used in addition to the more familiar two-way multiplicative models. Most importantly, a three-way generalization is given of the two-way singular value decomposition, that can be applied for closer inspection of three-way analysis of variance interaction, much in the same way as its two-way counterpart is used for two-way interaction. Two real data sets are analyzed to illustrate how two- and three-way multiplicative models for interaction jointly can provide extensive descriptions of genotype-by-environment interactions. To a major extent these descriptions are open to meaningful biological interpretations as well.

1. Introduction

A typical experiment in plant breeding consists of the evaluation of a number of genotypes under a range of circumstances supposedly related to future growing environments. Data summaries take the form of two-way tables with one way consisting of genotypes, and the other way containing factorial combinations of treatment and environmental factors. This kind of summary is a direct consequence of the interpretation of the phenotype as the joint product of genotype and environment, where the environment encompasses everything that is non-genetic.

Reducing genotypic evaluations to two-way tables invites researchers to fit two-way models for description. For the interaction between the genotypic and environmental factor, the genotype-by-environment interaction (GEI), a host of models is available between the additive two-way analysis of variance (ANOVA) model without terms for interaction and the full interaction model with a separate interaction parameter for every genotype-by-environment combination. Two approaches can be distinguished. In external modeling, use is made of information on genotypes and/or environments measured or observed additionally to the response which makes up the content of the two-way table to be analyzed. In internal

modeling only information present or to be derived from the response in the table is used.

A popular class of models to describe interaction in two-way tables is that of the multiplicative interaction models. The best known internal models in this class are the concurrence model (Mandel, 1961), the regression on the mean model (Yates and Cochran, 1938; Mandel, 1961; Finlay and Wilkinson, 1963), and the Additive Main effects and Multiplicative Interaction effects model or AMMI model (Gollob, 1968; Mandel, 1969; Freeman and Dowker, 1973; Gauch, 1988). Based on external modeling are factorial regression (Denis, 1988; Baril, 1992), and reduced-rank factorial regression (van Eeuwijk, 1992). All models have in common that they can be interpreted as describing GEI by differential sensitivity to either real or hypothetical environmental variables.

Though the incorporation of measured environmental variables in models for two-way tables deserves more attention (van Eeuwijk, Denis and Kang, 1995), in the past the emphasis has always been on models without external information. In Denis and Gower (1992) a survey is given of models for two-way tables without external information which include zero, one, or two main effects, and a number of multiplicative effects depending on the size of the table and the number of main effects already included. This class of models is called the class of bi-additive models, and includes besides the internal multiplicative interaction models mentioned above, the SHifted Multiplicative Model (SHMM), which consists of an intercept term followed by a number of multiplicative terms (Seyedsadr and Cornelius, 1992). The SHMM breaks with the traditional distinction between main effects and interaction effects. Its rationale is given by its parsimony, i.e. the small number of parameters necessary for data description, and prediction. Furthermore, it can be used to search for separability (Cornelius, Seyedsadr and Crossa, 1992), and identification of groups of genotypes and environments without environmental and genotypic rank change, respectively (Crossa *et al.*, 1993). However, in the majority of the cases there are still sufficient reasons to retain main effects, which favors models restricting multiplicativity to the interaction.

Nowadays, the AMMI model seems to be the most popular (internal) multiplicative interaction model. In the absence of external environmental information, or clear ideas about which information to take into account, it can suggest hypotheses about the interaction structure (see below). In many applications AMMI scores provide a means for sensible interpretation of GEI. This is especially true for environmental factors which possess no further factorial or nested structure, or for those complexes of factors for which the effects on the response variable may be expected to be qualitatively similar. For example, effects of years and locations are often assumed to be only quantitatively different, implying that lack of years could be compensated for by more locations (Schutz and Bernard, 1967, see also Discussion).

Ignoring factorial structure in the environmental dimension of two-way tables, however, is wasting information. A step forward would be the extension of two-way models to three-way models, including facilities for modeling two-way and three-way interactions multiplicatively. To this end, the environmental dimension of the two-way table can be

dissected in two complexes of environmental factors, e.g. location by year structure can be uncovered in multi-location multi-year variety trials, and strain by year structure in resistance breeding trials.

For bi-additive models the usual two-way singular value decomposition (SVD) of matrices offers the basis for an estimation procedure for the multiplicative terms (provided the two-way table is complete). Depending on the type of model, the two-way table of data is corrected for zero, one, or two main effects, and subsequently the matrix of residuals is decomposed multiplicatively. We will present a three-way generalization of the SVD which can be used for estimation of three-way multiplicative terms in three-way models, much in the same way as the familiar two-way SVD can be used in two-way modeling. There is no necessity to restrict three-way decomposition to decomposition of only the additive ANOVA three-way interaction, but application to three-way tables of data which have not been corrected for all main effects and two-way interactions immediately raises all sorts of questions with respect to marginality relations (Gower, 1977, 1995). This subject still needs further research. We will concentrate on three-way decomposition of the three-way interaction, and show how this device can be used in addition to multiplicative decompositions of the two-way interactions to describe and interpret GEI. Three-way interaction will be factorized as a sum of products between one set of genotypic descriptors and two sets of environmental descriptors, e.g. locations and years. The aim of decomposing the three-way interaction is to identify a small number of three-way contrasts and conditional outliers that are responsible for the major part of that interaction. The three-way decomposition firstly serves the goal of statistical parsimony and may secondly help the plant breeder to identify *combinations* of environmental conditions that are beneficial or detrimental to a small subset of genotypes.

General principles of three-way data analysis for genotype by environment trials will be illustrated on two example sets. Firstly, maize data from the official Dutch Maize Variety Trials, having variety by location by year structure, will be analyzed. After fitting a standard three-way ANOVA model, first the ANOVA two-way interactions are replaced by two-way multiplicative models for interaction. Subsequently, the ANOVA three-way interaction will be decomposed by a three-way generalization of the SVD. It is shown that three-way interaction in a three-way genotype by location by year table is essentially different from two-way genotype by location interaction, or genotype by year interaction.

A second example is taken from resistance breeding. The data are head blight incidences on a number of wheat genotypes over the period 1990-1993 due to infection with various *Fusarium* species and strains. Two-way and three-way ANOVA interactions are decomposed. Again, the nature of the three-way interaction was of a type that would have been hard to disclose by methods other than three-way decomposition.

2. Two-way models

In this section first the basics of two-way models with additive and multiplicative terms

will be reviewed. In section 3 three-way generalizations will be presented. For this section we assume that the data are arranged in a two-way table of genotypes by environments, and both genotypes and environments are fixed factors. The response, say yield, for a genotype g (g=1,..,G), in an environment e (e=1,...,E), is represented by the random variable Y_{ge}. The two-way ANOVA model including additive two-way interaction for the response reads

$$\mathcal{E}(Y_{ge}) = \mu + \alpha_g + \beta_e + \alpha\beta_{ge} \ . \qquad (1)$$

In this two-way ANOVA model, μ denotes the general mean, α_g the genotypic main effect, β_e the environmental main effect, and $\alpha\beta_{ge}$ the genotype by environment interaction. $\mathcal{E}(.)$ indicates the expectation operator. For identification there are sum-to-zero constraints over g for the genotypic main effect, over e for the environmental main effect, and over g and e for the interaction. The error is assumed to be independently identically distributed (i.i.d.) with zero mean and constant variance.

In (1) each cell has its own interaction parameter. Multiplicative models for GEI in two-way tables typically replace the additive interaction terms $\alpha\beta_{ge}$ (for which there are $(G\text{-}1)(E\text{-}1)$ independent parameters) by multiplicative terms whose parameters depend on either the genotypes or the environments, and so, usually, need considerably fewer parameters. The multiplicative single-indexed parameters are the natural complement of the single-indexed additive parameters, and allow for relatively simple interpretations. Mostly, multiplicative interaction is interpreted as differential genotypic sensitivity to real or hypothetical environmental variables.

A classic example of a multiplicative model for interaction is the row-regression or regression on the environmental mean model (Yates and Cochran, 1938; Mandel, 1961; Finlay and Wilkinson, 1963). Yield for the separate genotypes is regressed on the mean yield of all genotypes in a particular environment, i.e.

$$\mathcal{E}(Y_{ge}) = \mu + \alpha_g + \beta_e + \rho_g\beta_e \ . \qquad (2)$$

Interaction is described as differential genotypic sensitivity with respect to a biological measure of the environment, in particular, the mean of the genotypes or, equivalently, the environmental main effect, β_e. The sensitivity is expressed through the regression coefficients ρ_g, which for formulation (2) sum to zero.

When in (2) the environmental main effect, β_e, in the interaction term is replaced by a measured environmental variable, z_e, we get a simple form of a factorial regression model,

$$\mathcal{E}(Y_{ge}) = \mu + \alpha_g + \beta_e + \rho_g z_e \ . \qquad (3)$$

It is straightforward to extend the number of environmental variables, and also the inclusion of explanatory covariates on the genotypic dimension of the table presents no new features

(van Eeuwijk, Denis and Kang, 1995).

A simplification of (2), in which all genotypic response lines intersect in the same point, produces the concurrence model (Mandel, 1961). This model can be understood as the regression formulation of Tukey's one-degree-of-freedom for non-additivity test (Tukey, 1949). The model is of the form

$$\mathcal{E}(Y_{ge}) = \mu + \alpha_g + \beta_e + k\alpha_g\beta_e \ . \qquad (4)$$

The concurrence model can be derived from the regression on the mean model by taking the sensitivity to the environment, ρ_g, equal to a constant, k, times the genotypic main effect, α_g.

Models (2) and (4) provide simple internal descriptions of the interaction between genotypes and environments by means of a one-dimensional representation of the environments, and the genotypes are supposed to differ in sensitivity with respect to the environmental characterization implied by the model. However, in many instances the interaction will have a different form and higher dimensionality. An extension of (2), and of (4), is given by model (5),

$$\mathcal{E}(Y_{ge}) = \mu + \alpha_g + \beta_e + \sum_{s=1}^{S} a_{gs}^{*} b_{es}^{*} = \mu + \alpha_g + \beta_e + \sum_{s=1}^{S} \lambda_s a_{gs} b_{es} \ , \qquad (5)$$

which is a combination of ANOVA main effects for genotypes and environments, and multiplicative effects obtained from an SVD of the matrix of residuals from additivity (when there are no empty cells). Only the first S multiplicative terms are retained, where the maximum of S is equal to the minimum of G-1 and E-1 (see Number of multiplicative terms to retain). Important contributions are due to Gollob (1968), Mandel (1969, 1971), Freeman and Dowker (1973), and Gauch (1988). The multiplicative interaction parameters, or scores, for the genotypes are given by the a_{gs}^{*}'s or a_{gs}'s, and those for environments by the b_{es}^{*}'s or b_{es}'s, where the unstarred score vectors are not only orthogonal but also of unit length. The λ_s's are the singular values. The model is known under many names of which the Additive Main effects and Multiplicative Interaction effects model (AMMI) is perhaps the best known in agriculture. Recently Denis and Gower (1992, 1994, 1995) made a strong plea to call it a bi-additive model.

The genotypic and environmental scores belonging to a specific multiplicative term in (5), generally allow for one of three interpretations. The most common interpretation is that the environmental scores represent an environmental variable to which genotypes respond linearly, their sensitivities being expressed by the genotypic scores. For this *regression interpretation* to be valid the environmental scores have to be dispersed evenly over the range. When the environmental scores can be divided into a cluster of positive values on the one hand, and a cluster of negative values on the other hand (ignoring the near-zero values for the moment), a *contrast interpretation* is valid. The situation in which all but a few values are near-zero

requires an *outlier interpretation*. For higher dimensional interaction the situation becomes more complicated in that the multiplicative terms (axes) should be considered simultaneously. However, the above account remains valid for these situations, provided a rotation can be found that leads to a simple structure.

Combinations of various of the above two-way models for description of a particular additive ANOVA two-way interaction are also possible. One might think of a combination of row-regression and factorial regression, e.g. the environmental mean is used in combination with explicit measures of the environment. Another appealing combination consists in the combination of row-regression followed by SVD of the remainder. Similarly, factorial regression may be followed by SVD.

In contrast to the above two-way models, and especially to the AMMI model to which it is most similar in appearance, the SHMM does not include the additive main effects. It reads

$$\mathcal{E}(Y_{ge}) = \nu + \sum_{s=1}^{S} a_{gs}^{*} b_{es}^{*} = \nu + \sum_{s=1}^{S} \lambda_s a_{gs} b_{es} \ . \tag{6}$$

The parameter ν is called the shift parameter. Least squares estimation is not as easy as for the AMMI model. An alternating least squares algorithm can be used (Seyedsadr and Cornelius, 1992), but an exhaustive exploration algorithm seems to do better (Denis and Gower, 1992).

3. Two-way models and derivations thereof in three-way contexts

To describe models for three-way tables we assume that we are dealing with the response, Y_{gly}, for genotype g (g=1,...,G), at location l (l=1,...,L), in year y (y=1,...,Y). Data are arranged in a three-way G by L by Y table, and the factors *Genotype*, *Location*, and *Year* are taken as fixed. The three-way ANOVA model for this response is

$$\mathcal{E}(Y_{gly}) = \mu + \alpha_g + \beta_l + \gamma_y + \alpha\beta_{gl} + \alpha\gamma_{gy} + \beta\gamma_{ly} + \alpha\beta\gamma_{gly} \ . \tag{7}$$

Identifiability constraints for individual terms in (7) usually are of the sum-to-zero type, for each index attached to the term. The error term is again i.i.d. with zero mean and constant variance. In model (7) each combination of *Genotype*, *Location*, and *Year* is connected to four interaction parameters, one from each of the three two-way interactions and one from the three-way interaction.

A simple first move to a more parsimonious model is to replace each of the two-way interaction terms by one of the two-way multiplicative models for interaction discussed in the former section. As an example we mention the replacement of the genotype by location interaction, $\alpha\beta_{gl}$, by a regression on the location mean, and that of the genotype by year interaction, $\alpha\gamma_{gy}$, by a regression on the year mean,

$$\alpha\beta_{gl} = \rho_g \beta_l \quad \text{and} \quad \alpha\gamma_{gy} = \varrho_g \gamma_y \ . \tag{8}$$

The replacement of the location by year interaction, $\alpha\gamma_{ly}$, by a regression on the mean is not so evident from a physiological point of view. Comparing expressions (8) and (2), it will be clear how the same kind of substitution can be carried out using concurrence, factorial regression, or AMMI reformulations of the ANOVA two-way interactions in model (7). Also mixed forms may be chosen, e.g. SVD of the genotype by location interaction and factorial regression for the genotype by year interaction.

For a more parsimonious description of the three-way interaction, generalizing the concurrence model, the following rephrasal springs to mind,

$$\alpha\beta\gamma_{gly} = k\alpha_g \beta_l \gamma_y \ . \tag{9}$$

An early reference to this three-way extension of Tukey's one-degree-of-freedom test for non-additivity is Harter and Lum (1962), cited in Boik and Marasinghe (1989), where also the multiway case is discussed.

Various three-way generalizations of the regression on the environmental mean model can be thought of. The first one defines three-way interaction as a genotypic sensitivity to an environmental, location by year, characterization derived from the cross-product of the corresponding main effects;

$$\alpha\beta\gamma_{gly} = \rho_g \beta_l \gamma_y \ . \tag{10}$$

A second generalization interprets the three-way interaction as differential sensitivity to characterizations of the environment derived from the two-way location by year interaction;

$$\alpha\beta\gamma_{gly} = \rho_g (\beta\gamma)_{ly} \ . \tag{11}$$

More generalizations can be developed, but, taking into account the character of our data, it is not to be expected that either the three-way concurrence generalization, nor the three-way generalizations of the regression on the environmental mean will provide useful tools to model three-way genotype by environment ANOVA interaction, $\alpha\beta\gamma_{gly}$. We expect three-way ANOVA GEI to have the character of incidental corrections for genotypes performing excessively well or badly under particularly favorable or detrimental conditions respectively. Stress situations form a major source of three-way interactions, and as such subscribe to this interpretation of three-way GEI. Therefore, it is hard to imagine, that differential genotypic sensitivity to cross-products of environmental variables (or main effects) can give satisfactory descriptions of three-way interaction. Three-way GEI will be of the outlier or contrast type, seldom of the regression type. In this respect, three-way GEI clearly differs from two-way GEI, which will be more of the contrast and regression type. What is required is a method that can help us in

identifying specific combinations of locations and years that cause aberrant responses in some genotypes. For the two-way situation the two-way SVD has been shown to be a powerful exploratory tool. We will introduce a three-way generalization of the two-way SVD that provides parsimonious descriptions of three-way ANOVA interactions and facilitates the interpretation of these interactions.

4. A multiplicative three-way decomposition

There are various three-way generalizations of the two-way SVD but the one most useful for our purposes, i.e. identification of contrasts and outliers, was that due to Tucker (1966), who used it in what he called 'three-mode factor analysis'. Later Kroonenberg and De Leeuw (1980), and Kroonenberg (1983) elaborated this model in the context of what they called 'three-mode principal component analysis'. Basford, Kroonenberg and DeLacy (1991) demonstrated the use of three-mode principal components for the analysis of two-way genotype by environment interaction for a number of traits simultaneously.

For a better understanding of the Tucker three-way proposal it is good to reconsider the two-way case. Let $\mathbf{X}$ denote a IxJ two-way data table. (Although we will be dealing with data arrays in the form of tables and not with matrices, we will for convenience ignore this difference and use the terms interchangeably.) Fitting a multiplicative model with P terms to $\mathbf{X}$ is equivalent to finding a rank P approximation to $\mathbf{X}$ ($P \leq \min(I,J)$), or writing $\mathbf{X}$ as a product of two rank P matrices, $\mathbf{X} = \mathbf{AB}^T$, with $\mathbf{A}$ and $\mathbf{B}$ minimizing, in least square sense, the function $\phi_1(\mathbf{A},\mathbf{B}) = \|\mathbf{X} - \mathbf{AB}^T\|^2 = \text{trace}((\mathbf{X} - \mathbf{AB}^T)(\mathbf{X} - \mathbf{AB}^T)^T)$. Use of the SVD of $\mathbf{X}$ guarantees a unique solution to this minimization problem, apart from possible columnwise changes of sign (Eckart and Young, 1936; Rao, 1964). If the SVD of $\mathbf{X}$ is $\mathbf{UAV}^T$, with $\mathbf{U}$ the IxP matrix of left singular vectors, with $\mathbf{A}$ the PxP diagonal matrix of singular values (in decreasing order), and with $\mathbf{V}$ the JxP matrix of right singular vectors then the best rank P approximation to $\mathbf{X}$ is given by the product $\mathbf{AB}^T$ with $\mathbf{A} = \mathbf{U}_{(P)}\mathbf{\Lambda}_{(P)}^{c}$ and $\mathbf{B} = \mathbf{V}_{(P)}\mathbf{\Lambda}_{(P)}^{1-c}$ ($0 \leq c \leq 1$). The subscript indicates that only the first P columns of the pertinent matrices are retained.

The SVD of $\mathbf{X}$ also provides a direct solution to the related problem of finding an orthonormal IxP matrix $\mathbf{A}$, a diagonal PxP matrix $\mathbf{G}$, and an orthonormal JxP matrix $\mathbf{B}$ that minimize the loss function $\phi_2(\mathbf{A},\mathbf{B},\mathbf{G}) = \|\mathbf{X} - \mathbf{AGB}^T\|^2$; $\mathbf{A} = \mathbf{U}_{(P)}$, $\mathbf{B} = \mathbf{V}_{(P)}$ and $\mathbf{G} = \mathbf{\Lambda}_{(P)}$. Moreover, for two-way matrices, no improvement over the SVD solution is achievable by allowing $\mathbf{G}$ to be non-diagonal or non-square. Thus the number of columns of $\mathbf{A}$ and $\mathbf{B}$ will necessarily be equal and a column of $\mathbf{A}$ can only combine with the corresponding column of $\mathbf{B}$. The importance of such combinations is determined by the diagonal (non-zero) elements of $\mathbf{G}$.

Now let $\mathbf{X}$ denote a three-way table (matrix) with I rows, J columns and K layers. Analogous to the situation for a two-way table, one might think of the fitting of a

multiplicative model to a three-way table, **X**, as the search for a lower rank approximation to **X**. However, as there is for the three-way case no equivalent to the two-way result row-rank = column-rank, there is firstly no reason why the number of components for rows, columns and layers should be equal. Secondly, there is, in contrast to the two-way case, no a priori reason for restricting the components in one way to combine with only one of the components in the other ways. In Tucker's three-way decomposition of a three-way matrix **X**, the typical matrix entry x_{ijk} is written as a sum of product terms, each consisting of four multiplicative parameters; a scaling constant indexed by the component numbers for row ($p = 1,...,I$), column ($q = 1,...,J$) and layer ($r = 1,...,K$), λ_{pqr}; row scores indexed by row number ($i = 1,...,I$) and row component number (p), a_{ip}; column scores indexed by column number ($j = 1,...,J$) and column component number (q), b_{jq}; and layer scores indexed by layer number ($k = 1,...,K$) and layer component number (r), c_{kr};

$$x_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\lambda_{pqr}a_{ip}b_{jq}c_{kr} \ . \qquad (12)$$

In (12) every component in one way can combine with every component in another way and the number of components in the three ways can differ.

When a three-way multiplicative model is fitted to the three-way table, **X**, one tries to achieve a satisfactory fit with as few components as possible. Therefore, the number of components in the three ways (P,Q,R) will in a model usually be substantially lower than the maximum number of components in the full three-way decomposition (I,J,K).

Some notation. Let the row scores a_{ip}, the column scores b_{jq}, and the layer scores c_{kr} be the elements of the component matrices **A**, **B** and **C**, respectively. The scaling constants λ_{pqr} form the elements of the *core* matrix **G**. Denote by $\mathbf{X}_{<1;2<3>}$ the Ix(JxK) two-way rearrangement of the IxJxK three-way array **X**, where the rows of the three-way array correspond to the rows of the two-way array, while the columns of the two-way array represent combinations of the columns and layers of the three-way array (Cartesian product), with the index for the columns running fastest, i.e. three-way columns are nested within the three-way layers. In a similar vein define the two-way rearrangements $\mathbf{X}_{<2;3<1>}$ and $\mathbf{X}_{<3;1<2>}$.

Estimation of the parameters of a three-way model in the spirit of formulation (12) for fixed row, column and layer ranks, respectively P, Q, and R, can be based on minimization of the loss function $\phi_3(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{G}) = \left\| \mathbf{X}_{<1;2<3>} - \mathbf{A}\mathbf{G}_{<1;2<3>}(\mathbf{C}^T \otimes \mathbf{B}^T) \right\|^2$ (Kroonenberg and de Leeuw, 1980), with the subscript for **G** indicating appropriate arrangement of the core elements, and the operator $\otimes$ denoting the Kronecker product. Thus, $\mathbf{X}_{<1;2<3>}$ is approximated by $\hat{\mathbf{A}}\hat{\mathbf{G}}_{<1;2<3>}(\hat{\mathbf{C}}^T \otimes \hat{\mathbf{B}}^T)$, with **A** an IxP rank P row component matrix, **B** a JxQ rank Q column component matrix, and **C** a KxR rank R layer component matrix. Equivalently, $\mathbf{X}_{<2;3<1>}$ is approximated by $\hat{\mathbf{B}}\hat{\mathbf{G}}_{<2;3<1>}(\hat{\mathbf{A}}^T \otimes \hat{\mathbf{C}}^T)$ and $\mathbf{X}_{<3;1<2>}$ by $\hat{\mathbf{C}}\hat{\mathbf{G}}_{<3;1<2>}(\hat{\mathbf{B}}^T \otimes \hat{\mathbf{A}}^T)$.

The solution to the minimization of ϕ_3 will in general not be unique as the component

matrices can be multiplied by any non-singular matrix provided that the core matrix is multiplied by the inverse. This non-uniqueness can be used to impose orthonormality on the component matrices. Imposing orthonormality, $\mathbf{G}_{<1;2c3>}\mathbf{G}^{T}_{<1;2c3>}$, $\mathbf{G}_{<2;3c1>}\mathbf{G}^{T}_{<2;3c1>}$ and $\mathbf{G}_{<3;1c2>}\mathbf{G}^{T}_{<3;1c2>}$ will be diagonal and the squares of the elements of $\mathbf{G}$, λ^{2}_{pqr}, will represent the variation explained by the combination of the p-the component of the first way with the q-th component of the second way and the r-th component of the third way. It is especially this property of each component in one way combining orthogonally with all components in the other ways, that makes the Tucker three-way decomposition so useful for finding contrasts (outliers) in three-way ANOVA interactions.

As another consequence of the orthonormality constraints on the model parameters, $\mathbf{X}_{<1;2c3>}\mathbf{X}^{T}_{<1;2c3>} = \mathbf{A}\mathbf{G}_{<1;2c3>}\mathbf{G}^{T}_{<1;2c3>}\mathbf{A}^{T}$, $\mathbf{X}_{<2;3c1>}\mathbf{X}^{T}_{<2;3c1>} = \mathbf{B}\mathbf{G}_{<2;3c1>}\mathbf{G}^{T}_{<2;3c1>}\mathbf{B}^{T}$, and $\mathbf{X}_{<3;1c2>}\mathbf{X}^{T}_{<3;1c2>} = \mathbf{C}\mathbf{G}_{<3;1c2>}\mathbf{G}^{T}_{<3;1c2>}\mathbf{C}^{T}$. So, within every one of its ways this three-way decomposition is equivalent to a two-way spectral decomposition (Weesie and van Houwelingen, 1983).

An approximation to the number of degrees of freedom that corresponds to a particular three-way model can be obtained from the number of parameters estimated minus the number of constraints imposed. The numbers of parameters are IxP for **A**, JxQ for **B**, KxR for **C**, and PxQxR for **G**. Constraints amount to $P^2+Q^2+R^2$ for orthonormality.

An algorithm for estimation, given orthonormality of the component matrices, is given in Kroonenberg (1983, Chpt. 4). Related relevant references are Kroonenberg and De Leeuw (1980) and Kiers, Kroonenberg and Ten Berge (1992). First, note that if $\mathbf{X}_{<1;2c3>} = \mathbf{A}\mathbf{G}_{<1;2c3>}(\mathbf{C}^{T} \otimes \mathbf{B}^{T})$ for **A**, **B** and **C** of fixed rank then $\mathbf{G}_{<1;2c3>} = \mathbf{A}^{T}\mathbf{X}_{<1;2c3>}(\mathbf{C} \otimes \mathbf{B})$, so that **G** can be calculated from **A**, **B**, **C** and **X**, after **A**, **B** and **C** have been estimated. Substitution of $\mathbf{G}_{<1;2c3>} = \mathbf{A}^{T}\mathbf{X}_{<1;2c3>}(\mathbf{C} \otimes \mathbf{B})$ in ϕ_3 produces an expression that contains only the component matrices as unknowns. An iterative algorithm then consists in first solving for **A** given **C** and **B**, then for **B** given **A** and **C**, and finally for **C** given **B** and **A**. Within each cycle of the iterative process estimates for **A**, **B** and **C** are obtained as the eigenvectors of respectively

$$\begin{aligned}
&[\ \mathbf{X}_{<1;2c3>}(\hat{\mathbf{C}} \otimes \hat{\mathbf{B}})\ ][\ \mathbf{X}_{<1;2c3>}(\hat{\mathbf{C}} \otimes \hat{\mathbf{B}})\ ]^{T},\\
&[\ \mathbf{X}_{<2;3c1>}(\hat{\mathbf{A}} \otimes \hat{\mathbf{C}})\ ][\ \mathbf{X}_{<2;3c1>}(\hat{\mathbf{A}} \otimes \hat{\mathbf{C}})\ ]^{T},\\
&[\ \mathbf{X}_{<3;1c2>}(\hat{\mathbf{B}} \otimes \hat{\mathbf{A}})\ ][\ \mathbf{X}_{<3;1c2>}(\hat{\mathbf{B}} \otimes \hat{\mathbf{A}})\ ]^{T}.
\end{aligned} \qquad (13)$$

Starting values can be chosen to be the eigenvectors of $\mathbf{X}_{<1;2c3>}\mathbf{X}^{T}_{<1;2c3>}$ for **A**, of $\mathbf{X}_{<2;3c1>}\mathbf{X}^{T}_{<2;3c1>}$ for **B**, and of $\mathbf{X}_{<3;1c2>}\mathbf{X}^{T}_{<3;1c2>}$ for **C** (being the original Tucker (1966) solution). The process must be continued until a convergence criterion is met (change in residual sums of squares or values of the component scores). After convergence **G** can be calculated.

For the two-way case this algorithm would immediately lead to solution. Recall the SVD of **X** as $\mathbf{X} = \mathbf{U}\Lambda\mathbf{V}^{T}$. According to the algorithm we should postmultiply **X** by **V**; $\mathbf{X}^{*} = \mathbf{X}\mathbf{V} =$

UA. **U** follows directly from the spectral decomposition of $\mathbf{X}^{*}\mathbf{X}^{*T} = \mathbf{U}\mathbf{A}^{2}\mathbf{U}^{T}$. No iteration is necessary.

An alternative algorithm for the Tucker three-way decomposition was presented by Weesie and van Houwelingen (1983). They used the fact that this decomposition leads to a model that would be called quadri-additive, in the terminology developed by Denis and Gower (1995), to derive an estimation algorithm based on alternating regressions. Basically, each block of parameters (**A**,**B**,**C**,**G**) is calculated from a linear regression of the appropriately rearranged matrix **X** on a matrix of regressors constructed from the other three blocks of parameters. As an example, consider formula (12) with fixed values for the a_{ip}'s, b_{jq}'s and c_{kr}'s. What remains is a regression of x_{ijk} on regressors constructed from the products of the component scores, with as coefficients to be estimated the scaling constants λ_{pqr}.

The algorithm based on alternating spectral decompositions works especially well for complete tables and constitutes the basis for the *3Waypack* program for three-way analyses (Kroonenberg and Brouwer, 1993). When missing cells occur various strategies can be followed. In *3Waypack* an EM approach is implemented. A first estimation cycle starts with choosing arbitrary values for the missing cells, after which parameter estimates are obtained by applying the iterative scheme in (13). Subsequently, a second estimation cycle consists in replacing the missing values by the fitted values from the first cycle, and estimating a new set of estimates by again applying scheme (13). The procedure stops when no changes occur anymore in the missing cell estimates and the parameter estimates.

In a second approach the spectral decompositions of the rearranged and reduced three-way matrix in (13) are replaced by SVDs, and the solutions for these SVDs are calculated by means of alternating row and column regressions (Gabriel and Zamir, 1979; De Falguerolles and Francis, 1992, 1994; van Eeuwijk, 1995). Because regressions are placed at the heart of the estimation procedure missing cells cease to be a problem.

The alternating regression algorithm utilizing the quadri-additive nature of the Tucker three-way model of Weesie and van Houwelingen does not need modification going from complete to incomplete data sets, although convergence will require more iterations. Changing the individual regressions to generalized regressions at once creates an algorithm for generalized quadri-additive models (the two-way case is dealt with by De Falguerolles and Francis, 1992, 1994 and van Eeuwijk, 1995). Furthermore, generalizations to other metrics, e.g. for robust estimation, do not cause conceptual problems either, as remarked already by Weesie and van Houwelingen (1983).

5. Three-way modeling using three-way methodology

The simplest and most useful implementation of the three-way methodology of the last section is to arrange the three-way ANOVA interaction parameters, $\alpha\beta\gamma_{ijk}$, in a three-way array, and find a low rank approximation to that array;

$$\alpha\beta\gamma_{gly} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\lambda_{pqr}a_{gp}b_{lq}c_{yr} \ . \tag{14}$$

For successful modelling the numbers of row, column, and layer components necessary for adequate description of the three-way interaction (P,Q,R) should be clearly below the maximum number of components corresponding to the full decomposition (G-1,L-1,Y-1).

The three-way decomposition of the three-way interaction can be combined with multiplicative reformulations of the two-way ANOVA interactions as in

$$\mathcal{E}(Y_{gly}) = \mu+\alpha_g+\beta_l+\gamma_y+ \tag{15}$$
$$\sum_{u=1}^{U}\lambda_u a_{gu}b_{lu} + \sum_{v=1}^{V}\lambda_v a_{gv}b_{yv} + \sum_{w=1}^{W}\lambda_w a_{lw}b_{yw} +$$
$$\sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\lambda_{pqr}a_{gp}b_{lq}c_{yr} \ .$$

From the subscripts of the scores it will be clear to which ANOVA interactions the multiplicative terms refer. In this way all kinds of combinations of two-way multiplicative modeling and three-way multiplicative modeling are possible. To give an example in which regression on the mean formulations have been chosen for the $\alpha\beta_{gl}$ and the $\alpha\gamma_{gy}$ interaction, and the $\beta\gamma_{ly}$ interaction has not been decomposed at all;

$$\mathcal{E}(Y_{gly}) = \mu + \alpha_g + \beta_l + \gamma_y + \rho_g\beta_l + \varrho_g\gamma_y + \beta\gamma_{ly} + \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\lambda_{pqr}a_{gp}b_{lq}c_{yr} \ . \tag{16}$$

Models (15) and (16) are inspired by the full three-way ANOVA model, (7). In (15) and (16), two- and three-way ANOVA interactions are replaced by more parsimonious multiplicative formulations. A tacit assumption in the application of (15) and (16), is that all terms in the three-way ANOVA model were relevant. Models like (15) and (16), can be fitted by first fitting the three-way ANOVA model (7), and then decomposing the two- and three-way matrices of ANOVA interaction parameters individually. This stagewise procedure leads to the least squares solution only for complete tables. For incomplete tables this approach will not lead to the least squares solution, as the parameters in this case have to be estimated simultaneously. Gower (1977) already concluded that this problem is not trivial because of the complications arising from the imposition of constraints upon the parameters. For example, sum-to-zero constraints on the three-way components will automatically induce additional two-way components. For incomplete tables algorithms using alternating regressions may be useful, but in contrast to the situation for two-way tables it is questionable whether the conditions for identifiability of models like (15) and (16) are correctly taken into account by the regressions making up the estimation algorithm.

A three-way version of the SHMM can be defined as

$$\mathcal{E}(Y_{glv}) = \nu + \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\lambda_{pqr}a_{gp}b_{lq}c_{vr} \ . \qquad (17)$$

For fitting this model, a three-way analogue of the exhaustive search algorithm described by Denis and Gower (1992) for the two-way case may be used. During this search simple variations on (17) can be studied as well, like the models for which $\nu = 0$ (cf. Fisher and Mackenzie, 1923; this is also the default choice for three-way multiplicative modelling in psychometrics and chemometrics), and $\nu = \mu$, the general mean.

6. Number of multiplicative terms to retain after decomposition

6.1 Two-way interaction

For the assessment of the number of bi-additive multiplicative terms to retain various methods exist. Let **Z** be a *IxJ* matrix of two-way residuals from additivity. Under the null hypothesis of no multiplicative two-way interaction, the eigenvalues of the matrix $\mathbf{Z}^T\mathbf{Z}$ and $\mathbf{Z}\mathbf{Z}^T$ follow the distribution of the eigenvalues of a Wishart matrix with parameters *I*-1 and *J*-1. This result was obtained by, amongst others, Johnson and Graybill (1972), who derived a likelihood ratio test for the largest eigenvalue. Elaborating on this result, Hegemann and Johnson (1976) derived a similar test for the second eigenvalue, conditional on the first eigenvalue. Further work on a sequential conditional testing procedure was done by Yochmowitz and Cornell (1978), Marasinghe (1985), and Schott (1986). No estimate for error is required for these tests, i.e. they apply to unreplicated tables. The test by Johnson and Graybill was recently shown to perform well by Williams and Wood (1993) and Cornelius (1993).

Mandel (1971) developed an F-test for testing the significance of eigenvalues, by converting eigenvalues to mean squares through division by an approximate number of degrees of freedom (*df*). Mandel assumed that in the absence of interaction, the eigenvalues should follow a Chi-square distribition, whose expectation should give the corresponding *df*. Mandel's approach works well for the first eigenvalue (Williams and Wood, 1993), but is less reliable for later eigenvalues. Gauch (1992) showed that in the presence of multiplicative interaction, Mandel's procedure attributes too many *df* to the earlier eigenvalues, so that it is too conservative. He also found that when strong pattern dominates noise, it is better to attribute *df* somewhere between equi-proportional and equivalent to the number of independent parameters corresponding to a multiplicative term. For an optimal (best) allocation of *df*, for every individual data set the idiosyncratic distribution of pattern and noise should be taken into account. Gauch describes a simulation strategy to find these optimal *df*.

Following Gollob (1968), the number of independent parameters corresponding to specific multiplicative terms is determined as follows. For multiplicative term *s* there have to be

estimated as parameters: I row scores, J column scores, and one singular value, i.e. $I+J+1$ parameters. Simultaneously, there are 2 sum-to-zero constraints, 2 unit length constraints, and $2(s-1)$ orthogonality constraints, i.e. $2+2s$ constraints. Thus, term s has $I+J-1-2s$ df. Gollob also proposed an F-test for testing eigenvalues, by calculating mean squares constructed from the quotients of eigenvalues and *df*. The mean squares have to be tested against an independent estimate of error. Williams and Wood (1993) and Cornelius (1993) have shown Gollob's test to be too liberal. However, under the alternative hypothesis of multiplicative interaction, Gollob's rule *does* provide adequate *df* (Goodman and Haberman, 1990; Gauch, 1992). Recently, Cornelius (1993) presented two new F-tests, which seem to perform better than Mandel's and Gollob's procedures, and are not limited in their applicability by restricted tables of critical values as are the likelihood ratio tests. Still, an independent estimate of error is necessary.

Summarizing, likelihood ratio tests seem reliable, and don't need replication, but have as a disadvantage that special tables are necessary, however these are available only for a limited number of table sizes. Mandel's approach works fine for the first eigenvalue, but an error estimate has to be available. Gauch's methods are reliable, but imply a lot of work. Cornelius' tests are easy to apply and perform well. Again an error term should be available. This requirement of an independent error may cause problems (Milliken and Johnson, 1989, p.3).

An estimate for the error could be obtained from pooling non-significant multiplicative terms. Significance may be assessed by a likelihood ratio test, degrees of freedom for significant terms can be derived from Gollob's rule. When the first, or the first two eigenvalues stand out very clearly from the later eigenvalues, and when one is sure that the pertinent two-way ANOVA interaction is significant, it is a reasonably safe strategy to attribute *df* to the first eigenvalue(s) by Gollob's rule, and pool the later eigenvalues as error. The F-test(s) against the constructed error may be used as a means of verification.

6.2 Three-way interaction

For testing multiplicative three-way interaction not very much theory has been developed. Boik (1990) presents a likelihood ratio test for the first term, including a table of critical values for comparatively small three-way tables. For testing the first three-way term in larger tables a conservative critical value may be constructed from a product of critical values for the first eigenvalue in two-way tables, as tabulated in Johnson and Graybill (1972), and Milliken and Johnson (1989). We may interpret a three-way decomposition with one component for each way as in some sense an optimal succession of two-way decompositions. For example, to a close approximation the first three-way term for a IxJxK three-way table can be obtained from the application of a two-way decomposition to the two-way rearrangement Ix(JxK), followed by a two-way decomposition of the first (JxK) component rearranged as a JxK two-way table. As an approximation to the critical value for the first three-way term we may take the product of the critical values for the Ix(JxK) decomposition

and the *JxK* decomposition. Because there are three possibilities for collapsing pairs of three-way modes, (*IxJ*), (*IxK*), or (*JxK*), there are three possible approximate critical values. Usually these three values are quite close together. We advise choosing the lowest of the three.

For testing the first three-way term in small tables, Boik (1990) can be used. For the first three-way term in larger tables our approximate test may be used. As an alternative to these tests for the first term, and as the only possibility for higher terms, we may use a procedure similar to one of the procedures for two-way tables, attributing *df* equal to the number of independent parameters to terms that stand out in the amount of three-way interaction described. This usually concerns only the first three-way component for each way. Occasionally second terms are involved. In section 4 the model degrees of freedom for a model with P row components, Q column components, and R layer components for an *IxJxK* three-way array **X** was given to be $(IxP+JxQ+KxR+PQR) - (P^2+Q^2+R^2)$. This applies when **X** contains raw data. When **X** consists of three-way interaction parameters an additional $P+Q+R$ sum-to-zero constraints should be taken into account (subtracted).

7. Biplots

Row and column scores as derived by two-way decompositions (SVDs) of two-way tables (matrices) can be displayed in biplots (Gabriel, 1971, 1981; Kempton, 1984). A biplot is a graphical display of a matrix **X** with I rows and J columns by means of coordinate vectors $\mathbf{a}_1, \mathbf{a}_2,...,\mathbf{a}_I$ for the rows, and $\mathbf{b}_1, \mathbf{b}_2,...,\mathbf{b}_J$ for the columns. The inner products between row and column coordinate vectors, $\mathbf{a}_i^T\mathbf{b}_j$, should represent (approximate) the matrix elements, x_{ij}. The row and column scores from a decomposition provide the basic material for the row and column coordinate vectors in the biplot.

To display the results of a three-way multiplicative decomposition with row, column, and layer scores, one may construct a biplot by combining scores from two of three ways and plot these against the third. In particular, for a *Genotype* x *Location* x *Year* three-way table with the typical element x_{gly},

$$x_{gly} = \sum_{p=1}^{P} a_{gp}\left[\sum_{q=1}^{Q}\sum_{r=1}^{R}\lambda_{pqr}b_{lq}c_{yr}\right] = \sum_{p=1}^{P} a_{gp}d_{(ly)p} = \sum_{p=1}^{P} a_{gp}d_{ep}\ , \qquad (18)$$

where the indices l and y are combined into one index, e. The biplot is a genotype-by-environment one, but with extra multiplicative structure imposed on the environments. The number of parameters is considerably lower than in a direct Px(QxR) two-dimensional biplot of $x_{g(ly)}$.

8. Application 1: Dutch Maize Variety Trials

8.1 Problem and data description

In the Netherlands, as in other countries, new varieties of field crops must officially be tested before permission for release is given. Inclusion of a new variety in the Descriptive List of Field Crops is a prerequisite for release. In the continuous testing program that accompanies the composition of the Descriptive List, new and existing varieties are evaluated on a number of locations over time. In the list of varieties tested, every year new varieties are entered and poorly performing existing varieties dropped. Inclusion in the list can be interpreted as an implicit, though conditional recommendation of a variety. One of the most important causes for the conditional element in the recommendation stems from the differential sensitivity of the varieties to certain changes in the environment. More accurate recommendations require deeper insight into the various mechanisms causing variety by environment interactions. In this regard the interaction models discussed above have been proved to be very useful. In the example below we show how three-way multiplicative decomposition provides important additional means for more parsimonious descriptions of three-way variety by environment interactions.

This first application involves dry matter content data from maize, collected within the official Dutch maize variety trials for the Descriptive List of Field Crops. The data used are given in Table 1. These data form a subset of the data analyzed in van Eeuwijk, Keizer and Bakker (1995). For experimental details and background information the reader is referred to this paper. The selected three-way table of (mean) percentages includes 6 varieties (*Brutus, Splenda, Markant, Vivia, Dorina, Irla*), which were planted at 4 sites in the Netherlands differing in soil and location (*Southern Sand, Central Sand, Northern Sand,* and *River Clay*), during 7 years (*1980, 1981, 1982, 1984, 1985, 1986,* and *1987).* In Table 1 and below self-explanatory abbreviations are used for varieties and environments (sites, years, or combinations of both). All calculations were done with the special purpose package *3Waypack* (Kroonenberg and Brouwer, 1993), which can be ordered from the second author of this paper.

8.2 Analysis; error estimates

An ever recurring problem in genotype by environment analyses is the settlement upon a defendable estimate for error. As a first guess, when available, the intra-block error may be taken. For our maize data, however, we did not have this error at our disposal. Therefore, error estimates were obtained from those parts of the two-way and three-way interactions that were not modeled. More specifically, we took part of the *Variety.Site* interaction as representing one estimate for error (0.58, 5 *df*, Table 2), and part of the *Variety.Site.Year* interaction as representing another estimate (0.47, 66 *df*, Table 2). Thus, effects were judged against an error of roughly 0.50.

As we evaded a description of two-way variety by environment interactions in terms of

Table 1.
Three-way table of dry matter contents (%) for maize.

Year	Variety / Site	Brutus	Splenda	Markant	Vivia	Dorina	Irla
1980	SS	30.25	28.72	27.75	29.30	28.09	28.71
	CS	27.16	28.80	26.69	26.49	25.95	26.91
	NS	30.01	34.67	27.99	29.83	28.03	30.89
	RC	29.94	32.36	29.12	28.57	29.71	29.23
1981	SS	29.91	32.12	28.47	30.77	23.05	30.35
	CS	25.88	26.35	25.36	26.57	26.55	26.79
	NS	27.39	30.25	28.47	28.89	26.95	29.24
	RC	34.81	36.37	34.07	35.02	33.65	35.01
1982	SS	34.12	32.48	32.71	32.39	32.75	31.50
	CS	31.92	33.07	30.83	33.18	32.09	31.92
	NS	26.82	28.60	28.15	26.98	26.88	26.08
	RC	35.16	34.00	33.94	33.53	34.31	36.30
1984	SS	29.14	29.81	27.73	28.89	27.75	29.51
	CS	21.70	20.82	19.49	20.74	20.41	21.14
	NS	19.93	19.73	18.96	19.88	19.15	18.61
	RC	28.33	28.08	25.91	29.20	28.43	30.41
1985	SS	34.25	37.67	33.74	34.01	32.92	35.02
	CS	26.37	25.97	23.69	25.72	25.29	25.55
	NS	20.21	20.13	20.37	20.63	20.72	20.69
	RC	27.94	28.54	26.63	27.90	27.56	29.12
1986	SS	28.90	27.71	26.25	30.18	28.49	29.31
	CS	26.97	26.71	24.61	28.54	26.48	27.87
	NS	21.73	20.31	19.71	22.62	21.77	22.39
	RC	26.52	26.94	26.09	28.03	26.40	28.09
1987	SS	26.97	26.47	23.23	27.46	25.27	27.52
	CS	27.26	25.66	22.60	26.99	25.43	25.63
	NS	20.14	19.79	19.50	21.95	20.02	21.12
	RC	23.32	23.71	20.58	25.09	23.41	24.09

components obtained from two-way SVDs, we could avoid assessment of the number of multiplicative terms to retain for these interactions. In general, one should use the two-way SVD mainly as an exploratory tool for finding more substantive models.

8.3 Analysis; main effects

Table 2 gives the three-way ANOVA plus a number of partitionings of the variation due to (additive) three-way ANOVA terms according to models discussed above. The effect of the environmental main effect (*Environment* = *Site* + *Year* + *Site.Year*) strongly dominated the varietal main effect. Total genotype by environment interaction (*Variety.Environment* = *Variety.Site* + *Variety.Year* + *Variety.Site.Year*) had a magnitude (Sum of Squares) about

Table 2.
Analysis of variance table for maize data.

Source	Degrees of Freedom	Sum of Squares	Mean of Squares
Variety	5	80.18	16.04
Environment	27	3008.21	111.42
Site	3	962.19	320.73
SS-NS	1	717.27	717.27
Year	6	1194.65	199.11
Site.Year	18	851.37	47.30
SVD11	8	505.29	63.16
SVD22	6	256.18	42.70
SVD33	4	89.90	22.48
Variety.Env	135	161.88	1.20
Var.Site	15	17.01	1.13
Conc	1	2.76	2.76
Regr on Site Mean	5	4.91	0.98
SVD11	7	8.83	1.26
SVD22	5	5.55	1.11
SVD33	3	2.63	0.88
Var.(SS-NS)	5	7.27	1.45
Var.((SS+NS)-(CS+RC))	5	6.84	1.37
Deviations	5	2.90	0.58
Var.Year	30	64.90	2.16
Conc	1	3.33	3.33
Regr on Year Mean	5	24.42	4.88
SVD11	10	42.28	4.23
SVD22	8	14.68	1.84
SVD33	6	6.44	1.07
Var.(Year Mean + Radiance)	10	46.54	4.65
Deviations	20	18.36	0.92
Var.Site.Year	90	79.97	0.89
Conc	1	1.43	1.43
Var.(Site Mean x Year Mean)	5	6.50	1.30
Var.(Site.Year Interaction)	5	8.06	1.61
3WD111	12	31.28	2.61
3WD222 given 3WD111	12	17.93	1.49
Deviations	66	30.76	0.47
Total	167	3520.26	

twice the varietal main effect, making it worthwhile studying this interaction in more detail.

8.4 Analysis; two-way interactions

Two-way interactions were partitioned by concurrence, regression on the mean, two-way SVD, contrasts, and factorial regression. As the *Site.Year* interaction was by far the largest, we first took a look at this term, decomposing it by means of SVD. The idea was that main trends in this comparably large interaction might prove to be useful for the description of other interactions. *Site* and *Year* component scores derived from the decomposition of the *Site.Year* interaction can be used as concomitant variables on the *Site* and *Year* factor in the description of the *Variety.Site*, *Variety.Year*, and *Variety.Site.Year* interaction, much in the same way as environmental main effects are used as concomitant variables in the two-way regression on the mean models for interaction.

The leading term of the *Site.Year* interaction (SVD11) was a *Site* constrast between *SS* and *NS* times a *Year* contrast between *80+81* on the one hand and *85* on the other hand (Table 3). The *Site* contrast *SS-NS* was also found to describe an important part of the *Site* main effect (75%, Table 2). Furthermore, the same contrast accounted for a substantial amount of the variation of the *Variety.Site* interaction (43%, see Table 2 and below). The second multiplicative term of the SVD of the *Site.Year* interaction (SVD22) was less interesting for providing concomitant variables for other additive ANOVA terms. The third multiplicative term (SVD33) represented the *Site* contrast (*SS+NS*)-(*CS+RC*) times the *Year* contrast (*80-82*). This third *Site* contrast again did very well as concomitant variable for the *Variety.Site* interaction (see Table 2 and below).

The *Variety.Site* interaction could be described satisfactorily by contrasts derived from the first and third *Site* components obtained from decomposition of the *Site.Year* interaction. In fact, these contrasts were preferable to the contrasts derived from the SVD of the *Variety.Site* interaction itself. One argument was that the residual mean square for the deviations from the model with the *Site* contrasts, as derived from the *Site.Year* interaction, for the *Variety.Site* interaction (0.58) was very close to the error estimate derived from the deviations from the chosen three-way model for the three-way interaction (0.47, Table 2). Not surprisingly after the disappointing performance of the *Variety.Site* SVD, concurrence and regression on the mean did not succeed in giving a satisfactory description of the *Variety.Site* interaction either (Table 2).

For the description of the *Variety.Year* interaction the multiplicative terms from the *Site.Year* interaction were less useful. SVD of the *Variety.Year* interaction revealed that the leading *Year* component was related to the *Year* main effect (Pearson correlation coefficient = -0.73), an indication of the appropriateness of a regression on the mean model. As we had available extra climatological covariates for the years (see van Eeuwijk, Keizer and Bakker, 1995), we were able to find an adequate model for the *Variety.Year* interaction as a factorial regression model with the concomitant variables *Year mean* (main effect) and *radiance*

Table 3.
Component scores of two-way decomposition of Site.Year interaction for maize data. Components have unit length.

Site				Year			
	1	2	3		1	2	3
SS	0.72	-0.20	0.45	80	-0.53	0.17	0.63
CS	0.13	0.74	-0.44	81	-0.49	-0.44	-0.21
NS	-0.66	0.10	0.55	82	-0.03	0.12	-0.61
RC	-0.18	-0.64	-0.55	84	0.20	-0.57	0.02
				85	0.64	-0.15	0.37
				86	0.14	0.30	-0.23
				87	0.08	0.57	0.03

(Table 2). (One might also interpret this model to be a mix of regression on the environmental mean and factorial regression.) Concurrence was again ineffective.

Summarizing, we described the two-way variety by environment interactions in terms of regressions (regression on the mean and factorial regression) and contrasts.

8.5 Analysis; three-way interaction

For description of the three-way interaction we decided to retain two components for each of the modes *Variety, Site,* and *Year* (Table 2). For retention of the first components we could use the critical value for λ^2_{111}/{Sum of squares due to interaction} as tabulated in Boik (1990). The observed value of 0.40 was clearly above the tabulated critical value of 0.33 for a (6-1)x(4-1)x(7-1) three-way table ($P \leq 0.05$). (Our approximate critical value from two successive two-way decompositions was 0.35.) The second component for each way was retained on basis of the still substantial contribution to the description of the three-way interaction (22%), symmetry, and the size of the error estimate being comparable to the estimate from the *Variety.Site* estimate.

Subsequent three-way decompositions with regard to dimensionality are not nested in an orthogonal sense. Estimates of first components in a model with only one component for each way will differ from the estimates for the first components in a model with two components for each way. For that reason the amount of variation described by the inclusion of an extra component must be expressed with reference to a hierarchically simpler model. The gain of having two components per way instead of one is expressed in Table 2 by 3WD222 given 3WD111. The component scores corresponding to the 3WD222 solution are given in Table 4. The first *Variety* component represents the contrast *Do-Sp*. The second component stands for (*Sp+Do*)-(*Br+Vi*). The *Site* components represent the contrasts (*CS+NS*)-(2x*SS*) and *CS-NS*, respectively. For the years we find *81-80* and (*80+81*)-(*84+85+86+87*). The core-matrix elements (Table 5) suggest the prevailing importance of *Variety-1* x *Site-1* x *Year-1*, *Variety-1* x *Site-2* x *Year-2*, and *Variety-2* x *Site-1* x *Year-2*. The squares of the values of the core

Table 4.
Scores (unit length) for the first two multiplicative terms of the three-way decomposition of the three-way ANOVA interaction for maize data.

Variety			Site			Year		
	1	2		1	2		1	2
Br	0.09	-0.47	SS	-0.84	0.12	80	-0.61	0.47
Sp	-0.70	0.48	CS	0.32	-0.59	81	0.59	0.67
Ma	0.14	-0.12	NS	0.42	0.75	82	-0.33	0.05
Vi	-0.03	-0.46	RC	0.09	-0.28	84	0.07	-0.20
Do	0.67	0.57				85	0.38	-0.39
Ir	-0.18	-0.00				86	-0.14	-0.31
						87	0.04	-0.20

matrix are equal to the Sum of Squares explained by a particular combination of components. For example, the first components of *Variety*, *Site*, and *Year* account for $5.547^2 = 30.77$, or 38% of the total three-way interaction. The interpretation of this component is that conditional on the model including all main effects and two-way interactions, there was a difference between *Do* and *Sp*, which depended on the site, (*CS+NS*)-(2x*SS*), and the year, *81-80*. An interpretation using the signs of the component scores would lead us to conclude that there was positive interaction for *Sp* coupled with negative interaction for *Do* in the environments *SS.81* and (*CS+NS*).*80*, and negative interaction for *Sp* coupled with positive interaction for *Do* in *SS.80* and (*CS+NS*).*81*. However, as we concluded two components to be present for each of the modes, an interpretation focussing on individual component scores may be misleading. Therefore we prefer an interpretation that acknowledges the full dimensionality of the three-way interaction. As an intermediate instrument we use the biplot. The biplot in Figure 1 can best be thought of as being equivalent to a biplot display for two-way interaction with the proviso that the scores have been restricted to accommodate the three-way multiplicative structure of the Tucker three-way model. The latter does not affect the interpretational rules for the biplot. Figure 1 shows more detail than inspection of individual combinations of components could have revealed. We see that *Sp* had positive three-way interaction in *NS.80*, *SS.81*, and *SS.85*, against negative interaction in *NS.85* and *SS.80*. *Do* performed less well than expected on the basis of a model including main effects and two-way interactions in *SS.81*, but better in *CS.81*. *Vi* and *Do* did better than expected in *SS.80* and *SS.81*, and worse in *SS.85*, *SS.86*, *NS.80*, and *CS.81*. The biplot thus provides useful information for the plant breeder, because it identifies a number of specific environmental conditions in which some genotypes exhibit responses which cannot be fitted well by a model containing only two-way interaction terms. Further consideration of the genotype by environment combinations responsible for the three-way interaction may uncover very specific resistance or stress mechanisms.

In contrast to the two-way genotype by environment interactions, the three-way genotype

Table 5.
Values of the elements in the core-matrix for maize data. These are indicators of the importance of the combinations of multiplicative terms.

	Variety 1		Variety 2	
	Year 1	Year 2	Year 1	Year 2
Site 1	5.547	0.222	-0.216	3.088
Site 2	0.215	-2.933	-0.346	-0.200

by environment interaction had a contrast/outlier structure, certainly no regression structure. The attempts to fit regression models to the three-way interaction, as three-way concurrence, regression on the (environmental) mean, or regression on the (environmental) interaction, all failed (Table 2).

9. Application 2: Fusarium in wheat

9.1 Problem and data description

In resistance breeding it is important to know to what extent varieties show differential sensitivity to different strains of a pathogen, i.e. what is the specificity of the *biological interaction* between host and pathogen. To that purpose varieties are artificially inoculated (infected) at the beginning of the growing season with a number of strains of the pathogen. Usually these experiments are repeated in time. In the analyses principal attention is given to variety by strain interaction, and variety by strain by year interaction. Depending on the pattern of *biological interaction* to be expected, various *statistical* models for interaction are candidates to replace the little parsimonious additive ANOVA interaction. Though there may exist more parsimonious models, the multiplicative models following from two- and three-way multiplicative decompositions of the additive two- and three-way ANOVA interactions certainly form a good first approximation to the *biological* interaction structure, and thus should always be inspected.

For our second application we used resistance breeding data that came from Hungarian evaluations of 20 wheat varieties infected with 7 strains of *Fusarium* over the years 1990, 1991, 1992, and 1993. The variable that was analyzed, was severity of disease incidence due to *Fusarium* head blight. Data were recorded as ratings, and analyzed as logits. The data covering the years 1990, 1991, and 1992 were analyzed earlier as part of an international study on resistance of wheat to *Fusarium* (van Eeuwijk *et al.*, 1995). Experimental details and more information about varieties and strains can be found there, as well as the full names corresponding to the abbreviations used for varieties and strains (Tables 1 and 2 in van Eeuwijk *et al.*, 1995, respectively). In the abbreviations of Tables 7 and 8 the first letters of variety and strain codes refer to country of origin (N = Netherlands, F = France, G =

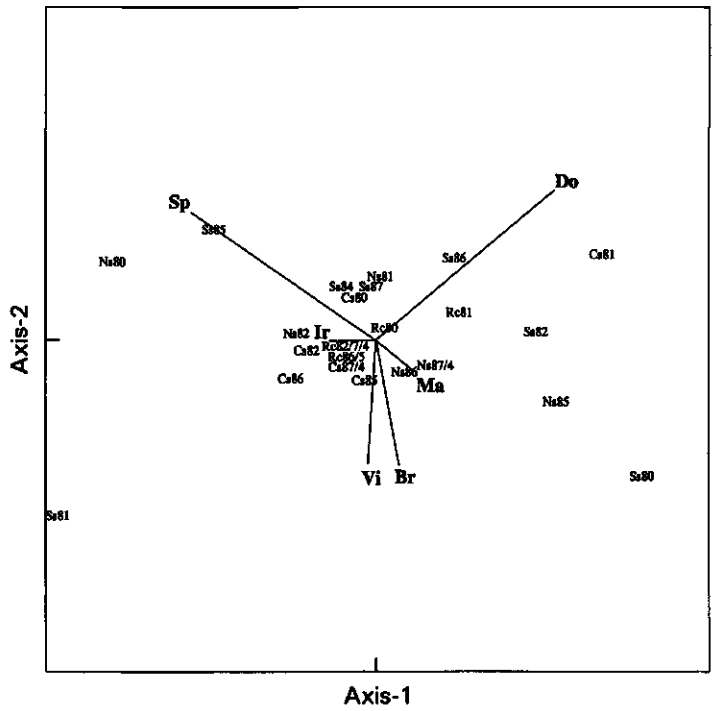


Figure 1. Biplot of three-way interaction for maize data, varieties are bold and large, environments consist of site-year combinations. Codes refer to Table 1 (slashes are used to indicate environments with a change in year only for a particular site). Axis-1 and axis-2 refer to scores on first and second component.

Germany, H = Hungary). The last letter of the strain codes distinguishes *Fusarium culmorum* strains (C) from *Fusarium graminearum* strains (G).

9.2 Analysis

Table 6 contains the analysis of variance table for the wheat data. We see that the *Variety.Strain* interaction is small (Sum of Squares) in comparison to other terms, of which the most relevant ones are the *Variety* main effect and the *Variety.Strain.Year* interaction, a rough indication of the low incidence of *biologically* specific host-pathogen interactions. Nevertheless, we decomposed the *Variety.Strain* interaction by two-way SVD to see whether any structure was to be discerned. The amounts of variation described by the successive multiplicative terms were 53%, 17%, 13%, 11%, 5%, and 2%. The first term clearly stands out, and was found significant ($P \leq 0.01$) by the likelihood ratio test of Johnson and Graybill (1972). Later terms were pooled to give an estimate for error of 0.18 (Table 6). Confidence intervals for scores (Table 7) were calculated using the formulae given by Goodman and Haberman (1990) and Chadoeuf and Denis (1991). The bounds were calculated quite roughly,

Table 6.
Analysis of variance table for wheat data.

Source	Degrees of Freedom	Sum of Squares	Mean Square
Variety	19	152.46	8.02
Strain	6	91.26	15.21
Year	3	321.00	107.00
Variety.Strain	114	33.79	0.30
SVD11	24	17.65	0.74
Deviations	90	16.14	0.18
Variety.Year	57	137.74	2.42
Strain.Year	18	232.07	12.89
Var.Strain.Year	342	93.00	0.27
SVD111	26	29.72	1.14
Deviations	316	63.27	0.19
Total	559	1061.31	

with $\lambda^2 = 17.65$, $\sigma^2 = 0.2$, and $t = 2$. Their main function was to see whether 0 was included in the interval, in which case no further attention had to be given to the particular variety or strain for the description of *Variety.Strain* interaction. Only 4 out of the 20 varieties were found to have non-zero scores: G47/83, HZombr, HBence, and HSgv/G. Of the 7 strains, the following 4 had non-zero scores: FLeR4C, H216-G, H377-G, G207-G. Combinations of these varieties and strains form candidates for specific host-pathogen interactions. Thus, three Hungarian varieties exhibited negative interaction with respect to two Hungarian strains (negative products of scores), the latter two both of the *Fusarium graminearum* species, i.e. they were relatively weakly infected. This could be a consequence of positive selection for resistance of Hungarian wheat lines against Hungarian strains of *Fusarium*. The same Hungarian lines were relatively heavily infected by the French strain FLeR4C, and the German strain G207-G. The German variety G47/83 did relatively well with respect to the German strain G207-G. This might again be an expression of selection for resistance.

Three-way decomposition of the *Variety.Strain.Year* interaction led us to retain only the first multiplicative term, λ^2_{111}/{Sum of squares due to three-way interaction} was 0.320, which was clearly above the nearest tabulated value of Boik (1990) of 0.270 (tabulated 9x5x3 table, needed (20-1)x(7-1)x(4-1) table, our two times two-way approximation was about 0.18). Further terms did not add substantially to the description of the three-way interaction, and were pooled to give another estimate for error of 0.19, about the same as the estimate obtained

Table 7.

Variety and strain scores (unit length) with lower and upper bounds ($P \leq 0.95$), for two-way Variety by Strain interaction.

Variety	Lower Bound	Score	Upper Bound
NSVP75	-0.07	0.14	0.34
NArina	-0.28	-0.07	0.14
N72005	0.00	0.20	0.41
N72017	-0.14	0.07	0.28
N75079	-0.10	0.11	0.32
FCopai	-0.17	0.03	0.24
FRescl	-0.01	0.20	0.40
FRC103	-0.12	0.09	0.29
F82/F3	-0.10	0.10	0.31
F81/F3	-0.28	-0.07	0.13
G25/83	-0.19	0.02	0.23
G47/83	0.06	0.26	0.46
G77/82	-0.07	0.13	0.34
G16381	-0.06	0.14	0.35
G20481	-0.23	-0.02	0.18
HZombr	-0.82	-0.67	-0.51
HSzoke	-0.16	0.05	0.26
HBence	-0.62	-0.43	-0.25
H85/92	-0.15	0.05	0.26
HSgv/G	-0.53	-0.33	-0.14

Strain	Lower Bound	Score	Upper Bound
FLeR4C	-0.44	-0.25	-0.06
H216-G	0.17	0.35	0.53
H377-G	0.47	0.62	0.77
H375-G	-0.25	-0.05	0.15
H551-C	-0.11	0.08	0.28
G223-C	-0.31	-0.12	0.08
G207-G	-0.78	-0.64	-0.50

from the part of the *Variety.Site* interaction that was not modeled (0.18). The extracted multiplicative three-way term was interpreted by the values of the scores (Table 8). The *Strain* component was a contrast between foreign (FLeR4C, G223-C, G207-G) and domestic strains (especially H377-G and H375-G). The *Year* component represented a contrast between 1992 and the other years. The varieties most sensitive to this strain by year pattern of environmental interaction were the very resistant Dutch variety Arina (coded NArina) and the two Hungarian varieties Zombre (coded HZombr) and Bence (coded HBence). Conditional upon a model including main effects and two-way interactions, these three varieties showed very low infection incidences in 1992, relative to the other years, with respect to the Hungarian strains, when contrasted with the foreign strains. This three-way structure could be

Table 8.

Component scores (unit length) for the three-way decomposition of three-way interaction for wheat data.

Variety		Strain		Year	
NSVP75	0.09	FLeR4C	-0.36	1990	-0.26
NArina	-0.45	H216-G	0.08	1991	-0.21
N72005	0.07	H377-G	0.56	1992	0.86
N72017	0.12	H375-G	0.45	1993	-0.39
N75079	0.25	H551-C	0.10		
FCopai	-0.05	G223-C	-0.46		
FRescl	0.13	G207-G	-0.35		
FRC103	0.16				
F82-F3	-0.01				
F81-F3	-0.27				
G25/83	0.25				
G47/83	0.20				
G77/82	0.11				
G16381	0.24				
G20481	0.25				
HZombr	-0.40				
HSzoke	-0.03				
HBence	-0.40				
H85-92	-0.08				
HSgv/G	-0.15				

understood quite well, as it turned out that due to a disturbance in the storage of the inoculae of the foreign strains in 1992, these strains were unable to cause infection in the field. (inoculum = disease causing solution containing, in this example, a specific *Fusarium* strain, that is sprayed over the plants). The varieties responding most clearly to the environmental (strain by year) structure determining the three-way interaction, were varieties that were neither infected by foreign strains nor domestic strains in 1992.

The use of a three-way multiplicative decomposition identified a disturbance in the experiment that had nothing to do with the resistance mechanisms that were of interest. Therefore, the biological interpretation of the experiment could be limited to the model with two-way interactions, more specifically to the interpretation of the two-way SVD of the *Variety.Strain* interaction, which proved to be straightforward and interesting. This is an important result because a similar unequivocal interpretation will be hard to come by with a conventional genotype by environment two-way analysis of the data in which the environmental dimension would have been compounded of the *Strain* and *Year* factor.

10. Discussion

10.1 Statistical considerations

An important assumption underlying practically all of the methodology and reasoning presented in this paper is that the three-way ANOVA model is in principle a reasonable model for the data to be analyzed. First a three-way ANOVA model is fitted, and subsequently individual two- and three-way ANOVA interactions that are represented by additive parameters in the ANOVA context are replaced by multiplicative formulations that are thought to either represent the biological reality more closely or otherwise lead to a more insightful interpretation. Two types of objections may be made. Firstly, our linear predictor scale may not be the identity scale, while our error distribution may not be normal. The answer to this objection would require extension of the presented methodology to the framework of the generalized linear models. In principle this would mean no insurmountable complications. In van Eeuwijk (1995), two-way decompositions of interactions are generalized to the class of generalized linear models, using the property of bi-additivity of the two-way interaction for constructing an alternating generalized regression algorithm for estimation of the interaction parameters. Elaborating the three-way algorithm based on quadri-additivity of Weesie and van Houwelingen (1983), a similar generalization of three-way methodology to the class of generalized linear models, seems straightforward.

A second objection involves the necessity of all additive ANOVA terms for adequate prediction of responses. For example, in the three-way generalization of the SHMM (16), the ANOVA structure is replaced by a sum of an intercept term and one or more multiplicative terms. There may be instances where a SHMM formulation to describe the observed variation in the biological response will be more parsimonious than an ANOVA formulation supplemented by multiplicative formulations for the interactions. We feel, however, that this gain in parsimony in general cannot compensate for the loss in interpretability. Furthermore, the difference between the models is for a great deal illusory, because imposition of sum-to-zero constraints will induce two-way interactions and main effects (Gower, 1977). Even when data were generated by a completely multiplicative model, fitting of an additive ANOVA model would reveal that a very large proportion of the variation is captured by the main effects. Therefore, for non-pathological data, the ANOVA model (including multiplicative interactions) can be expected to perform quite well, even when it is not the most parsimonious model.

10.2 Biological considerations related to modeling GEI.

Basing ourselves on subject matter considerations, we think that for the *two-way* situation it is plausible that GEI takes the form of differential sensitivity to environmental factors (Federer and Scully, 1993; van Eeuwijk, Denis and Kang, 1995). The concurrence model, as a model for biological interaction, is suspect, as it is more related to a test for transformation than to a serious model (Atkinson, 1982). The regression on the mean model is a very useful model as a starting model for investigation of GEI. However, it should always be followed by

genotypes and environments. *Heredity* **30,** 97-109.

Gabriel, K.R. (1971). Biplot display of multivariate matrices with application to principal component analysis. *Biometrika* **58,** 453-467.

Gabriel, K.R. (1981). Biplot display of multivariate matrices for inspection of data and diagnosis. In: *Interpreting multivariate data,* V. Barnett (ed.), 147-173. Chicester: Wiley.

Gabriel, K.R., and Zamir, S. (1979). Lower rank approximations of matrices by least squares with any choice of weights. *Technometrics***21,** 489-498.

Gauch, H.G. Jr. (1988). Model selection and validation for yield trials with interaction. *Biometrics* **44,** 705-715.

Gauch, H.G. Jr. (1992). *Statistical analysis of regional yield trials: AMMI analysis of factorial designs.* Amsterdam: Elsevier.

Gollob, H.F. (1968). A statistical model which combines features of factor analysis and analysis of variance techniques. *Psychometrika* **33,** 73-115.

Goodman, L.A. and Haberman, S.J. (1990). The analysis of non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **85,** 139-145.

Gower, J.C. (1977). The analysis of three-way grids. In: *Dimensions of intra personal space,* Vol. 2: *The measurement of intra personal space by grid technique.* Ed. P. Slater. Chicester: Wiley. 163-173.

Gower, J.C. (1995). Comments on J.A. Nelder 'The statistics of linear models: back to basics'. *Statistics and Computing* **5,** 91-92.

Harter, H.L., and Lum M.D. (1962). An interpretation and extension of Tukey's one degree of freedom for non-additivity. *Technical Report* 62-313, Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Hegemann, V.J., and Johnson, D.E., (1976). On analyzing two-way AOV data with interaction. *Technometrics* **18,** 862-868.

Johnson, D.E. and Graybill, F.A. (1972). An analysis of a two-way model with interaction and no replication. *Journal of the American Statistical Society* **67,** 862-868.

Kempton, R.A. (1984). The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science, Cambridge* **103,** 123-135.

Kiers, H.A.L., Kroonenberg P.M., and Ten Berge, J.M.F. (1992). An efficient algorithm for TUCKALS3 on data with large numbers of observational units. *Psychometrika* **57,** 415-422.

Kroonenberg, P.M. (1983). *Three-mode principal components analysis. Theory and applications.* Leiden: DSWO-Press.

Kroonenberg, P.M., and Brouwer, P. (1993). *3Waypack user's manual.* Department of Education, Leiden University.

Kroonenberg, P.M., and De Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika* **45,** 69-97.

Lin, C.S., and Binns, M.R. (1988). A method of analyzing cultivar x location x year experiments: a new stability parameter. *Theoretical and Applied Genetics* **76,** 425-430.

Mandel, J. (1961). Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **56,** 878-888.

Mandel, J. (1969). The partitioning of interaction in analysis of variance. *Journal of Research of the National Bureau of Standards - B, Mathematical Sciences* **73,** 309-328.

Mandel, J. (1971). A new analysis of variance model for non-additive data. *Technometrics* **13,** 1-18.

Marasinghe, M.G. (1985). Asymptotic tests and Monte Carlo studies associated with the multiplicative interaction model. *Communications in Statistics A - Theory and Methods* **14,** 2219-2231.

Milliken, G.A., and Johnson, D.E. (1989). *Analysis of messy data.* Vol. 2: *Nonreplicated experiments.* New York: Van Nostrand Reinhold.

Rao, C.R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhya A* **26,** 329-358.

Seyedsadr, M., and Cornelius, P.L. (1992). Shifted multiplicative models for non-additive two-way tables. *Communications in Statistics B- Simulation and Computation* **21,** 807-832.

Schott, J.R. (1986). A note on the critical values in stepwise tests for multiplicative components of interaction. *Communications in Statistics A- Theory and Methods* **15,** 1561-1570.

Schutz, W.M., and Bernard, R.L. (1967). Genotype x environment interactions in the regional testing of soybean strains. *Crop Science* **7,** 125-130.

Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* **37,** 3-27.

Tukey, J.W. (1949). One degree of freedom for non-additivity. *Biometrics* **5,** 232-242.

van Eeuwijk, F.A. (1992). Interpreting genotype-by-environment interaction using redundancy analysis. *Theoretical and Applied Genetics* **85,** 89-100.

van Eeuwijk, F.A. (1995). Multiplicative interaction in generalized linear models. *Biometrics* **51,** 1017-1032.

van Eeuwijk, F.A., Denis, J.-B., and Kang, M.S. (1995). Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: *Genotype by environment interaction: New Perspectives,* M.S. Kang and H.G. Gauch Jr. (eds.), 15-49. Boca Raton: CRC-Press.

van Eeuwijk, F.A., Mesterhazy, A., Kling, Ch. I., Ruckenbauer, P, Saur, L., Bürstmayr, H., Lemmens, M., Keizer, L.C.P., Maurin, N., and Snijders, C.H.A. (1995). Assessing non-specificity of resistance in wheat to head blight caused by inoculation with European strains of *Fusarium culmorum, F. graminearum* and *F. nivale,* using a multiplicative model for interaction. *Theoretical and Applied Genetics* **90,** 221-228.

van Eeuwijk, F.A., Keizer, L.C.P, and Bakker, J.J. (1995). Linear and bilinear models for the analysis of genotype by environment interaction, with an application to data from the Dutch Maize Variety Trials. *Euphytica* **84,** 9-22.

Weesie, J., and van Houwelingen, J.H. (1983). *GEPCAM user's manual: Generalized principal components analysis with missing values.* Technical report, Institute Mathematical Statistics, University of Utrecht.

Williams, E.R., and Wood, J.T. (1993). Testing the significance of genotype-environment interaction. *Australian Journal of Statistics* **35,** 359-362.

Yates, F., and Cochran, W.G. (1938). The analysis of groups of experiments. *Journal of Agricultural Science, Cambridge* **28,** 556-580.

Yochmowitz, M.G., and Cornell, R.G. (1978). Stepwise tests for multiplicative components of interaction. *Technometrics* **20,** 79-84.

Chapter XVI

Between and beyond addivity and non-additivity; the statistical modelling of genotype by environment interaction in plant breeding

F.A. van Eeuwijk
DLO - Centre for Plant Breeding and Reproduction Research (CPRO-DLO)
P.O.Box 16
6700 AA Wageningen
The Netherlands

Between and beyond additivity and non-additivity; the statistical modelling of genotype by environment interaction in plant breeding

1 Introduction

In the previous chapters many instances of genotype by environment interaction (GEI) have been presented. One way to classify these instances is by the crop from which the data were obtained; white cabbage (chapter II), sugar beet (chapter III), perennial ryegrass (chapter IV), lettuce (chapter V), wheat (chapters VI, XII, XIII and XV), maize (chapters XI and XV), and potato (chapter XII). Another way of classification can be based on the type of variable that was analysed; field emergence (II), yield and quality parameters (III), seed yield (IV), nitrate concentration (V), *Fusarium* head blight incidence (VI, XII, XIII, XV), dry matter content (XI, XV), number of potato cyst nematodes (XII). A third way of classification may concern the discipline in which the experiment was carried out; 'classical plant breeding' (IV, V), resistance breeding (III, VI, XII, XIII, XV), value for cultivation and use research (XI), and seed technology research (II). Then the model used for analysis may determine classification; analysis of variance models (all applied chapters), factorial regression models (II, III, IV, V, XI, XV), bilinear models (II, III, IV, V, VI, XI, XII, XIII, XV), reduced rank factorial regression models (VI, XI), generalized bilinear models (XII), quadrilinear models (XV), and mixed models (XI). The list of classifying factors could easily be extended to, for example, the identity of the environmental factors (locations, years, pathotypes, and combinations of the former), or, the type of concomitant information available for genotypes and environments, etc., etc..

An important and basic question that arises from the consideration of these attempts to classify instances of GEI is whether the meaning of GEI is the same for all those instances, and, if no, which are the factors that determine what GEI means in a particular instance. Is GEI in maize (inbred lines and hybrids) different from GEI in potato (vegetatively propogated lines) or GEI in perennial ryegrass (synthetic varieties)? Is GEI for yield characteristics different from GEI for quality characteristics? Is GEI for yield (end of development) different from GEI in field emergence (begin of development). Is GEI in resistance breeding trials searching for vertical resistances different from GEI in variety trials assessing value for cultivation and use? Is cultivar times location interaction different from cultivar times year interaction? Is GEI as described by bilinear models different from GEI as described by factorial regression models?

Are the questions of the former alinea sufficiently specific? Should not questions be asked of the form: Is GEI for percentage of sugar in sugar beet different from GEI for potassium concentration in sugar beet? Or, should it be: Is GEI for yield in sugar beet different from GEI for yield in maize? Or: Is the bilinear representation of the cultivar by location interaction for seed yield in perennial ryegrass different from the representation by factorial regression for the same interaction?

All these questions drive us automatically to *the* question of: *What is it that we are talking about when we talk about GEI*? For answering that question we have to study models that contain representations of GEI; entities, structures, processes, terms, parameters, etc.. A division could be made between biological (plant breeding/ physiology/ genetics) and statistical models containing representations of GEI. Whether the structure of biological models is necessarily very different from the structure of statistical models is doubtful, but generally it will be clear which types of models may be called biological and which statistical. Biological are those models often used by biologists of which they say that they are biological, and whose structural features can be described in biological language. Statistical models are also used by biologists, but they will not try to attribute to every structural part of a statistical model a biological meaning. This is not to say, that it would not be worthwhile to have a fully biological interpretation of a statistical model, i.e. change the statistical model into a biological model. However, for many purposes it does not pay off to develop a fully biological model, as a statistical model provides sufficient control over the biological process or phenomenon of interest.

The major aim of this chapter is to elucidate what GEI means in various models and contexts, and how these models and contexts relate to each other, a conceptual clarification. Both biological and statistical models should give adequate descriptions of the data at hand. However, biological models are granted a more robust character, a larger range of applicability, because they are assumed to have been verified at a larger number of independent occasions. Statistical models are more prone to optimize local fit. This assumption of wider applicability of biological models should be verified, because it may still be false. Nevertheless, the position taken is that for the construction of statistical models for GEI it can be advantageous to take biological models as guidelines. Thus obtained statistical models may be valid over a wider range of conditions.

The composition of the rest of the chapter is as follows. In section 2, first some general ideas will be posited on what models are and what purposes they should serve. Then some terminology is introduced that should help in relating statistical models to biological models and vice versa. Section 3 presents a summary of biological models describing GEI. This section ends with a rephrasal of biological models in a form amenable to statistical modelling. Section 4, being the body of this chapter, contains a discussion of the meaning of GEI in various statistical models and how statistical GEI relates to biological GEI. This section culminates in the description of a strategy for the statistical modelling of GEI in plant

breeding. Section 5 shows a glimpse of the horizon, coming nearer and becoming increasingly bright.

2 Models and relations between them

2.1 Models and embedment

A model is a structure or body of language that makes statements about observable and non-observable phenomena. That what should decide upon the appropriateness of a model is the agreement of its observable consequences with data (van Fraassen, 1980, 1989). That a model can be wrong with regard to its non-observable consequences is rather immaterial. What we cannot observe, we cannot decide on. Models are not required to give a 'truthful representation of reality'. They have to fit adequately and are then accepted, without it being necessary for them to represent any reality.

A central notion with respect to the interpretation of model structures and for relating different models to each other, e.g. statistical and biological models, is the notion of *embedment*. Embedment applies when parts or the whole of one model can be identified (interpreted) as a substructure of another model. The discussion then centres on the justifiability of the conjectured similarities (isomorphisms, van Fraassen, 1980; homomorphisms, Lloyd, 1987).

Good models provide interesting predictions, i.e. the right functional values for other configurations of data than those used for constructing the model. Furthermore, good models direct future research, for example by suggesting new embedment relations. Generally, models become more interesting when adequate fit and prediction are maintained with increasing parsimony.

2.2 Explanation, description and prediction

Embedment is a very useful concept to define relations between models. A substantial part of scientific activity consists in the search for justifications and refutations of proposed embedments. As data cannot be separated from a model of observation (selection of data, use of specialized monitoring equipment, etc.), even the fitting of statistical models to data already involves an act of embedment.

Other concepts whose meaning at least partly coincides with that of embedment are 'explanation' and 'description'. Embedment of one model within another can count as an act of explanation, while a description of a phenomenon using particular terms can be seen to imply an act of embedment. Embedment is a preferable notion over explanation because it is less value-laden, less charged with unwanted connotations, and less context-dependent. In this section a brief argument is given to show how explanation is a context-dependent phenomenon, which is not essentially different from description. Also the related notion of prediction will be discussed.

When characterizing models and their relations with other models and data, it is common

to classify them in terms of descriptive, predictive and explanatory models, with merit usually increasing from descriptive to explanatory. Without further qualification, however, it is rather obscure what should distinguish, for example, explanatory models from descriptive models. Are statistical model descriptive and biological models explanatory?

Explanations are often assumed to tell us something more profound than descriptions, and in a restricted sense this can be true. According to van Fraassen (1980) asking for explanations is asking for answers to questions of the type; Why X, rather than Y? An example: Why does genotype A benefit more from increasing nitrogen availability than other genotypes? The same why-question can ask for different kinds of information in different contexts. For example, two answers to the above why-question are; (1) Genotype A outcompetes others, because it has the highest positive regression coefficient with respect to the chosen univariate description of the environmental conditions (available nitrogen); (2) Genotype A has a physiology that makes it utilize available nitrogen best. Would we think of (1) as a description and of (2) as an explanation?

The collection of alternatives from which the observed state of affairs is a member, is called the contrast-class (van Fraassen, 1980), or reference class (Salmon, 1984). For the first case, the contrast-class, rather trivially, contains the range of possible values for the regression coefficients in the particular regression model. For the second case, the contrast-class involves alternative physiological mechanisms. That for which an explanation is wanted is called the topic. The answer (explanation) to a why-question (the request for an explanation) must have the right relevance-relation to the topic and the contrast-class, it must give a reason why exactly this state of affairs was observed and not another one from the contrast-class.

Are explanations essentially different from descriptions? No, explanations are just adequate and informative descriptions. The difference resides, if anywhere, in the contrast classes. Scientific explanations are scientific for no other reason than that they are formulated in scientific language, they represent a choice from a particular contrast class which should represent the actual state of knowledge in the particular discipline.

Just as the difference between explanation and description is a context dependent phenomenon, so is the difference between prediction and description only intelligible within an explicitly defined context. In a statistical context different loss functions may have been defined for each of them. It is then imaginable that a model combines good descriptive qualities with poor predictive qualities, although this would cast doubt on the adequacy of the chosen loss functions.

A useful classification of predictions that we will use in the remainder is that between interpolatory predictions and extrapolatory predictions (Ehrenberg and Bound, 1993). Interpolations concern predictions for combinations of predictor variables that are judged to be elements of the same set of configurations as that used for actually fitting the model. Extrapolations refer to configurations outside the modelling set.

2.3 Embedment, causation and guidance

Embedding models in more elaborate models is sometimes taken to accompany the transition from statistical *association* to *causation* (Fisher as reported by Cochran, cited in Cox, 1992). Causation is an important notion with regard to the generalizability of associations. Causation suggests unconditional association, high predictability, and global validity (robustness). Embedding statistical models in more elaborate biological models may thus lead to more robust models combined with higher predictive power. Furthermore, embedment can provide guidance in the construction of statistical models.

It is not necessary to believe that something essential changes going from association to causation (Occam's razor). Causation is best stripped of any metaphysical connotations. One argument to do this was given by van Fraassen (1980). Accepting a model and assigning truth values to statements entailed by the model, logical relations follow from the relations among the truth values. Thus, causality (causation) is a derived concept, understandable only within the confines of a model. About its existence elsewhere nothing sensible can be said. However, there is no need to do so either.

Assuming a continuum, causation can be interpreted as an ideal, unconditional, association. It is valuable to know that a particular predictor variable always occurs in the prediction equation for a particular response variable independent of the other predictor variables included (and independent of the selected modelling data). Unconditional association might form a kind of necessary condition for the attribution of the predicate 'causal' (Cox, 1992). Qualification is required with regard to the ordering of the variables. For something to be a cause it should precede its alleged effect, while being posterior to other candidate causes. Instead of temporal proximity, also spatial proximity can be used as the primary ordering device. Both temporal and spatial ordering requirements can be comprised within orders on the basis of subject matter knowledge.

An association that is causal, unconditional, should have high predictive power. If *all* relevant predictors are included in the *right* functional form, the quality of the prediction should be insensitive to environmental changes, provided the right values are furnished for the predictors. Of course, for biological systems we will never be able to model the variation in the response variables perfectly. For those parts escaping our acuity we posit random model terms. Nevertheless, we should attempt to cover the relevant predictors as good as possible when we purport to construct reliable prediction functions, and try to distinguish between interpolatory and extrapolatory situations.

Biological subject matter knowledge, in the form of biological models, can guide the ideas about candidate functional forms, the set of predictor variables to include plus the range over which they should be varied, and the set of nuisance variables for which appropriate corrections should be carried out. Although we often do not know whether these more elaborate models exhibit adequate goodness-of-fit over the full domain of application, usually substantial parts of the larger model have been confirmed independently of each other, which

could give us some confidence that the global fit is reasonable and that the (biological) subject matter model may be used as guidance. Statistical models are prone to optimize local fit. By embedding statistical models in biological models we hope to increase the range of validity of our models and to achieve better global properties.

3 Biological models for phenotypic expression

3.1 Crop physiology and crop growth simulation

Crop physiology is the area of biology to look at first when searching for guidance in the construction of statistical models for phenotypic expression and GEI, both with respect to functional forms and variables to be included. More specifically we are interested in crop growth simulation models that are built to predict crop production, in particular potential yield. With these models crop production can be increased by adjusting the phenology of the plant as good as possible to the local environmental conditions, by utilizing available resources, and avoiding stresses. Recent illustrations are given by Hammer and Vanderlip (1989 a,b), Hammer *et al.* (1989), Palanisamy *et al.* (1993), Ritchie (1993), Wallace *et al.* (1993 a,b), and Wallace, Zobel and Yourstone (1993). The summary below is inspired by Ritchie (1993).

In plants the quantity to be predicted, potential biomass yield, is the product of the rate of biomass accumulation times duration of growth. The rate of accumulation is principally influenced by the amount of intercepted light over a fairly wide range of temperatures. Duration of growth is in the first place dependent on the temperature, and to some extent on photoperiod during floral induction. Highest potential yield for annual crops is achieved in temperate regions where duration of the growing season is maximized. Radiation may not be too low then.

Genetic variation between plants is most evident for duration of growth and least evident for growth rate. Older cultivars with the same duration of phases as modern ones may differ in allocation patterns. For basic physiological processes as photosynthesis and respiration there seems to exist little genetic variation under 'normal' conditions. Duration of growth is most conveniently expressed in thermal time (degree days), the plant its way of experiencing time. Genotypes differ primarily in the length of the vegetative phase. In the period before floral induction, the juvenile phase, when the genotypes are not photoperiod sensitive, genotypic differences occur mainly in the form of differences in required thermal time until floral induction. The actual thermal environment eventually determines how long (real time) this period is going to be in a particular instance. During floral induction plants are photoperiod sensitive. Genetic differences for photoperiod sensitivity determine the progress in development to the reproductive phase. The rate depends on daylength. As before, the actually encountered daylengths determine the time needed to the reproductive phase in a particular instance. The rest of the life-cycle until maturity then once again is principally dependent on thermal time, but in contrast to the situation before floral induction, genotypes do not differ

greatly in the thermal time required for achieving maturity after floral induction. The times to the major phenological events are thus primarily dependent on thermal time and photoperiod sensitivity (vernalization) on the part of the genotype, and temperature and daylength on the part of the environment.

When the phenology of a plant is adequately modelled, consequences of growth under specified environmental conditions can be investigated. The phenology must be made to match the seasonal cycle of temperatures, rainfall, and soil water balance. Drought and temperature extremes should be avoided, maturity should occur before drought, heat, or frost. For quantifying resource utilization and risk, the distribution of historical weather data for a region can be used.

Crop growth simulation models focus on yield prediction. Plant varieties, however, have to meet more requirements than just high potential yield. Nevertheless, these models can provide useful information on which types of interplay between genotype and environment can lead to particular phenotypic realizations. Measurement of the environmental characteristics is probably easier than the assessment of genotypic coefficients. Especially the simultaneous assessment of the genotypic coefficients in their mutual interdependency is an underdeveloped area of crop growth simulation.

With the use of historical weather data it is possible to assess the utility of the multi-location multi-year trials that are so common in variety testing. Muchow and Carberry (1993), comparing different crops for temporally variable environments, concluded that unconditional recommendations over years were hard to make. First the low number of years used in their trials did not permit any clear prediction, and secondly what was to count as a desirable variety, depended heavily on the attitude towards risk. For example, earlier maturity may improve yield in poor years and reduce fluctuations between years, but in better years yield would be sacrificed. The use of the historical distribution of years can give insight into what may be expected from a variety over the weather conditions representative for the region.

3.2 Developmental biology

Phenotypic expression is at any time the result of previous development. For modelling phenotypic expression and GEI developmental biological insights should therefore be taken into account.

For many traits plants exhibit a clear dependency of the phenotypic expression on the environmental conditions, i.e. they have a high degree of plasticity. This is supposed to be due to the typical, modular growth of their undifferentiated parts. As these undifferentiated parts are kept on the outside, it is relatively easy to branch and add new parts. By that mechanism adaptation to changing environmental conditions is enhanced (Schmid, 1992).

Plasticity differs markedly between traits. Highly plastic are measures of size and numbers of modules, that develop over long periods of modular growth. Plasticity for biomass

allocation and partitioning to plant tissues is a major means of plant adjustment to the environment. Less plastic characters, and developing more quickly, are rules of growth as allometries, positional relationships, and branching types. Least plastic are characters related to reproduction, like flowers. Floral structures have to attract and accomodate specific visitors, and thus should be recognizable. These structures have relatively closed organization and some parts are fixed in early development (style and ovary). Seed size is often the least plastic character.

Plasticity is limited by internal and external constraints. External constraints are formed by limitations of the range in which critical variables appear in the environment. Internal constraints concern developmental constraints whose nature is in general still poorly elucidated. They are responsable for developmental stability. Development always occurs along an ordered sequence of stages and within a particular stage plants are able to assume a limited set of trait configurations (covariance matrices). The suggestion by Kauffman (1992) that developmental stability is a direct consequence of the the type of regulatory circuitry that controls phenotypic expression (genetic hierarchies with regulatory genes at the top) and not of selection, would imply bad prospects for artificial selection on plasticity itself.

Knowledge of processes determining development is accumulating. Plants may control cell fate (differentiation) by controlling cell size (Meyerowitz, 1994). Genetic regulatory mechanisms for developmental pathways seem to be strongly hierarchic and controlled by only a few regulatory loci responding to internal, developmental, and external, environmental, signals (Doebley, 1993). The evidence, however, is fragmentary and restricted to a few exemplary species. Developmental biology may help in the construction of statistical models for phenotypic expression and GEI by formulating expectations for the plasticity of traits to be modelled and the limitations to be faced within specified developmental stages.

3.3 Biological guidelines for statistical models for genotype by environment interaction

The biological models for phenotypic expression of the sections 3.1 and 3.2 share the emphasis on development. Development is a contingent process in which the actual phenotype is the outcome of complex interactions between the past phenotype, the genotype and the environment (physical and biological). During development a plant passes through a number of developmental stages. Within each stage another part of the genome is active, and different sets of environmental variables control phenotypic expression.

We now propose a mathemathical generalization of the phenotypic expression at stage T, P_T, that roughly summarizes the biological models from the previous sections and that can serve as reference point for the statistical models of the next sections:

$$P_T = \Xi_{[1,\ldots,T]} (\Omega_t (\phi_t(G),\psi_t(E),P_{t-1})).$$

Here the overall phenotypic function $\Xi_{[1,\ldots,T]}(.)$ expresses that the phenotype at stage T is a

cumulative function over all developmental stages up till now, 1,...,T. Within developmental stage t the phenotypic expression is determined by the stage-specific phenotypic function $\Omega_t(.)$ with as arguments the genotype at stage t, i.e. the stage-specific ensemble of activated genes, $\phi_t(G)$, the environment at stage t, i.e. the stage-specific ensemble of critical environmental variables, $\psi_t(E)$, and the phenotype at the previous developmental stage, P_{t-1}, representing epigenetic regulation (biological interactions between genetic materials mediated by their products). The genotype, G (to be interpreted as the total collection of genes), is modulated by the function $\phi_t(.)$ to indicate that not the whole of the genotype is active all of the time, but that there is a stage-specific activity regulated by environmental and epigenetic factors. The function $\psi_t(.)$ allows different critical environmental variables for different stages, plus stage-specific transformations.

In its most general form, statistical modelling would involve finding the functions Ξ, Ω, ϕ and ψ, besides genotype and environment specific parameters. Luckily enough, for many purposes simpler models suffice and various models often give comparable fit. Furthermore, the quality of phenotypic data very rarely allows the estimation of more than a few simple functions and a limited amount of parameters. Nevertheless, when modelling phenotypic expression statistically we should try to combine adequate goodness-of-fit for the modelling data (local goodness-of-fit) with sufficient embedding in biological models to achieve validity over a larger range of application (robustness).

It is difficult to give general biologically inspired guidelines for the choice of functional forms Ξ, Ω, ϕ, and ψ. They are available for some exceptional cases only. About the strategy for finding critical environmental variables more can be said. First of all, variable selection procedures should preferably be applied within developmental stages. If this is not possible the phenotypic response should somehow be corrected for differences in development. The candidate set of variables to be included can be deduced from physiological knowledge. Temperature sums and photoperiods are obligatory choices, whereas rainfall, soil water availability, maximum and minimum temperatures are little less than obligatory. Besides the composition of the candidate set of variables, the scale on which the variables should be included can be problematic. The scale on which a variable is measured by researchers may be very different from the scale on which a variable is relevant to the plant.

Model construction should be distinguished from model prediction. After a good fitting model has been constructed, predictions can be calculated to answer many kinds of what-if questions, like 'what will be the change in average yield of cultivars A and B when average temperature is increased by one degree?' or 'what will happen if the cultivars A and B are grown over the whole of the Netherlands for the next five years?'. For making useful predictions, ecologically meaningful descriptions of the environment are a prerequisite. Of the key environmental factors, the statistical distribution must be known to define weighing regimes for calculating predictions.

A last point of caution concerns the fact that in statistical analyses genotypes are

standardly compared at one point in time only, for example harvest time. Since traits can vary in relation to age, size, and developmental stage, common age comparisons of traits can be drastically different from common size, or common stage comparisons (Coleman, McConaughay and Ackerly, 1994). Appropriate choice of time-scale in (statistical) modelling is therefore crucial. Common age comparisons are most relevant for 'real time' processes such as plant-plant competition, but for comparing biomass partitioning common size comparisons are more suitable, while for leaf development aspects it is best to perform common stage comparisons.

4 Statistical models for studying genotype by environment interaction

4.1 Analysis of variance

A marked difference between biological and statistical models for phenotypic expression is the almost complete neglect of developmental aspects in statistical models. A clear illustration of this point is given by the analysis of variance model. The analysis of variance, or ANOVA, model is still the most influential and widely applied statistical model for describing phenotypic responses. One reason for the popularity of the ANOVA model is that the 'natural' design-structure for evaluating a number of discrete genotypes in a number of discrete environments lends itself extremely well to subsequent analysis by an ANOVA model, which is a model apt at describing responses as functions of nominal variables. It probably is no coincidence that the epistemological basis of the ANOVA model traces back to a model for quantifying polygenic gene action. In ANOVA models phenotypic responses are written as sums of genotypic, environmental and joint contributions. An ANOVA model for the mean of the phenotypic response of genotype i in environment j is $\mathscr{E}(P_{ij}) = \mu + g_i + e_j + ge_{ij}$. The expected phenotypic response is the sum of the general mean, μ, the genotypic main effect, g_i, the environmental main effect, e_j, and the *statistical* genotype by environment interaction, ge_{ij}. When the analysis of variance is going to be used for producing statistical inferences, usually a stochastic element is added to the ANOVA model in the form of an uncorrelated error term with zero mean and constant variance. It should be acknowledged, however, that the analysis of variance can also be presented as a purely descriptive tool without any reference to stochastic errors. In that case the analysis of variance can be understood as a generalization of the Pythagorean theorem (Eisenhart, 1947; Scheffé, 1959; Kempthorne, 1975). A vector of observations, the response vector, is decomposed as the sum of a number of projections on (orthogonal) spaces. The squared lengths of the projections add up to the squared length of the response vector, or, the sum of squares of the response is partitioned into contributions due to individual sources of variation.

The spaces are spanned by combinations of indicator variables that represent the levels of treatment factors. Projections on individual treatment spaces produce the main effects. Projection on the sum space of the treatment factors provides the fit corresponding to the so-called additive model. The additive model is adequate when its fit is close to the observed

response. Usually an additive model does not give a satisfactory fit and the response vector is further projected on the product spaces of pairs of factors, triples, quadruples, etc.. The difference between product and sum space is called non-additivity, or interaction. Thus, when a phenotypic response cannot be reconstituted from the projection on the sum space of the genotypic and environmental factor, further projection on the product space is necessary, and non-additivity, or statistical interaction, is invoked.

From the way in which the ANOVA model parameters are estimated it can be deduced that they are defined exclusively in terms of the phenotypic observations, there is no reference to genotypic and environmental information beyond that of the customary nominal labels for factor levels in ANOVA. This complicates their biological interpretation (embedding). Since, non-additivity is always relative to an additive reference model, the interpretation of non-additivity parameters in ANOVA models will be more ambiguous than that of main effect parameters.

4.2 Interpretation of ANOVA model parameters

ANOVA models are frequently applied in over-parameterized form. As an example we can think of a two-way ANOVA model with interaction for our response for $\mathcal{E}(P_{ij})$ that uses 1+I+J+IJ parameters for describing IJ cell means. For statistical inference it is necessary to make a choice between the use of generalized inverses for solving the normal equations and the imposition of indentifiability constraints on the parameters. Two common types of constraints are the sum-to-zero constraints $g_. = e_. = ge_{i.} = ge_{.j} = 0$ (a dot denotes averaging over that index), and the corner stone constraints $g_1 = e_1 = ge_{i1} = ge_{1j} = 0$.

Because the interpretation of the parameters changes with the choice of constraints, some authors have argued that inference should not be based on 'unobservables' like individual parameters, but on *estimable* quantities only (Nelder, 1977, 1994). Estimable are those linear combinations of parameters that can be written as expectations of linear combinations of observations, of 'observables'. For example, estimable are $\mu + g_1 + e_1 + ge_{11} = \mathcal{E}(P_{11})$, and $(g_1 + ge_{1.}) - (g_2 + ge_{2.}) = \mathcal{E}(P_{1.} - P_{2.})$. Non-estimable are μ, g_1, e_1, ge_{11}, and $\mu + g_1 + ge_{11}$. The opinion to restrict inference to estimable functions implies the abolition of potentially useful modelling opportunities. Therefore, other authors have pleaded to extend inference to parameters (unobservables) on the condition that the choice of constraints can be motivated from (biological) subject matter knowledge (Harville, 1978, 1991; Hocking, 1973).

The most popular constraints for main effects identification in ANOVA models are the sum-to-zero constraints. A main effect then represents the deviation of the general mean. This quantity is generally considered interpretable, sensible, and meaningful. Some qualification is necessary. Firstly, when an additive model needs to be supplemented by additional non-additivity terms for achieving adequate fit, the main effect parameters lose interest proportional to the size of the non-additivity. Secondly, strictly speaking the meaning of the main effects is confined to the experiment as it is done. This is a consequence of the way in

which ANOVA model parameters are defined. A genotypic main effect gives the deviation of the general mean for the set of environments included, where the general mean itself is determined by the set of genotypes and environments too. The usefulness of these parameters depends on the considerations with which the genotypes and environments have been chosen. The differences between genotypic means constitute estimable functions, and their interpretation is only subject to provisos with respect to the choice of environments.

With some effort main effects can be given biological meaning. For non-additivity or interaction this is definitely far more difficult. Non-additivity terms must be interpreted conditional on the sum of the main effects. Assuming sum-to-zero constraints again to be most natural, they are deviations from deviations (two-factor interactions), or deviations from deviations from deviations (three-factor interactions), etc.. The artificial nature of these model terms with respect to biological phenomena is evident. Nevertheless, for balanced factorial data, non-additivity terms are still fixed functions of marginal means and so entertain in a weak sense a qualified (choice of right scale) concrete existence. For unbalanced data identification constraints become harder to justify, and the pattern of missing cells will influence the interpretation of model terms.

Two necessary conditions for arriving at biologically meaningful interpretations of non-additivity parameters may be formulated. Firstly, it must be biologically meaningful to condition on the corresponding reference model, i.e the additive model for two-way interactions, the model including main effects and two-way interactions for three-way interactions, etc.. Secondly, the non-additivity parameters must be *directly* relatable to biological phenomena. The adverb 'directly' is used to indicate that further modelling is thought obligatory to make anything out of the non-additivity at all. Because the statistical noun 'interaction' suggests causal interdependency in biological contexts, we propose to reserve 'interaction' for those occasions in which non-additivity indeed is further modelled. With respect to non-additivity things again become simpler when upon deciding on the necessity of including non-additivity terms in the model, we focus subsequently on estimable functions only, like predicted responses over environments, or the basic two-way interaction contrasts (tetrads), $\theta_{ii'jj'} = \mathcal{E}((P_{ij} - P_{ij'}) - (P_{i'j} - P_{i'j'})) = \mathcal{E}(P_{ij}) - \mathcal{E}(P_{ij'}) - \mathcal{E}(P_{i'j}) + \mathcal{E}(P_{i'j'})$.

4.3 Additivity

Additivity applies when an additive model gives an adequate fit to observed data. Additivity is a desired property. Firstly, because of the simplicity of the additive model, where a sum of discrete variables provides an approximation to whichever function. Secondly, because additivity suggests causal independence of the row and column factor (Cox, 1984), implying that the predicted response can *always* be obtained from the sum of the contributions of the determining factors. (One might wish to argue that also the more elaborate row regression model, $\mathcal{E}(P_{ij}) = g_i + \beta_i e_{ij}$, with g_i, β_i, and e_{ij} parameters, still suggests causal independence.)

Causal independence of factors may follow from additivity in technological applications, but it certainly does not in the life sciences. When causality is interpreted as unconditional association, additivity of the genotypic and environmental factor for the description of the phenotypic response would mean that under all circumstances knowledge of the genotypic effect and the environmental effect would suffice for adequate prediction. This seems unlikely. As explained before, the genotypic and environmental effects in ANOVA models are heavily dependent on the conditions in which the experiment was done. Including other genotypes or environments might completely change the effects. For example, in a resistance breeding trial measuring disease incidence, a lightly resistant variety in a set of susceptible varieties will have a negative genotypic value, whereas the same variety in a set of very resistant varieties will have a positive genotypic value. As there are no external reference points for the genotypic and environmental values in ANOVA models, it is unclear what these values mean outside of the experiment. The quality of the information breeders can extract from the application of ANOVA depends on the carefulness with which genotypes and environments have been chosen.

It may be objected that this may be all true for fixed sets of genotypes and environments, but that the situation changes if the assumption is made that the genotypes and environments form a representative sample of some population. To refute this objection we appeal to a second argument for which it is immaterial whether our genotypes and environments are fixed or random. In the generalization of section 3.3, the phenotypic expression at time T depends in a complicated way on the phenotypic expression at previous stages. Also different parts of the genotype were active at the various stages. The sequence of interconnected developmental stages makes it in principle impossible to separate unconditional genotypic effects from unconditional environmental effects in an analysis at a fixed point in time. A possible exception may be the phenotypic expression at the first stage, although there the effects of our idealization of time as being discrete may discredit the conclusion of additivity.

Another argument against coupling additivity and causality stems from the interpolation-extrapolation distinction. Let causality still mean unconditional association. Assume that analysis of experimental data leads to the conclusion of additivity of genotype and environment. Since environments within ANOVA are characterized solely by nominal labels, it is difficult to tell whether predictions involve interpolations or extrapolations. The objection that we should have sampled our environments representatively is rather weak when it is not clear at all in which respects the sampling should have been representative.

For the modelling of the phenotypic response additivity is desirable from the point of view of parsimony. Conclusions with respect to biological causality are impossible as this would require an embedment that cannot be verified with data collected at only one point in time. However, there is no need for an embedment that guarantees causality as long as our predictions are adequate enough for our purposes. Since there is no clear connection between the additive model and causality, failure of statistical additivity primarily means that extra

non-additivity terms are necessary for prediction, but it does not necessarily mean that the nature of the statistical model suddenly touches upon an essentially different biological mechanism.

4.4 Non-additivity

The ANOVA model for the phenotypic response in section 3.3 is dramatically simpler then the general expression for phenotypic expression of section 3.3. There is no counterpart for development in the ANOVA model, or, $\Xi_{[1,...,T]}(.) = I_T$, the identity function at T. The time index can be dropped altogether. The genotypic function ϕ and the environmental function ψ at first sight seem to be identity functions as well, but are not well defined because they have no clear argument as g_i and e_j are defined in phenotypic terms only and are not referring to anything that is explicitly genotypic or environmental. This argument holds even stronger for the non-additivity term, ge_{ij}. The best we can do is to replace $\Omega(\phi(g_i),\psi(e_j))$ as a whole by $g_i + e_j + ge_{ij}$, without claiming one-to-one correspondence between functions and arguments.

For reasons of parsimony we prefer ANOVA models consisting of main effect terms only, $\mathcal{E}(P_{ij}) = g_i + e_j$, and to leave out non-additivity terms as much as possible. If non-additivity terms should be included we want them to have a biological interpretation. However, non-additivity terms can arise through a number of biological as well as statistical mechanisms.

A familiar method for trying to remove non-additivity terms from ANOVA models is transformation of the response. It is hoped that the non-additivity term necessary for an ANOVA description of the response, $\mathcal{E}(P_{ij}) = g_i + e_j + ge_{ij}$, can be removed by using a function of the response instead of the response itself, $\mathcal{E}(f(P_{ij})) = g'_i + e'_j$. A necessary, but not sufficient, condition for removable two-way non-additivity is that both rows and columns of the two-way table of expectations can be consistently ordered (Scheffé, 1959). Two rows x_{11}, $x_{12},...,x_{1J}$ and x_{21}, $x_{22},...,x_{2J}$ can be consistently ordered when the differences $(x_{1j} - x_{2j})$ are either all >0 or =0, or all <0. Similarly so for columns.

Transformation not only influences the scale properties (first order properties) of the response, but also affects the variance (second order properties). Heterogeneity of variance is often purported to be a main cause of non-additivity itself. Transformations may change heterogeneity of variance into homogeneity. One and the same transformation need not cure both scale and variance problems simultaneously. Therefore, transformations of the response usually involve some kind of compromise between optimal scale and variance properties. These compromises can be circumvented by using generalized linear models (GLMs) that allow separate modelling of the scale and variance properties (McCullagh and Nelder, 1989). In GLMs not the response itself is transformed to become linear in the parameters, but the expectation of the response. The GLM equivalent of the additive two-way ANOVA model with interaction consists for the expectation in $f'(\mathcal{E}(P_{ijk})) = g_i + e_j$. The function $f'(.)$ that transforms the expectation so that it becomes linear in the parameters is called the link

function. The GLM specification for the variance describes the variance as a function of the mean. For the standard ANOVA model f'(.) is the identity function and the variance function is a constant.

A well-known example of removable interactions involves the transformation of biologically multiplicative processes to statistically additively behaving responses. Many biological processes are multiplicative instead of additive. Fitting an ANOVA model to multiplicative data makes non-additivity terms appear necessary, while the error will not have constant variance. Logarithmic transformation remedies this situation. Non-additivity terms will no longer be necessary and the variance will have been stabilized. An example of a biologically multiplicative response is a phytopathological model for host-pathogen interaction that describes the disease reaction of a host to a pathogen by the product $1/2.n_r.n_v$, where n_r is the number of susceptibility alleles in the host and n_v is the number of aggressiveness alleles in the pathogen (Model B in Carson, 1987). For data generated according to this mechanism, logarithmic transformation would remove the apparent interaction present at the scale of observation.

Another important (statistical/biologial) source of non-additivity are outliers. Outliers, located outside the model, can induce inclusion of non-additivity terms, inside the model. For example, in resistance breeding it is of interest to find genotypes that are specifically resistant to particular isolates (genotypes) of a disease causing agent. Non-additivity can then be very local, caused by a few outlying cells. To find these cells either robust outlier detection techniques are recommended, or a robust ANOVA. For the biological conclusion it does not matter whether we interpret the specific resistances as being part of the statistical model in the form of non-additivity, or as outliers.

There exist mutual trade-offs between non-additivity, transformations, outliers, and heterogeneity of variance (Atkinson, 1982). Complicated situations can arise where some cells can indicate a need for transformation that disappears when the cells are interpreted as outliers. Alternatively, cells may become outliers after transformation. As there is no such thing as the right model, very different models can give almost equivalent fit to the same data. The only way to distinguish between these models is to add data supposed to influence the fit differentially. Alternatively, parts of the data can be left out to check whether certain features are not due to a minority of deviating observations.

4.5 Separability

In plant breeding GEI tends to be considered to be roughly equivalent to non-additivity in ANOVA models and is thus primarily a statistical entity. One attempt to connect statistical and biological interaction in a more meaningful way supplants the statistical notion of interaction as non-additivity by the mathematical notion of separability (Gregorius and Namkoong, 1986, 1987). The argument asserts to extend an argument by Lewontin (1974) that for describing genotypic response functions ANOVA models are inappropriate and even

misleading, because the statistical genotypic and environmental main effects can be mistaken to refer to biological genotypic and environmental affects (causes). Gregorius and Namkoong interpret the problem as that of linear ANOVA models being inadequate for describing non-linear response functions. Phenotypes are response functions of genotypic and environmental variables, but for non-linear responses it is not so evident how to define 'genotypic' and 'environmental', and their joint operation in a sensible way. The suggested solution was to think of the phenotype as $P_{ij} = \Omega(\phi(g_i),\psi(e_j))$, with g_i and e_j being the genotypic and environmental effect, $\phi(.)$ and $\psi(.)$ being a genotypic and environmental transformation function, and $\Omega(.,.)$ being a joint operator function of the transformed genotypic and environmental effect. When the phenotypic response can be expressed in the above way the effects of genotype and environments are said to be separable, and no biologically interesting interaction should occur. Failure of separability would imply biological interaction.

For univariate responses a necessary condition for separability is non-intersection. This becomes a sufficient condition with monotonicity or certain symmetry requirements. Separability entails that a reference response can be designated, $\Omega(\phi(g_0),.)$, from which all other responses $\Omega(\phi(g_i),.)$ can be derived by a one-to-one mapping. In practice this means that when the responses are plotted against the basic response, no intersection of responses can be observed. Separability, as defined by Gregorius and Namkoong, is almost equivalent to the absence of so-called crossover interactions (Baker, 1988), where the ranking-order of genotypes changes going from one environment to another.

Contrary to the claims, separability of genotypic and environmental effects still has little to do with the desired separation of the genotypic and environmental 'causes' that shape the phenotype. The main reason for this has been given already at the start of section 4.4. The functions ϕ and ψ are poorly defined, because they contain no explicit arguments. But even if the arguments had been explicit, solutions for Ω, ϕ, and ψ would still not have been unique, as there will always exist considerable freedom to move between the various forms of Ω, ϕ, and ψ. Another, by now familiar, reason for the failure of the separability concept to separate genotypic and environmental causes is the obvious neglect of development. This not to say that the concept of separability has no useful aspects, but these coincide, if not wholly then to a large extent, with the theory for finding transformations for removable interactions in two-way tables (Scheffé, 1959, p. 95-98).

4.6 Random terms in ANOVA models

In the previous sections it has implicitly been assumed that all model parameters had a fixed value (except the random error). However, model parameters may also be defined to follow a distribution. ANOVA models often consist of a sum of fixed and random terms. The random terms in ANOVA models are typically normal, independent, and identically distributed, with mean zero. There are various reasons for defining model terms to be random, of which the desirability of recovery of information and shrinkage towards the mean are the

most important ones. For example, by taking locations and years random in variety trials, variety information can be recovered from location and year totals. Taking variety effects random will shrink individual variety effects towards the mean of all varieties effects, thus correcting for selection bias. A statistical technical reason for choosing a term to be random is to avoid the estimation of a great many individual (fixed) parameters. Another technical consideration leads to random terms for modelling (intra-class) correlations. The choice between fixed and random is for the largest part a pragmatic one (Robinson, 1991). When considering a term to be random, the effects should at least look like as if they could have come from the assumed distribution.

For random terms individual parameter estimates and variance components are of interest. Their interpretation depends on whether and which identification constraints have been used. The imposition of identification constraints for random terms has been even more controversial than for fixed terms. Nelder (1977, 1994) proposed to abstain from any constraints altogether, with the argument that expected mean square tables then would consistently indicate how to test main effects irrespective of the choice of the treatment factors as fixed or random. For example, for a balanced two-way table both main effects are then always tested over the interaction. With constraints, the random main effect in mixed models cannot be tested over the interaction anymore. For unbalanced data only the constraints-free version of mixed and random models is feasible, so the problem can then be discarded. Less severe in their condemnation of constraints were Hocking (1973) and Harville (1978), who defended the viewpoint that the choice of constraints should be made dependent on extra-statistical arguments.

A good example of how the research question can guide the choice of constraints is given by Fry (1992), who compared two popular versions of the two-way mixed model. One version has sum-to-zero constraints imposed on the random interaction parameters within the levels of the fixed factor, causing correlation. The second version is free of constraints. In the constraints-imposed version the variance component of the random main effect can be interpreted as the variance of the marginal means. In the constraints-free version the corresponding variance component expresses a covariance between the responses for the levels of the random factor across the fixed factor. When the random factor represents genotypes and the fixed factor environments, the constraints-imposed version is useful for the calculation of genetic variances and heritabilities, whereas the constraints-free version is useful for assessing the correlation between genotypic performances across environments. For a concrete example we can think of experiments in resistance breeding where genotypes are evaluated on their resistance to a specific set of isolates of a pathogen.

4.7 Factorial regression models

Though ANOVA models are generally useful for modelling phenotypic responses, they should be taken as a first step in the modelling process, definitely requiring follow-up

environmental covariables for describing two-way non-additivity between genotypes and isolates (genotypes of a disease causing agent). Additive parameters are estimated that are typical for the intersections of groups of genotypes with groups of isolates. The groups can be obtained from subject matter knowledge (confirmatory approach), or from the application of a clustering algorithm, like that described by Corsten and Denis (1990) (exploratory approach). Genotypes are then clustered by their pattern of specific resistances and susceptibilities, and pathotypes by their pattern of specific (a)virulences (Arntzen and van Eeuwijk, 1992).

An area of application where meaningful genotypic covariables are present by definition, is in the development of greenhouse tests for field traits, as diverse as field emergence and disease incidence. We give two examples. Firstly, non-additivity for field emergence in white cabbage could very well be described in terms of genotypic greenhouse test assessments and environmental potentialities (II). The presumed absence of large developmental differences in this kind of data may underlie this success. Secondly, yield and quality figures in sugar beet trials with varying level of Rhizomania infection could very well be described by a factorial regression model including only one cross-product term; namely that between a greenhouse resistance test for the genotypes and a field infestation assessment for the trials (III). Only a scale constant remained to be estimated. The consequence of the interaction term was that susceptible cultivars were corrected downward in infested fields and upward in clean fields. Cross-products between genotypic and environmental variables can have very simple interpretations in terms of compensations.

Factorial regression models offer the opportunity to model developmental differences. Development may be included as a continuous genotypic covariable in the form of earliness, or as a categorical genotypic covariable in the form of maturity classes. Baril *et al.* (1995) analysed yield in potato and found an important factorial regression term to consist of the product of the genotypic covariable earliness (Variety List figures) and the environmental covariable mean temperature over the growing season. Early potato genotypes were found to benefit from high temperatures during the growing season, whereas later genotypes did relatively poorly under these conditions, because they suffered from the drought stress brought about by these higher temperatures later on in the season. Low temperatures caused slow growth, which affected early genotypes most. Later genotypes could compensate through a longer growing season.

When maturity is not incorporated as an explicit genotypic variable, it can still turn up as sensitivity to included environmental variables. Interaction for seed yield in perennial ryegrass could be modelled by the differential sensitivity of cultivars to the minimum temperature in the developmental period preceding ear emergence (IV). Upon closer inspection these sensitivities reflected a contrast between early and late cultivars. Similarly, non-additivity for dry matter content in maize could be described by a factorial regression model including environmental variables as mean dry matter content (cf. regression on the

mean) and irradiation. Maturity groups of maize cultivars were identified on basis of the genotypic sensitivities to the included environmental covariables (XI).

When using statistical models for GEI it is important to distinguish between the model *building* and model *prediction* phase (Nelder, 1977, 1994; Lane and Nelder, 1982). In the model building phase we try to find estimates for the model parameters by minimizing a particular loss function (local fit), for example a least squares criterion. In the model prediction phase we try to answer questions of the what-if type (global fit/ robustness). Examples are predictions for the mean phenotypic response at a particular location, or for a region. For calculating these predictions some kind of averaging over years must take place. The weighing regime for calculating predictions need not be equal to that of the estimation process. For local and regional predictions the historical distribution of years constitutes a more logical choice than a uniform distibution over the sampled years, attributing equal weight to each sampled year.

For defining the 'historical' distribution of years it is necessary to characterize individual years. An obvious choice are the meteorological variables that are selected for a factorial regression model describing the set of phenotypic responses in their dependence on the environmental conditions. The joint historical distribution of meteorological variables defines a relevant weighing regime for calculating local and regional predictions. However, for studying the consequences of climatological changes, alternative weighing regimes should be used.

For predicting varietal year means over locations, locations should be characterized by soil variables, with the spatial distribution of these soil characteristics providing a basic weighing regime. For general cultivation and use value predictions, combinations of temporal climatological and spatial edaphic distributions are appropriate.

Crop growth modellers regularly study the consequences of different weighing regimes in the model prediction stage (for a recent example see van Noordwijk, Dijksterhuis and van Keulen, 1995). In plant breeding weighing regimes receive attention in the risk assessment approach propounded by, among others, Eskridge (1990). However, ample opportunities exist for further elaboration of this concept. The design of (multi-environment) variety trials is a first major candidate theme that would benefit from an approach consisting of the use of factorial regression models incorporating spatially and temporally varying environmental variables coupled to empirically assessed weighing regimes (for a cautious start, see Muchow and Carberry, 1993).

The types of models encompassed in the class of factorial regression models are sufficiently close to biological models to define embedding relations. Factorial regression models for GEI ascribe interaction as being due to differential genotypic sensitivity to critical environmental variables. The idea of differential genotypic sensitivity has obvious biological connotations. The modelling on explicit covariables for a major part removes the drawbacks of ANOVA models. The replacement of the non-additivity, ge_{ij}, by terms of the form $\beta_i z_j$,

furnishes the phenotypic function Ω with clear meaning: $\Omega(a,b) = a.b$ (assuming for the moment that the main effect interpretations of g_i and e_j are unproblematic). The function Ω includes transformations (both of the response and its expectation). The environmental modulation function, ψ, not only allows a selection of environmental variables, but also their transformation. The genotypic modulation function, ϕ, has meaning in so far as it addresses transformations of genotypic covariables, the selection of genes is beyond the scope of the ordinary factorial regression models (although quantitative trait loci, or QTL-, representations of the genotype bring this possibility within scope). Factorial regression models are still directed at modelling phenotypic responses at only one point in time, making the developmental function $\Xi_{[1,...,T]}$ of section 3.3 effectively equal to I_T, just as for ANOVA. However, with factorial regression some crude forms of correcting for development have become available.

From the embedment in biological theory suggestions can be obtained for the choice of functions, the set of covariables to include, and the kind of transformations for response and covariables. The high similarity between the structure of biological and factorial regression models should guarantee good global properties, i.e. accurate predictions over a wide range of environmental conditions. The structure of factorial regression models together with the distribution of relevant environmental variables even allows an empirical assessment of the interpolation zone.

4.8 Multiplicative models for parts or the whole of the phenotypic response based on multiplicative decompositions of two- and three-way tables

Factorial regression models are very useful for modelling interaction when concomitant information is present, and when the non-additivity can be approximated by a series of clearly distinguishable factorial regression terms. For various reasons factorial regression is not always feasible. Concomitant information may be partly or completely missing, or covariables may be so intricately interrelated that a sort of marginal approach is preferable over the conditional approach of factorial regression. An approach in between the local, unparsimonious approach of ANOVA and the controllable interpolation approach with external orientation of factorial regression consists in the use of multiplicative models for interaction based on multiplicative decompositions of tables of ANOVA parameters. This type of multiplicative models is usually referred to as multilinear models. The most applied models from this class are the bilinear models (Gollob, 1968; Mandel, 1969; Gauch, 1988; Denis, 1991), also called biadditive models (Denis and Gower, 1992, 1994). Bilinear models are especially popular for the parsimonious modelling of the structureless ANOVA non-additivity by means of a few multiplicative terms. The individual terms in bilinear multiplicative models for interaction can, just like those in factorial regression models, be interpreted as embodying differential sensitivities to environmental variables, but for bilinear models both genotypic sensitivities and environmental potentialities have to be estimated from the data themselves

by minimizing a least squares criterion (Gollob, 1968; Mandel, 1969; Gabriel, 1978) or least deviance criterion (for generalized bilinear models, XII). Bilinear terms consist of hypothetical environmental characterizations that discriminate best, in a least squares sense, between genotypes. Parsimony is guaranteed for bilinear models, but global properties are less clear than those for factorial regression. For biological interpretations the genotypic sensitivities, or genotypic scores, and the environmental characterizations, or environmental scores, should be inspected on emerging patterns, and be related to available genotypic and environmental concomitant information, even if the latter information is only nominal and/or incomplete.

Bilinear models for interaction approximate the non-additivity by a sum of (orthogonal) products. For example, $ge_{ij} = \gamma_{1i}\delta_{1j} + \gamma_{2i}\delta_{2j}$ (ignoring a residual term), where the γ's are genotypic scores and the δ's are environmental scores. The product terms are called bilinear to express the fact that upon fixing the row parameters the terms become linear in the column parameters, and upon fixing the column parameters they become linear in the row parameters. An alternation of row and column regressions can be used as a general algorithm for the estimation of parameters in bilinear models (Denis, 1991; XII).

Bilinear reformulations of two-way non-additivity can be interpreted as low rank approximations to the original two-way tables of non-additivity parameters. Therefore, the interpretation of the bilinear reformulation should proceed by considering the multiplicative terms simultaneously. The biological processes underlying the non-additivity may, but do not need to coincide with individual product terms. However, as a whole these processes should be captured by the *joint* configuration of product terms. For inspection of this joint configuration the graphical device of the biplot is an essential tool (Gabriel, 1971; Kempton, 1984; X).

Genotypes and environments are simultaneously represented in the biplot by vectors whose end point coordinates are determined by the genotypic and environmental scores. The origin is usually an imaginary additive genotype (environment). Distances between genotypic (environmental) vectors represent interaction between genotypes (environments), angles between genotypic (environmental) vectors are proportional to correlations. Projections of genotypic vectors on environmental vectors are helpful for approximating specific non-additivities. Inspection of biplots can lead to interpretations of the interaction as being due to differential genotypic sensitivity to further to be identified environmental variables, to genotypes differing with respect to particular contrasts of environments, and to (conditional) outliers (VII).

Besides biplots, i.e. joint plots of genotypic and environmental scores, plots containing either genotypic or environmental scores can be useful. A good example concerns phytopathological research investigating (non-)specificity of genotypic resistance. Assume phytopathological genotypic evaluations are summarized in the form of two-way tables, with the environmental dimension being the product of strains, locations and years. The table is corrected for the main effects of genotypes and environments and the residual two-way non-

additivity is approximated by a bilinear model. The environmental scores are depicted and inspected for clusters. When the environments do not cluster on the basis of common strains, other factors are more important for the interaction and there is no indication for the resistance being specific (VI, XIII).

Bilinear reformulations need not be restricted to two-way non-additivity. Reformulations may be given of two-way tables that contain; 1) data as such, i.e. are uncorrected for any main effect; 2) data that are corrected for an intercept term only, as for shifted multiplicative models, or SHMMs (Seyedsadr and Cornelius, 1992a), with the concurrence model being a special case (Tukey, 1949; Mandel, 1961); and 3) data that are corrected for the row or column main effect, with the least squares solution form of the regression on the mean as a special case (Gabriel, 1978). In fact, bilinear terms may be fitted to the two-way residuals from whichever two-way model. Thus, also residuals from factorial regressions may be fruitfully inspected for further structure in this way.

SHMMs are said to facilitate biological understanding of phenotypic expression via their detection power for separability (Cornelius, Seyedsadr and Crossa, 1992), and the identification of sets of genotypes (environments) without cross-over interactions (Crossa *et al.*, 1993; Cornelius, Van Sanford and Seyedsadr, 1993). Separability is present when a $SHMM_1$ or concurrence model, $P_{ij} = v + \gamma_i\delta_j$, fits the data well and all multiplicative parameters are of the same sign (Cornelius, Seyedsadr and Crossa, 1992), so that the responses consist of a number of converging or diverging monotonic lines without intersection. In that case no cross-over interactions will be found.

Generalized bilinear models allow separate modelling of a bilinear expectation and a non-homogeneous variance, and thus integrate the advantages of multiplicative modelling of the expectation with facilities to deal in an explicit way with the problem of heterogeneity of variance. Generalizing bilinear models to generalized bilinear models is rather straightforward when acknowledging the alternating regression interpretation of standard bilinear models, with identity link and constant variance. By replacing the regressions with generalized regressions an estimation algorithm for generalized bilinear models is created (XII). The interpretation of interaction in generalized bilinear models is not so straightforward, but the same principles for simplification of the model may be used, i.e. regression and contrast reformulations, and outlier identification. Again, resistance breeding applications seem most appealing, as many measurements in resistance breeding do involve non-normal variables like counts and proportions.

A generalization of two-way bilinear multiplicative models for two-way tables to three-way quadrilinear multiplicative models for three-way tables can be effectuated by generalizing the alternating row and column regressions for bilinear models to a four step alternating regression scheme (XV). This quadrilinear reformulation is very effective in the identification of the few specific three-way contrasts and (conditional) outliers that are usually responsible for the majority of the three-way non-additivity. The joint use use of two- and three-way

multiplicative models for interaction may help to identify the most important sources of non-additivity using only a fraction of the degrees of freedom (parsimony). Hypotheses that are generated could be further tested with factorial regression.

Multilinear multiplicative reformulations of ANOVA model terms suffice as a means of analysis when the research question to be answered can be expressed directly in terms of a pattern to be expected in plots and biplots. An example is given by the patterns to be expected in environmental scores in case of the (non-)specificity of resistance. Multilinear reformulations are also useful for simplifying non-additivity structures by identifying interactions due to differential genotypic sensitivities to known or unknown environmental variables, due to interacting genotypic and environmental contrasts, and conditional outliers. Complexities in the interaction may thus be resolved, leading to better factorial regression models for prediction. The possibilities for embedment of multilinear models in biological models vary between the limits imposed by ANOVA models on the one hand and factorial regression models on the other hand. The amount of a posteriori external referencing is decisive for the exact position along this continuum. The (global) predictional merits of models with multilinear terms that do not refer to concomitant information outside the model are doubtful. Indeed, in a sense multiplicative decompositions separate structure from noise and predictions may be supposed to receive higher accuracy from this process (Gauch, 1988, 1992; Seyedsadr and Cornelius, 1992b). However, without external references for the multilinear multiplicative parameters it will remain hard to differentiate between interpolations and extrapolations.

4.9 Reduced rank regression

Bilinear and quadrilinear models for interaction approximate tables of ANOVA effects by tables of lower rank. To give an example, the rank of an IxJ two-way non-additivity table is the minimum of (I-1) and (J-1). This table is approximated by a table whose rank is equal to the number of bilinear terms used. The rank m bilinear approximation to two-way non-additivity should capture the essence of the structure present in the (I-1)(J-1) independent non-additivity parameters by m(I+J-2-m) independent parameters. The gain is most notable for large tables which can be approximated by rank one or two reformulations. Fewer parameters enhance interpretation. Furthermore, low rank multiplicative reformulations allow biplot representations of the interaction, which are important devices for discerning complicated interaction patterns.

Factorial regression reformulations of non-additivity allow a direct interpretation of interaction in the form of differential genotypic sensitivity to environmental variables. The use of biplots for discerning complex interaction patterns is limited to the situation with only two concomitant variables on either the genotypic or the environmental factor, and even there the interpretation can be hindered by non-orthogonality of the covariables.

We would like to have a means for combining the strong point of factorial regression,

direct interpretation with regard to genotypes and environments, with the strong points of bilinear and quadrilinear formulations, parsimony and biplots for uncovering complex patterns. A synthesis of factorial regression reformulations with multilinear reformulations is given by reduced rank factorial regression (Denis, 1988; V, X). Reduced rank factorial regression is a generalization of factorial regression as well as multilinear reformulations. We will restrict ourselves to the one-way reduced rank factorial regression model (for a two-way genotype by environment table). This model can be understood as a factorial regression model for which the matrix of genotypic regression coefficients (number of genotypes x number of environmental covariables) is approximated by a matrix of lower rank. Simultaneously, it can be interpreted as a reduced rank approximation to the non-additivity table where the environmental scores not only have to comply with the least squares criterion of generating maximal genotypic discrimination, but also with the restriction of having to be linear combinations of (explicitly measured) environmental variables. This form of reduced rank factorial regression is thus still bilinear. The results of reduced rank factorial regression can be displayed in biplots, where covariable vectors are added to the already familiar genotypic and environmental vectors (ter Braak and Looman, 1994; X). Besides relations between genotypic and environmental vectors, a relation is defined between the genotypic vectors and the covariable vectors. The projections of the genotypic vectors on the covariable vectors are proportional to the factorial regression coefficients. Thus, reduced rank factorial regression biplots combine the best of both worlds (factorial regression models and bilinear models).

Reduced rank factorial regression has been applied only rarely in plant breeding (V, XI), but the potential of the method seems great enough to raise research efforts.

4.10 A strategy for the statistical modelling of GEI

The main conclusions of the previous sections will be used to formulate recommendations and a strategy for modelling phenotypic responses in which GEI is suspected. It will be evident that we assume that all relevant biological knowledge has been listed. The individual steps in the modelling process have no fixed order. A cyclical pattern is usually needed to arrive at a satisfactory model (see section 4.4).

A first choice in the statistical modelling of phenotypic expression involves the scale on which we want to model our effects, i.e. choice of link (GLMs) or transformation (standard ANOVA). Many scales are possible, but use of biological knowledge should limit the choice to a few serious candidates only. In the absence of elaborate biological knowledge, acknowledgement of the measurement scale of the variable may determine this choice. For example, for modelling counts log link or log transformation are the first candidates.

A second choice concerns the variance function. Historically this is a constant function, implying that the variance is independent of the mean. With the advent of GLMs other choices have become available. For counts a Poisson variance might be tried. Recent developments posit models for the variance reminiscent of the models for the expectation (for

an example, see Engel and Keen, 1994).

Although the fitting of an ANOVA model should never be the end point of an analysis of a group of phenotypic responses, it certainly provides a good starting point (including its GLM counterpart). The ANOVA decomposition of variation gives first insight into the structure of the data. Absence of significant non-additivity does not imply absence of GEI, as all of the GEI may be present in a few degrees of freedom. Alternatively, significant non-additivity terms may vanish when some points are identified as outliers and subsequently are removed from the model. Choices of model terms as random instead of fixed should follow from the pragmatic reasons listed in section 4.6 (recovery of information, shrinkage).

After fitting an ANOVA model, parsimony can be increased by replacing parts of the ANOVA model by multilinear formulations, thereby effectively reducing the rank of individual ANOVA effects tables or linear combinations of them. Statistical/ mathematical criteria have been developed for assessing the amount of rank reduction that can be imposed without too severe information loss (XII,XV). These formal criteria are preferably not followed too rigorously, as they tend to lead to the retention of too many multilinear terms. A more sound approach may consist in the incorporation of only those multilinear terms for which a biological interpretation can be found. For balanced data, a simple procedure for achieving rank reduction is to decompose the tables of two- and three-way non-additivity. For unbalanced data slightly more complicated procedures must be applied (Denis, 1991; XII). Usually a considerable gain in parsimony results and biplots of scores may indicate underlying mechanisms. Sometimes this can be the end point of the analysis, as for phytopathological research on (non-)specificity of resistance. At other times hypotheses should be tried to be tested further within factorial regression models.

An alternative way of searching for more parsimony in models for two-way tables consists in the application of diagnostic biplots. Diagnostic biplots provide model indentification keys for a number of rank two and three models based on special geometric properties of their expectation structure (Bradu and Gabriel, 1978; Gower, 1990; IX). The same keys that are valid for the whole of the table also apply for parts of the table. Thus simpler models may be identified for different parts of the table, whereas the whole of the table would require a more complicated model. For example, additivity may exist for parts of the table, but remain hidden when one and the same model is fitted to the whole of the table. Partial additivity can also be found by the application of clustering algorithms, as for example the one due to Corsten and Denis (1990), that create subsets of genotypes and environments that are internally homogeneous (additive), relegating interaction to between group contrasts.

When present, explicit genotypic and environmental concomitant information should always be used to model GEI. Information on developmental differences between genotypes, whether nominal or continuous, may routinely be incorporated without further testing. For other variables a preselection based on subject matter knowledge is possible (see section 3.3).

For genotype by location interaction edaphic information (soil characteristics, water availability) can be considered, for genotype by year interaction meteorological information (daylength, temperature). From the point of view of parsimony, rank reduction of the matrix of regression coefficients becomes more interesting with increasing numbers of incorporated covariables. However, the possibility of biplots merits the standard application of reduced rank factorial regression for all situations with more than one covariable.

After having replaced non-additivity by parsimonious factorial regression and/or multilinear structures, the local fit of the model can be checked by inspection of the residuals. Extra structure may be suspected (further multiplicative decompositions of residuals), outliers may be identified (outlier detection methods), or residual variation may be found to vary with genotype (tests for heteroscedasticity). Heterogeneity of residual variance might lead to the inclusion of extra variance components (Shukla, 1972; Piepho, 1995). After having finished the model building phase, predictions can be calculated. Weighing regimes should be chosen in accordance with the questions to be answered.

The joint use of GLM-, ANOVA-, factorial regression- and multilinear techniques should enable us to fit adequate models to the majority of genotype by environment tables, and should allow us to find answers to most of our relevant questions.

5 Future research directions

Though many types of phenotypic responses can be represented adequately with the methods described much work remains to be done, statistically as well biologically. On the expectation side of our statistical models multivariate extensions are urgently needed. Various approaches are possible. Firstly, we could try to develop more complicated network-like representations of dependencies between variables in the expectation structure, similar to those which are common now in the modelling of LInear Structural RELationships (LISREL), to arrive at closer correspondence with causal mechanisms ("make your models more elaborate"). In this respect the developing theory on graphical chain models seems to provide interesting prospects (Wermuth and Lauritzen, 1990). Secondly, methods should be investigated that fit common multiplicative structures for a number of traits simultaneously (various options exist: Kroonenberg, 1983; Basford *et al.* 1991; Denis and Moro, 1995; van Eeuwijk and Kroonenberg, 1995).

On the variance side of our statistical models various generalizations merit further study. Firstly, interesting prospects await the development and implementation of *mixed* factorial regression, with extensions for dealing with heteroscedasticity. Research has been initiated already in this direction (Denis, Piepho and van Eeuwijk, 1996). Secondly, extra attention is needed for the development of bilinear models in which one or both of the row and column scores are taken random. First results along this line can be found in Oman (1991) and Gogel, Cullis and Verbyla (1995). A third direction in variance structure research worthwhile pursuing concerns the integration of generalized linear mixed models with mixed factorial

regression and mixed (random) bilinear models.

Currently, a very promising avenue for future research is given by the implementation of more refined genotypic representations in statistical models for phenotypic expression and GEI, thereby increasing the biological content of these models. The rapid development of the quantitative trait loci technology has brought a new challenge. Quantitative trait loci expression may be related directly to environmental variables in further elaborated, more 'causally' orientated, factorial regression models. Regression based methods for quantitative trait loci mapping (Jansen and Stam, 1994) present an ideal starting situation for generalizations in that direction.

Acknowledgements

Thanks are due to Piet Stam, Jean-Baptiste Denis, Jan-Peter Nap, Michiel Jansen, Robert van Loo and Hans-Peter Piepho for comments and discussion.

Account of chapter numbers

II van Eeuwijk, F.A., and Kraak, H.L. (1995). Two classes of multiplicative models for the analysis of genotype by environment interaction in seed technology experiments assessing field emergence, with an application to white cabbage. Submitted.

III Paul, H., van Eeuwijk, F.A., and Heijbroek, W. (1993). Multiplicative models for cultivar by location interaction in testing sugar beet for resistance to beet necrotic yellow vein virus. *Euphytica* **71**, 63-74.

IV van Eeuwijk, F.A., and Elgersma A. (1993). Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial regress. *Heredity* **70**, 447-457.

V van Eeuwijk, F.A. (1992). Interpreting genotype-by-environment interaction using redundancy analysis. *Theoretical and Applied Genetics* **85**, 89-100.

VI Snijders, C.H.A., and van Eeuwijk, F.A. (1991). Genotype x strain interactions for resistance to Fusarium head blight caused by Fusarium culmorum in winter wheat. *Theoretical and Applied Genetics* **81**, 239-244.

VII van Eeuwijk, F.A. (1992). Multiplicative models for genotype-environment interaction in plant breeding. *Statistica Applicata* **4**, 393-406.

VIII van Eeuwijk, F.A. (1993). *Genotype by environment interaction; Basic ideas and selected topics.* In: *Advances in longitudinal and multivariate analysis in the behavioral sciences*, J.H.L. Oud and R.A.W. van Blokland-Vogelesang (eds), 91-104. Nijmegen: ITS.

IX van Eeuwijk F.A., and Keizer L.C.P. (1995). On the use of diagnostic biplots in model screening for genotype by environment tables. *Statistics and Computing* **5**, 141-153.

X van Eeuwijk, F.A. (1995). Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models. *Euphytica* **84**, 1-7.

XI van Eeuwijk F.A., Keizer L.C.P. Keizer, and Bakker J.J. (1995). Linear and bilinear models for

the analysis of multi-environment trials: II. An application to data from the Dutch Maize Variety Trials. *Euphytica* **84**, 9-22.

XII van Eeuwijk, F.A. (1995b). Multiplicative interaction in generalized linear models. *Biometrics* **51**, 1017-1032.

XIII van Eeuwijk, F.A., Mesterhazy, A., Kling, Ch. I., Ruckenbauer, P, Saur, L., Bürstmayr, H., Lemmens, M., Keizer, L.C.P., Maurin, N., and Snijders, C.H.A. (1995). Assessing non-specificity of resistance in wheat to head blight caused by inoculation with European strains of *Fusarium culmorum*, *F. graminearum* and *F. nivale*, using a multiplicative model for interaction. *Theoretical and Applied Genetics* **90**, 221-228.

XIV van Eeuwijk, F.A., Denis, J.-B., and Kang, M.S. (1995). *Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables.* In: *Genotype by environment interaction: New perspectives*, M.S. Kang and H.G. Gauch Jr. (eds), 15-49. Boca Raton: CRC-Press.

XV van Eeuwijk, F.A., and P.M. Kroonenberg. (1995). Multiplicative decompositions of interactions in three-way analysis of variance, with applications to plant breeding. Submitted.

References

Arntzen, F.K., and van Eeuwijk, F.A. (1992). Variation in resistance levels of potato genotypes and virulence level of potato cyst nematode populations. *Euphytica* **62**, 135-143.

Atkinson, A.C. (1982). Regression diagnostics, transformations and constructed variables. *Journal of the Royal Statistical Society B* **44**, 1-36.

Baker R.J., (1988). Tests for crossover genotype-environmental interactions. *Canadian Journal of Plant Science* **68**, 405-410.

Baril, C.P., Denis, J.-B., Wustman, R., and van Eeuwijk, F.A. (1995). Analysing genotype by environment interaction in Dutch Potato Variety Trials using factorial regression. *Euphytica* **82**, 149-155.

Basford, K.E., Kroonenberg, P.M., DeLacy, I.H., and Lawrence, P.K. (1991). Three-way methods for multiattribute genotype x environment data: an illustrated partial survey. *Field Crops Research* **27**, 131-157.

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

Carson, M.L. (1987). Assessment of six models of host-pathogen interaction in horizontal pathosystems. *Phytopathology* **77**, 241-246.

Coleman, J.S., McConaughay, K.D.M., and Ackerly, D.A. (1994). Interpreting phenotypic variation in plants. *Trends in Ecology and Evolution* **9**, 187-191.

Cornelius, P.L., Seyedsadr, M., and Crossa, J. (1992). Using the shifted multiplicative model to search for "separability" in crop cultivar trials. *Theoretical and Applied Genetics* **84**, 161-172.

Cornelius, P.L., Van Sandford, D.A., and Seyedsadr, M.S. (1993). Clustering cultivars into groups without rank-change interactions. *Crop Science* **33**, 1193-1200.

Corsten, L.C.A. and Denis, J.-B. (1990). Structuring interaction in two-way tables by clustering. *Biometrics* **46**, 207-215.

Cox, D.R. (1984). Interaction. *International Statistical Review* **52**, 1-31.

Cox, D.R. (1992). Causality: some statistical aspects. *Journal of the Royal Statistical Society A* **155**, 291-301.

Crossa, J., Cornelius, P.L., Seyedsadr, M., and Byrne, P. (1993). A shifted multiplicative model cluster analysis for grouping environments without gentypic rank change. *Theoretical and Applied Genetics* **85**, 577-586.

Denis, J.-B. (1988). Two-way analysis using covariates. *Statistics* **19**, 123-132.

Denis, J.-B. (1991). Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Revue de Statistique Appliquée* **39**, 5-24.

Denis, J.-B. and Gower, J.C. (1992). Biadditive models. *Technical Report*, pp. 33. Laboratoire de Biométrie, INRA-Versailles.

Denis, J.-B., and Gower, J.C. (1994). Biadditive models. Letter to the editor. *Biometrics* **50**, 310-311.

Denis, J.-B. and Moro, J. (1995). Multivariate generalizations for modeling two-way interaction; Part I: Defining and estimating models. *Biuletyn Oceny Odmian (Cultivar Testing Bulletin)*, in press.

Denis, J.-B., Piepho, H.-P. and van Eeuwijk, F.A. (1996). Heteroscedastic mixed factorial regression models for analysing genotype by environment data. To be submitted.

Doebley, J. (1993). Genetics, development and plant evolution. *Current Opinion in Genetics and Development* **3**, 865-872.

Eisenhart, C. (1947). The assumptions underlying the analysis of variance. *Biometrics* **3**, 1-21.

Ehrenberg, A.S.C. and Bound, J.A. (1993). Predictability and prediction. *Journal of the Royal Statistical Society A* **156**, 167-206.

Engel, B. and Keen, A. (1994). A simple approach for the analysis of generalized linear mixed models. *Statistica Neerlandica* **48**, 1-22.

Eskridge, K.M. (1990). Safety-first models useful for selecting stable cultivars. In: *Genotype-by-environment interaction and plant breeding*, M.S. Kang (ed), 151-168. Baton Rouge: Louisiana State University.

Finlay, K.W. and Wilkinson, G.N. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research* **14**, 742-754.

Fry, J.D. (1992). The mixed-model analysis of variance applied to quantitative genetics: biological meaning of parameters. *Evolution* **46**, 540-550.

Gabriel, K.R. (1971). Biplot display of multivariate matrices with application to principal component analysis. *Biometrika* **58**, 453-467.

Gabriel, K.R. (1978). Least squares approximation of matrices by additive and multiplicative models. *Journal of the Royal Statistal Society B* **40**, 186-196.

Gauch, H.G. Jr. (1988). Model selection and validation for yield trials with interaction. *Biometrics* **44**, 705-715.

Gauch, H.G. Jr. (1992). *Statistical analysis of regional yield trials: AMMI analysis of factorial designs*. Amsterdam: Elsevier.

Gogel, B.J., Cullis, B.R. and Verbyla, A.P. (1995). REML estimation of multiplicative effects in multi-environment variety trials. *Biometrics* **51**, 744-749.

Gollob, H.F. (1968). A statistical model which combines features of factor analysis and analysis of variance techniques. *Psychometrika* **33**, 73-115.

Gower, J. C. (1990). Three-dimensional biplots. *Biometrika* **77**, 773-785.

Gregorius, H.-R. and Namkoong, G. (1986). Joint analysis of genotypic and environmental effects.

Theoretical and Applied Genetics **72,** 413-422.

Gregorius, H.-R. and Namkoong, G. (1987). Resolving the dilemmas of interaction, separability, and additivity. *Mathematical Biosciences* **85,** 51-69.

Hammer, G.L. and Vanderlip R.L. (1989a). Genotype-by-environment interaction in grain sorghum I. Effects of temperature on radiation use efficiency. *Crop Science* **29,** 370-376.

Hammer, G.L. and Vanderlip, R.L. (1989b). Genotype-by-environment interaction in grain sorghum. III. Modeling the impact in field environments. *Crop science* **29,** 385-391.

Hammer, G.L., Vanderlip, R.L., Gibson, G., Wade, L.J., Henzel, R.G., Younger, D.R., Warren, J., and Dale, A.B. (1989). Genotype-by-environment interaction in grain sorghum. II. Effects of temperature and photoperiod on ontogeny. *Crop Science* **29,** 376-384.

Harville, D.A. (1978). Alternative formulations and procedures for the two-way mixed model. *Biometrics* **43,** 441-453.

Harville, D.A. (1991). Comment on Samuels, Casella, and McCabe, Interpreting blocks as random factors. *Journal of the American Statistical Association* **86,** 812-815.

Hocking, R.R. (1973). A discussion of the two-way mixed model. *American Statistician* **27,** 148-152.

Jansen, R.C. and Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136,** 1447-1455.

Kauffman, S.A. (1992). Origins of order in evolution: self-organisation and selection. In: *Theoretical biology: epigenetic and evolutionary order from complex systems*, B.C. Goodwin and P.T. Saunders (eds), 67-88. London: John Hopkins Press Ltd.

Kempthorne, O. (1975). Fixed and mixed models in the analysis of variance. *Biometrics* **31,** 473-486.

Kempton, R.A. (1984). The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science, Cambridge* **103,** 123-135.

Lane, P.W. and Nelder, J.A. (1982). Analysis of covariance and standardization as instances of prediction. *Biometrics* **38,** 613-621.

Lewontin, R.C. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics* **26,** 400-411.

Lloyd, E.A. (1987). Confirmation of ecological and evolutionary models. *Biology and Philosophy* **2,** 277-293.

Mandel, J. (1961). Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **56,** 878-888.

Mandel, J. (1969). The partitioning of interaction in analysis of variance. *Journal of Research of the National Bureau of Standards - B, Mathematical Sciences* **73,** 309-328.

McCullagh, P., and J.A. Nelder. (1989). *Generalized linear models.* London: Chapman and Hall.

Meyerowitz, E.M. (1994). Pattern formation in plant development: four vignettes. *Current Opinion in Genetics and Development* **4,** 602-608.

Muchow, R.C. and Carberry, P.S. (1993). Designing improved plant types for the semiarid tropics: agronomists' viewpoints. In: *Systems approaches for agricultural development.* F.W.T. Penning de Vries, P.S. Teng, and K. Metselaar (eds), 37-61. Dordrecht: Kluwer.

Nelder, J.A. (1977). A reformulation of linear models. *Journal of the Royal Statistical Society A* **140,** 48-76.

Nelder, J.A. (1994). The statistics of linear models: back to basics. *Statistics and Computing* **4,** 221-234.

Oman, S.D. (1991). Multiplicative effects in mixed model analysis of variance. *Biometrika* **78,** 729-739.

Palanisamy, S., Penning de Vries, F.W.T., Mohandass, S., Thiyagarajan, T.M., and Abdul Kareem, A. (1993). In: *Systems approaches for agricultural development.* F.W.T. Penning de Vries, P.S. Teng, and K. Metselaar (eds), 63-75. Dordrecht: Kluwer.

Piepho, H.-P. (1995). Detecting and handling heteroscedasticity in yield trial data. *Communications in Statistics B: Simulation and Computation* **24,** 243-274.

Ritchie, J.T. (1993). Genetic specific data for crop modelling. In: *Systems approaches for agricultural development.* F.W.T. Penning de Vries, P.S. Teng, and K. Metselaar (eds), 77-93. Dordrecht: Kluwer.

Robinson, G.K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science* **6,** 15-51.

Salmon, W.C. (1984). *Scientific explanation and the causal structure of the world.* Princeton: Princeton University Press.

Scheffé, H. (1959). *The Analysis of Variance.* New York: Wiley.

Schmid, B. (1992). Phenotypic variation in plants. *Evolutionary trends in plants* **6,** 45-60.

Seyedsadr, M., and Cornelius, P.L. (1992a). Shifted multiplicative models for non-additive two-way tables. *Communications in Statistics B: Simulation and Computation* **21,** 807-832.

Seyedsadr, M., and Cornelius, P.L. (1992b). Computational procedures for estimation and hypothesis testing in analyzing two-way tables with interaction and no replication. In: *Proceedings Compstat 1992,* Y. Dodge and J. Whittaker (eds), 89-94, Heidelberg: Physica-Verlag.

Shukla, G.K. (1982). Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* **29,** 237-245.

ter Braak, C.J.F. and Looman, C.W.N. (1994). Biplots in reduced-rank regression. *Biometrical Journal* **36,** 983-1003.

Tukey, J.W. (1949). One degree of freedom for non-additivity. *Biometrics* **5,** 232-242.

van Eeuwijk, F.A., and Kroonenberg, P.M. (1995). The simultaneous analysis of genotype by environment interaction for a number of traits using three-way multiplicative modelling. *Biuletyn Oceny Odmian (Cultivar Testing Bulletin),* in press.

van Fraassen, B.C. (1980). *The scientific image.* Oxford: University Press.

van Fraassen, B.C. (1989). *Laws and symmetry.* Oxford: Clarendon Press.

van Noordwijk, M., Dijksterhuis, G.H., and van Keulen, H. (1995). Risk management in crop production and fertilizer use with uncertain rainfall; how many eggs in which baskets. *Netherlands Journal of Agricultural Science* **42,** 249-269.

Wallace, D.H., Baudoin, J.P., Beaver, J., Coyne, D.P., Halseth, D.E., Masaya, P.N., Munger, H.M., Myers, J.R., Silbernagel, M., Yourstine, K.S., and Zobel, R.W. (1993a) Improving efficiency of breeding for higher crop yield. *Theoretical and Applied Genetics* **86,** 27-40.

Wallace, D.H., Yourstone, K.S., Masaya, P.N., and Zobel, R.W. (1993). Photoperiod gene control over partitioning between reproductive and vegetative growth. *Theoretical and Applied Genetics* **86,** 6-16.

Wallace, D.H., Zobel R.W., and Yourstone, K.S. (1993b). A whole-system reconsideration of paradigms about photoperiod and temperature control of crop yield. *Theoretical and Applied Genetics* **86,** 17-26.

Wermuth, N. and Lauritzen, S.L. (1990). On substantive research hypotheses, conditional independence graphs and graphical chain models. *Journal of the Royal Statistical Society, Series B* **52,** 21-50.

Yates, F., and Cochran, W.G. (1938). The analysis of groups of experiments. *Journal of Agricultural Science, Cambridge* **28,** 556-580.

Summary

The phenotype of a plant is the ensemble of its observable characteristics. The phenotype is the result of an interplay between the genetic constitution, the genotype, and the environmental conditions over time. A central theme of research in plant breeding is the description and prediction of phenotypic responses under changing environmental conditions. Different genotypes often react differently to the same environmental change. Phenotypic responses then converge, diverge, or intersect in dependence on the environment. Genotype by environment interaction (GEI) is the name for the concept that covers all instances of differential phenotypic responses, i.e. whenever phenotypic responses are non-parallel GEI is said to occur.

Essential for an adequate statistical description of GEI are the facilities for modelling non-parallelism of responses. Traditional analysis of variance (ANOVA) models allow for non-parallelism by non-additivity parameters. The additive ANOVA model, without non-additivity parameters, implies parallel responses. Deviations from parallelism induce, within an ANOVA context, non-additivity parameters. These parameters appear as higher dimensional tables of multiply indexed additive parameters. Non-additivity address all types of non-parallelism simultaneously. Testing for GEI by means of testing for non-additivity in ANOVA models may therefore be not very powerful, while description of GEI will be little parsimonious and hard to interpret. Interpretational problems arise because of the multitude of multiply indexed parameters that need interpretation and the neglect of explicit external genotypic and environmental information in the ANOVA formulation for interaction.

This thesis tries to improve on ANOVA descriptions for GEI. The high dimensional tables of multiply indexed non-additivity parameters are replaced by lower dimensional tables built on multiplicative formulations for GEI that contain only single indexed parameters. The multiplicative parameters either bear a direct relation to external information, or otherwise generally can be made to bear a close relation to external information. Thus, there are gains in parsimony, specificity and interpretability.

Two main types of multiplicative models for interaction are distinguished; factorial regression models and multilinear models. Factorial regression models are multiplicative models that incorporate external concomitant information on either or both of genotypes and environments. Because they are ordinary linear models, estimation and testing present no complications. Parameters can be interpreted as 1) genotypic sensitivities to environmental

covariables; 2) environmental potentialities that combine multiplicatively with genotypic covariables, and; 3) scaling constants for products of genotypic and environmental covariables. Factorial regression models provide the means for testing hypotheses about the biological mechanisms underlying GEI, and thus facilitate biological interpretation. This is a strong argument for their use. This thesis is mainly concerned with the application of the two-way form of factorial regression in fixed ANOVA models, it only touches upon higher way factorial regression, mixed factorial regression and generalized factorial regression.

Multilinear models replace non-additivity terms in ANOVA models by sums of multiplicative terms, without having recourse to external concomitant information. The product terms consist exclusively of parameters that are estimated using multiplicative decompositions of tables of non-additivity parameters. These tables are approximated by tables of lower rank by retaining only the most important terms of the decompositions, resulting in parsimonious descriptions of GEI. Statistical inference regarding the number of terms to retain requires special procedures, because standard linear theory is inapplicable. Interpretation of multilinear parameters can be in terms of genotypic sensitivities to theoretical environmental variables that are best in the sense of maximizing interaction between genotypes. However, there is a danger in interpreting multilinear terms individually. The models provide low rank fits to originally higher dimensional tables. Therefore, the terms of a low rank fit should be inspected simultaneously. For this, the graphical device of the biplot is an indispensable aid. In a biplot genotypes and environments are represented as vectors starting at the origin, and with the end point coordinates determined by the values for the genotypic and environmental interaction parameters. Distance and inner product relations between the genotypic and environmental vectors furnish the basis for interpretational rules. The patterns in biplots often lead to biologically interesting conclusions. It always pays off to fit low rank approximations to tables of non-additivity parameters and to visually display the results in biplots. This procedure, minimally guarantees an increase in parsimony, and, more often than not, an even more valuable result is obtained when a biological interpretation of the GEI can be given.

The most popular form of multilinear models is the bilinear reformulation of two-way non-additivity in fixed ANOVA models, but extensions to higher way non-additivity are available, e.g. quadrilinear models for three-way non-additivity. Bilinear models need not be restricted to identity links and constant variance functions. Generalized bilinear models have been developed and shown to be useful.

A noteworthy descendant from a cross between the classes of factorial regression models and bilinear models is the class of reduced rank factorial regression models (strictly speaking, also bilinear), which combines the advantage of directly modelling GEI on external variables with the advantage of visually displaying GEI in the form of biplots.

This thesis contains four categories of papers; 1) those that emphasize the use of especially factorial regression and bilinear models to arrive at biologically interesting conclusions for GEI (chapters II, III, IV, VI, XI and XIII); 2) those that emphasize the

theoretical aspects of a method or model (V, IX, XII and XV); 3) those that principally review models (VII, X and XIV), and; 4) those that aim at conceptual clarification (VIII and XVI).

The applied papers are constructed around practical problems, but enough attention is given to model descriptions and theoretical elaborations to make the papers largely self-contained. The papers illustrate how multiplicative models for GEI can be implemented successfully for very different variables in very different crops; field emergence in white cabbage (II), yield and quality characteristics in sugar beet (III), seed yield in perennial ryegrass (IV), *Fusarium* head blight incidence in wheat (VI, XIII), and dry matter content in maize (XI). The chapters II, III, IV and XI illustrate the power of an approach based on the complementarity of factorial regression and bilinear models. The chapters VI and XIII show how for some questions bilinear models alone can suffice for an analysis, i.e. patterns in plots of bilinear environmental parameters reveal the specificity of resistances.

Principally theoretically orientated considerations and elaborations of GEI are present in both the theory and the review chapters. The extension of (two-way) factorial regression to its reduced rank equivalent is presented in chapter V (under the name of redundancy analysis). The model is described, a test is given for determining the dimensionality of the interaction, and an example is included in which nitrate content in lettuce is analysed. The chapters VII and X contain brief reviews of the theory for reduced rank factorial regression, with extra attention for biplots in chapter X. Besides reduced rank regression, chapter VII discusses other multiplicative models for two-way interactions and develops some thoughts on the merits of bilinear and factorial regression models in comparison to those of the regression on the mean model. Chapter X describes in some detail the possibilities of linear models, including mixed models and factorial regression, besides those for bilinear models including reduced rank factorial regression. The models described in chapter X are jointly illustrated in chapter XI.

Chapter XIV presents an exhaustive review of all types of factorial regression models that are available for modelling GEI in two-way tables. The differentiating features that were chosen for categorizing the models were: 1) whether the genotypic dimension, the environmental dimension, or both were modelled; 2) the type of covariables that was used (continuous/ discrete); 3) whether random model terms were included, and; 4) whether rank constraints were imposed.

Theoretical generalization of bilinear models determines the content of the chapters XII and XV. In chapter XII the limitations of bilinear models of identity link and constant variance are removed. Generalized bilinear models are introduced, their structure is discussed, and an algorithm is given to estimate the parameters. The principles are illustrated for the numbers of potato cyst nematodes on potatoes and for *Fusarium* head blight incidences in wheat. Chapter XV shows one way of proceeding from multilinear models for two-way ANOVA interactions to multilinear models for three-way ANOVA interactions. A quadrilinear model for three-way interaction is presented, with an algorithm for estimation, and applications to, again, *Fusarium* head blight incidences in wheat, and dry matter content in

maize.

Parsimony and interpretability are central themes throughout the thesis. Chapter IX describes the theory and evaluation by means of simulation of an informal graphical method for finding parsimonious models for two-way tables based on the geometrical properties of some rank-two and three models. Interpretability and related conceptional questions determine the discussions in the opinion papers of the chapters VIII and XVI. It is concluded that GEI is recommendably approached by regression based methods incorporating explicit genotypic and environmental covariables. Multilinear models provide a welcome complement. Their exclusive application may be sufficient for answering some types of research questions.

Samenvatting

Het fenotype van een plant bestaat uit het geheel van zijn observeerbare kenmerken. Het fenotype is het resultaat van de interactie tussen de genetische constitutie, het genotype, en de omgeving. Een centraal thema binnen de plantenveredeling behelst de beschrijving en voorspelling van fenotypische responsies in relatie tot veranderende milieu-omstandigheden. Verschillende genotypen reageren vaak verschillend op een zelfde milieuverandering en fenotypische responsies kunnen dan convergeren, divergeren of elkaar snijden. Genotype-bij-milieu-interactie is de naam voor het concept dat alle instanties van differentiële fenotypische responsies insluit. Voor elk geval waarvoor fenotypische responsies niet parallel zijn wordt gesteld dat genotype-bij-milieu-interactie optreedt.

De faciliteiten voor de modellering van non-parallellisme bepalen de mate waarin een adequate statistische beschrijving van genotype-bij-milieu-interactie kan worden gegeven. In traditionele variantie-analysemodellen wordt non-parallellisme gemodelleerd met behulp van non-additiviteitsparameters. Geldigheid van het additieve variantie-analysemodel, zonder non-additiviteitsparameters, impliceert parallelle responsies. Afwijkingen van parallellisme induceren in een variantie-analysecontext non-additiviteitsparameters. Deze parameters treden op als hoogdimensionale tabellen van meervoudig geïndiceerde additieve parameters. Non-additiviteit omsluit alle typen van non-parallellisme tegelijk. Daardoor kunnen testen op de aanwezigheid van genotype-bij-milieu-interactie middels testen op non-additiviteit wel eens weinig krachtig uitvallen. De bijbehorende beschrijvingen van genotype-bij-milieu-interactie blinken niet uit in spaarzaamheid en interpreteerbaarheid. Interpretatieproblemen ontstaan door de veelheid aan meervoudig geïndiceerde parameters die interpretatie vereisen en het negeren van expliciete, externe genotypische en omgevingsinformatie in de variantie-analyseformulering van interactie.

In dit proefschrift worden alternatieven aangedragen voor de variantie-analysebeschrijvingen van genotype-bij-milieu-interactie. De hoogdimensionale tabellen van meervoudig geïndiceerde non-additiviteitsparameters worden vervangen door lager dimensionale tabellen gebaseerd op multiplicatieve formuleringen voor genotype-bij-milieu-interactie waarin alleen enkelvoudig geïndiceerde parameters figureren. Deze multiplicatieve parameters hebben of direct betrekking op externe informatie, of ze kunnen er op indirecte wijze nauw mee in verband gebracht worden.

Twee typen multiplicatieve modellen voor interactie worden onderscheiden; factoriële

regressiemodellen en multilineaire modellen. Factoriële regressiemodellen zijn multiplicatieve modellen waarin externe informatie met betrekking tot genotypen en/of milieus geïncorporeerd is. Omdat het gewone lineaire modellen betreft creëren schatten en testen geen problemen. Parameters kunnen als volgt worden geïnterpreteerd: 1) als genotypische gevoeligheden met betrekking tot milieucovariabelen; 2) als milieupotenties die multiplicatief combineren met genotypische covariabelen; 3) als schaalconstanten voor produkten van genotypische en milieucovariabelen. Factoriële regressiemodellen verschaffen de mogelijkheid hypotheses te testen betreffende de biologische mechanismen die aan genotype-bij-milieu-interactie ten grondslag liggen. De biologische interpretatie van genotype-bij-milieu-interactie wordt hiermee bevorderd, wat een belangrijk argument is voor het gebruik van deze modellen. In het proefschrift wordt vooral aandacht geschonken aan de twee-weg vorm van factoriële regressie, in variantie-analysemodellen met vaste effecten. Andere verschijningsvormen als meer-weg factoriële regressie, gemengde factoriële regressie en gegeneraliseerde factoriële regressie komen slechts terloops aan de orde.

In multilineaire modellen zijn de non-additiviteitsparameters van variantie-analysemodellen vervangen door sommen van multiplicatieve termen, waarbij deze termen nu niet verwijzen naar externe informatie. De produkttermen bestaan uitsluitend uit parameters verkregen middels multiplicatieve decomposities van non-additiviteitstabellen. Deze tabellen worden benaderd door tabellen van lagere rang door behoud van alleen de leidende termen van de decomposities. Dit resulteert in spaarzame beschrijvingen van genotype-bij-milieu-interactie. Omdat de standaard lineaire modeltheorie niet geldig is voor multilineaire modellen zijn speciale procedures vereist voor statistische gevolgtrekkingen, zoals het bepalen van het aantal multiplicatieve termen dat gehandhaafd moet worden voor een adequate beschrijving van de genotype-bij-milieu-interactie. De interpretatie van multilineaire parameters kan gesteld worden in termen van genotypische gevoeligheden ten aanzien van theoretische milieuvariabelen. De theoretische milieuvariabelen maximaliseren de interactie tussen de genotypen. Er kleeft echter een gevaar aan het individueel interpreteren van multilineaire termen. De modellen geven lage-rangaanpassingen aan hoger dimensionale tabellen. De termen van de lage-rangaanpassing moeten daarom simultaan bekeken worden. Een onmisbaar grafisch hulpmiddel daarbij is de biplot. In een biplot worden genotypen en milieus gerepresenteerd als vectoren die beginnen in de oorsprong, terwijl de eindpuntcoördinaten vastgelegd worden door de waarden van de genotypische en milieu-interactieparameters. Interpretatieregels volgen uit de afstands- en inproduktrelaties tussen de genotypische en milieuvectoren. De in biplots aanwezige patronen leiden vaak tot biologisch interessante conclusies. Het loont onder alle omstandigheden de moeite om lage-rangaanpassingen uit te voeren op tabellen van non-additiviteitsparameters en de resultaten te visualiseren in biplots. Deze werkwijze garandeert minimaal een verhoogde spaarzaamheid, maar nog waardevoller is het resultaat wanneer een biologische interpretatie gegeven kan worden aan de genotype-bij-milieu-interactie.

De populairste vorm van multilineaire modellen is de bilineaire herformulering van twee-weg non-additiviteit in variantie-analysemodellen met vaste effecten. Uitbreidingen naar meer-weg non-additiviteit zijn echter ook beschikbaar, bijvoorbeeld quadrilineaire modellen voor drie-weg non-additiviteit. Bilineaire modellen behoeven niet uitsluitend op te treden met identiteitslink en constante variantie. Gegeneraliseerde bilineaire modellen zijn ontwikkeld en werden met succes toegepast.

Een opvallende nakomeling van een kruising tussen de klasse van de factoriële regressiemodellen en de klasse van de bilineaire modellen is de klasse van de gereduceerde-rang factoriële regressiemodellen (strikt genomen ook bilineair), die de voordelen combineert van directe modellering van de genotype-bij-milieu-interactie op externe covariabelen en visualisering van de genotype-bij-milieu-interactie in de vorm van biplots.

Dit proefschrift bevat vier categorieën hoofdstukken: 1) die waarin vooral de toepassing van factoriële regressiemodellen en bilineaire modellen leidt tot biologisch interessante conclusies met betrekking tot de genotype-bij-milieu-interactie (hoofdstukken II, III, IV, VI, XI en XIII); 2) die welke de theoretische aspecten van een model of methode benadrukken (V, IX, XII en XV); 3) die welke hoofdzakelijk een overzicht geven van modellen (VII, X en XIV); 4) die waarin conceptuele verheldering wordt nagestreefd (VIII en XVI).

De toegepaste hoofdstukken zijn opgebouwd rondom praktische problemen, maar er wordt voldoende aandacht gegeven aan modelbeschrijvingen en theoretische uitwerkingen om de hoofdstukken onafhankelijk van elkaar leesbaar te laten zijn. Deze hoofdstukken tonen hoe multiplicatieve modellen voor genotype-bij-milieu-interactie succesvol gebruikt kunnen worden voor de analyse van uiteenlopende variabelen in uiteenlopende gewassen; veldopkomst in witte kool (II), opbrengst- en kwaliteitskarakteristieken in suikerbiet (III), zaadopbrengst in Engels raaigras (IV), *Fusarium*-aantasting in tarwe (VI, XIII) en droge stofgehalte in maïs (XI). De hoofdstukken II, III, IV en XI illustreren de kracht van een benadering die gebaseerd is op de complementariteit van factoriële regressie en bilineaire modellen. De hoofdstukken VI en XIII laten zien hoe voor sommige vragen volstaan kan worden met de toepassing van bilineaire modellen alleen. De patronen in de plots van de bilineaire milieuparameters onthullen hier de non-specificiteit van de resistenties.

Voornamelijk theoretisch georiënteerde beschouwingen en uitwerkingen van genotype-bij-milieu-interactie vullen de theorie- en overzichtshoofdstukken. De uitbreiding van (twee-weg) factoriële regressie naar zijn gereduceerde-rangequivalent wordt beschreven in hoofdstuk V. Het model wordt gepresenteerd, een test wordt gegeven voor de bepaling van de dimensionaliteit van de interactie en in een voorbeeld worden nitraatgehaltes in sla geanalyseerd. De hoofdstukken VII en X bevatten beknopte overzichten van de theorie voor gereduceerde-rang factoriële regressie, met speciale aandacht voor biplots in hoofdstuk X. In hoofdstuk VII worden verder ook ander multiplicatieve modellen voor twee-weg interacties bediscussieerd en worden enkele gedachten ontwikkeld over de sterke punten van de bilineaire en factoriële regressiemodellen in vergelijking tot die van het regressie-op-het-gemiddelde-

model. Hoofdstuk X beschrijft de mogelijkheden van lineaire modellen, inclusief gemengde modellen en factoriële regressiemodellen, samen met die van bilineaire modellen, inclusief gereduceerde-rang factoriële regressie. Hoofdstuk XI laat zien hoe de modellen van hoofdstuk X in de praktijk gezamenlijk en aanvullend toegepast kunnen worden.

Hoofdstuk XIV presenteert een uitputtend overzicht van alle typen factoriële regressiemodellen die beschikbaar zijn voor de modellering van genotype-bij-milieu-interactie in twee-weg tabellen. De onderscheidende kenmerken die gekozen werden om de modellen te categoriseren waren: 1) het modelleren van de genotypische dimensie, de milieudimensie, of beide; 2) het type covariabelen dat werd gebruikt (continu of discreet); 3) het al dan niet opnemen van extra toevalstermen; 4) het al dan niet opleggen van rangrestrikties aan de matrix van regressiecoëfficiënten.

De inhoud van de hoofdstukken XII en XV wordt bepaald door theoretische generalisaties van bilineaire modellen. In hoofdstuk XII worden de beperkingen van een identiteitslink en constante variantie opgeheven. Gegeneraliseerde bilineaire modellen worden geïntroduceerd, hun structuur wordt bediscussieerd en er wordt een schattingsalgorithme voor de parameters gegeven. De principes worden toegelicht aan de hand van twee voorbeelden. Het eerste analyseert het aantal aardappelcyste-aaltjes op aardappelen, het tweede *Fusarium*-aantasting in tarwe. Hoofdstuk XV demonstreert hoe men van multilineaire modellen voor twee-weg interacties naar multilineaire modellen voor drie-weg interacties kan geraken. Een quadrilineair model voor drie-weg interactie wordt gepresenteerd, met een schattingsalgorithme en twee toepassingen: opnieuw *Fusarium*-aantasting in tarwe en droge stofgehalte in maïs.

Spaarzaamheid en interpreteerbaarheid vormen de centrale thema's in dit proefschrift. Hoofdstuk IX beschrijft de theorie en evaluatie door middel van simulatie van een informele, grafische methode voor het vinden van spaarzame modellen voor twee-weg tabellen. De methode is geconstrueerd op basis van de geometrische eigenschappen van sommige rang-twee en rang-drie modellen. Interpreteerbaarheid en gerelateerde conceptuele kwesties leiden de discussies in de opiniërende hoofdstukken VIII en XVI. De conclusie luidt dat genotype-bij-milieu-interactie bij voorkeur wordt geanalyseerd met op regressie gebaseerde methoden. De opname van expliciete genotypische en milieucovariabelen in het statistische model is daarbij cruciaal. Multilineaire modellen vormen een welkome aanvulling. In voorkomende gevallen kan zelfs met hun exclusieve gebruik worden volstaan.

Curriculum vitae

Fredericus Antonius van Eeuwijk (Fred) werd geboren op 25 Februari 1959 te Kerkdriel, doorliep de lagere school in Velddriel en vervolgens het Atheneum-B aan het Buys Ballot College te Zaltbommel. Hij startte in 1976 een studie biologie aan de toenmalige Rijksuniversiteit Utrecht en behaalde in 1980 het kandidaatsexamen biologie B1* (biologie, ecologische variant). Zijn doctoraalprogramma bevatte de hoofdvakken ethologie en populatiegenetica. Daarnaast werd de onderwijsbevoegdheid voor biologie behaald, alsmede het propaedeuse diploma filosofie (1983; cum laude). De studie biologie voltooide hij in 1985 (cum laude). Aansluitend werd hij aangesteld als statisticus bij de Stichting voor Plantenveredeling, dat later opging in het DLO-Centrum voor Plantenveredelings- en Reproduktieonderzoek. In 1990 werd een begin gemaakt met het onderzoek dat leidde tot dit proefschrift.

Dankwoord

Velen hebben bijgedragen aan dit proefschrift. Nogmaals wil ik hier alle co-auteurs bedanken, alsmede iedereen uit de dankwoorden van de individuele hoofdstukken. Enkele personen hebben een dermate grote bijdrage geleverd dat zij een aparte vermelding verdienen. In de allereerste plaats is dat Paul Keizer. Zonder zijn tomeloze inzet op velerlei gebied had dit proefschrift zeker nog wel een tijdje op zich laten wachten. Zijn bijzondere verdiensten op zowel het redactionele als vakinhoudelijke vlak kunnen op vele plaatsen in dit proefschrift teruggevonden worden. Bovendien heeft hij al die jaren voor een aangenaam werkklimaat gezorgd. Een tweede persoon die hier genoemd moet worden betreft Jean-Baptiste Denis. Hij was degene die me wegwijs heeft gemaakt in de theoretische aspecten van de genotype-bij-milieu-interactie. Zijn scherpzinnigheid en generalisatievermogen zullen me ook in de toekomst tot voorbeeld blijven strekken. De laatste persoon die ik hier bij naam wil noemen is Piet Stam, vanwege het geschonken vertrouwen in de door mij uitgezette onderzoekslijnen en de rustgevende manier waarop de incidentele voortgangsbesprekingen plaatsvonden.

Vele anderen hebben minder tastbaar en misschien soms juist niet bijgedragen. Desalniettemin wil ik de volgende groepen bedanken; CPRO-DLO (SVP), de collega's van de afdeling Populatiebiologie, familie en vrienden, y el departamento de Bruselas.