Published August 30, 2016

RESEARCH

What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype × Environment Interactions?

Fred A. van Eeuwijk,* Daniela V. Bustos-Korts, and Marcos Malosetti

ABSTRACT

A good statistical analysis of genotype × environment interactions (G × E) is a key requirement for progress in any breeding program. Data for G × E analyses traditionally come from multi-environment trials. In recent years, increasingly data are generated from managed stress trials, phenotyping platforms, and high throughput phenotyping techniques in the field. Simultaneously, and complementary to the phenotyping, more elaborate genotyping and envirotyping occur. All of these developments further increase the importance of a sound statistical framework for analyzing G × E. This paper presents considerations on such a framework from the point of view of the choices that need to be made with respect to the content of short academic courses on statistical methods for G × E. Based on our experiences in teaching statistical methods to plant breeders, for specialized G × E courses between three and 5 d are reserved. The audience in such courses includes MSc students, PhD students, postdocs, and researchers at breeding companies. For such specialized courses, we propose a collection of topics to be covered. Our outlook on G × E analyses is two-fold. On the one hand, we see the G × E problem as the building of predictive models for genotype-specific reaction norms. On the other hand, the G × E problem consists in the identification of suitable variance-covariance models to describe heterogeneity of genetic variance and correlations across environments. Our preferred class of statistical models is the class of mixed linear-bilinear models. These statistical models allow us to answer breeding questions on adaptation, adaptability, stability, and the identification and subdivision of the target population of environments. By a citation analysis of the literature on G × E, we show that our preference for mixed linear-bilinear models for analyzing G × E is supported by recent trends in the types of methods for G × E analysis that are most frequently cited.

F.A. van Eeuwijk, D.V. Bustos-Korts and M. Malosetti, Biometris, Wageningen Univ., PO Box 16, 6700 AA Wageningen, The Netherlands. Received 17 June 2015. Accepted 25 May 2016. Assigned to Associate Editor Jean-Luc Jannink. *Corresponding author (fred.vaneeuwijk@wur.nl).

Abbreviations: AMMI, additive main effects and multiplicative interactions; G × E, genotype × environment interaction; GGE, genotypic main effects and genotype × environment interactions; MET, multi-environment trial; QTL, quantitative trait locus; REML, REstricted Maximum Likelihood; SVD, singular value decomposition; TPE, target population of environments; TPG, target population of genotypes; VCOV, variance-covariance matrix.

Setting Up a Course on G × E

Motivation

Approximately 15 yr ago, the prediction of phenotypes from high-density marker information was recognized as a potential game changer in animal breeding (Meuwissen et al., 2001). In plant breeding, the current development of high throughput genotyping techniques alongside with similar techniques for phenotyping and envirotyping (Cobb et al., 2013; Araus and Cairns, 2014; Cooper et al., 2014b) provides opportunities for large scale phenotypic predictions of new, and therefore, untested genotypes in new environments under a wide spectrum of genotype × environment interaction (G × E) scenarios (Burgueño et al., 2008; Dawson et al., 2013; Heslot et al., 2013; Jarquín et al., 2013; Bustos-Korts et al., 2016; Malosetti et al., 2016). The increased volumes of phenotypic, genotypic and environmental data give a stimulus to the development of new statistical approaches for more precise description and prediction of G × E phenomena. A better modeling of G × E will undeniably contribute to a higher efficiency of

Published in Crop Sci. 56:2119–2140 (2016).
doi: 10.2135/cropsci2015.06.0375

© Crop Science Society of America | 5585 Guilford Rd., Madison, WI 53711 USA
All rights reserved.

breeding programs. In the light of the current developments, it is obvious that the study of G × E will become even more important in the near future than it was already in the past. Future generations of students and researchers in plant breeding will require substantial training in the statistical aspects of G × E. In this paper, we present and discuss our views on the topics that need to be covered in a course on G × E as well as some ideas about how to teach such a course.

From our experience, for specialized courses about the statistical aspects of G × E, typically a period between three to 5 d is reserved. The population of students for such courses covers the spectrum from MSc students to researchers in plant breeding companies, with in-between PhD students and postdocs. For this paper, we base ourselves on the following experiences. We gave a number of courses as part of the Integrated Breeding Platform Multi-Year Course program in the years 2012, 2013, and 2014 (https://www.integratedbreeding.net). The students in these courses came from Africa and Asia and ranged from PhD students in plant biology to experienced plant breeders in charge of breeding programs. A course with MSc students from the Mediterranean area was part of the MSc Plant Breeding program at the Mediterranean Agronomic Institute of Zaragoza (http://masters.iamz.ciheam.org/en/plantbreeding/). Further courses were given to MSc and post-doctoral students in plant breeding at Wageningen University (http://www.wageningenur.nl/biometris) and to MSc students in applied statistics at Leiden University (http://en.mastersinleiden.nl/programmes/statistical-science-for-the-life-and-behavioural-sciences/en/introduction). Finally, various in-house courses were organized for employees of international plant breeding companies.

Prerequisites

The target audience for our courses consists of plant breeding students who should be familiar with the following areas and concepts:

- Quantitative genetics (estimation of genetic and environmental variances, heritability, genetic correlation, response to selection, correlated response)
- Statistical design of plant breeding trials (obligatory: completely randomized designs, randomized complete block designs, incomplete block designs; desired: resolvable designs, cyclic designs, row-and-column designs, p-rep designs, augmented designs, split plot designs)
- Statistical analysis of plant breeding trials (one-way and two-way analysis of variance (ANOVA), development and testing of contrasts and interactions, regression, analysis of covariance, one- and two-way linear mixed models, inference for fixed and random terms)

Preferentially, students have followed a course on the design and analysis of individual plant breeding trials before going to a G × E course. On the design side, students should be able to choose an appropriate design for plant breeding experiments given particular objectives (e.g., estimation of heritability versus comparison of a small number of genotypes) and limitations (number of replicates, rows, and columns, spatial variation, trends). On the analysis side, students should be able select an appropriate model for an experiment, making choices on which terms to take fixed and which to take random, and assessing the necessity to include individual model terms by F-tests and Wald tests for fixed terms and log-likelihood ratio (deviance) tests for random terms.

Familiarity with quantitative trait locus (QTL) mapping is useful, because it allows G × E analyses to be connected directly to QTL × environment (QTL × E) analyses. For QTL mapping, it is best, when both linkage and linkage disequilibrium mapping have been discussed, for single and multiple traits and environments, preferentially within a mixed model framework. Simple examples of genome-enabled prediction methods can be included in a QTL mapping course as well. From our perspective, an important objective of QTL analysis is the identification of the genetic basis of G × E in the form of QTL × environment interaction (QTL × E).

To finish off the quantitative education of plant breeding students, a course in decision support is desired in which students learn how to integrate information from statistical analyses and formulate breeding strategies. Specific topics in such a course are marker assisted back crossing, gene pyramiding, marker assisted selection, genomic selection, and various types of index selection. A course on G × E will benefit from knowledge on decision support, although a decision support course is not an absolute prerequisite for a G × E course.

In our three-to-five day course on G × E, the emphasis is on the formulation and building of genotype-to-phenotype prediction models. From the identified genotype-to-phenotype models, we estimate genotype-specific statistics for mean performance under general and specific environmental conditions besides genotype-specific stability and risk measures. Selection strategies for genotypic means, sensitivities, stability variances and risks can be part of the G × E course as long as they concern phenotypic selection. We would prefer to discuss marker assisted selection strategies for such G × E related statistics in a separate decision support course.

G × E Concepts and Perspectives

Before we describe the statistical-technical details of our proposal for a statistical course on G × E for future plant breeders, we want to introduce a number of breeding concepts that are useful for a better communication between plant breeders and statisticians. These concepts help to define pertinent breeding questions in terms that allow unequivocal translation to statistical models and

parameters. We acknowledge that the definitions of the breeding concepts below may look biased toward the direction of statistical clarity at the expense of biological width. We have on purpose narrowed down breeding definitions to guarantee one-to-one relations to statistical parameters.

Target Population of Genotypes and Target Population of Environments

As a preliminary to models for G × E and breeding concepts related to G × E (adaptation, adaptability, stability, etc.), it is useful to introduce the concepts of target population of genotypes (TPG) and target population of environments (TPE). The TPG and TPE define the set of genotypes and environments for which we want our inference and predictions to be valid and precise. The TPG contains all possible genotypes we hope to develop and grow the coming years. Statistically speaking, we aim at coincidence of the TPG and the genetic design space of our prediction models. In the genomic selection literature, the target population of genotypes coincides with the notion of the set or population of selection candidates (Jannink et al., 2010; Schulz-Streeck et al., 2012; Albrecht et al., 2014). The TPE delineates the future growing conditions of the genotypes in the TPG (Comstock, 1977; Cooper and Hammer, 1996; Cooper et al., 2014b). The TPE can be defined by geography, soil and meteorological conditions, management choices, and the incidence of biotic and abiotic stresses. We want the TPE to be reflected in the environmental design space of our prediction models.

As the phenotype is an integrated outcome of interactions between genetic and environmental factors during development, TPG and TPE cannot be chosen as to be independent. To give an example for abiotic stress, if we define a TPE on a geographical basis that includes drought and well-watered conditions, we wish to develop genotypes that perform well under both drought stress and well-watered conditions, or, in other words, the TPG consists of genotypes with wide adaptation. However, if we want to interpret drought stressed conditions as a TPE by itself, the composition of the TPG will have to change in reaction to the redefinition of the TPE. Therefore, the width of the TPE has consequences for the definition of the TPG and vice versa. For biotic stresses, the same arguments will hold when we replace the drought stress in the above example with infection pressure for a particular disease.

Reaction Norm

For individual genotypes, we want to describe their phenotypic behavior across the full TPE and therefore we introduce the concept of the reaction norm: the genotype-specific functional relationship between phenotype and environmental gradient(s) (Woltereck, 1909; DeWitt and Scheiner, 2004). In practice, environmental gradients are sampled in a limited number of experiments. The observations made in those discrete environments are called character states (Schlichting and Pigliucci, 1998; Pigliucci, 2001).

For phenotypic prediction across a range of environmental conditions, we need to fit statistical models that represent the reaction norms of individual genotypes, that is, the main environmental drivers for phenotypic differences need to be identified together with suitable functional forms for the reaction norms. Phenotypic data can come from a series of field trials that represent a draw from the TPE. Such a draw from the TPE is often equivalent to or part of a multi-environment trial (MET) (Smith et al., 2001, 2005; van Eeuwijk et al., 2010). More informative data for modeling reaction norms can come from managed stress trials (Cooper et al., 2014a) and phenotyping platforms (Tardieu and Tuberosa, 2010; Cobb et al., 2013; Araus and Cairns, 2014; Kuijken et al., 2015).

Adaptedness and Adaptation

Within the framework of reaction norms, a genotype shows adaptedness when its reaction norm is superior to that of a standard genotype or when it is close to that of an ideotype (Van Oijen and Höglind, 2015). In the plant breeding literature, adaptedness, a state, is not always distinguished from adaptation, a process. For example, it is common to talk about wide adaptation for genotypes showing adaptedness across the full TPE, versus narrow adaptation for genotypes showing adaptedness for part of the TPE (Ceccarelli, 1989; Ceccarelli et al., 1996; Braun et al., 1996; Cooper and Hammer, 1996; Cooper, 1999; Araus et al., 2008; Sadras and Rebetzke, 2013). Below, we will use both terms to some extent interchangeably, but our intention is to refer to a state, so adaptedness would be the more correct term to use.

Adaptation and adaptedness usually pertain to yield or biomass. Traits different from yield itself are instrumental for realizing adaptation in yield, a very important one being phenology or earliness. The development of a genotype should match with the timing of resource availability and the absence or low incidence of stresses in its environment. The reaction norm for yield depends on the reaction norms for the yield components. The joint reaction norm of yield and yield components is a multivariate function of phenotypes that mutually affect each other and genetic and environmental inputs. For interesting elaborations of this multi-trait idea of reaction norms for plant breeding purposes, see Podlich and Cooper (1999), Messina et al. (2011), Chapman et al. (2012), Cooper et al. (2014a, 2014b), and Harrison et al. (2014). A good understanding of the processes leading to adaptedness and G × E requires observations on yield together with its main component traits as a function of (developmental) time. As multi-trait developmental data are still rare, G × E analysis methods for single traits observed at single time points dominate the literature.

Adaptability and Sensitivity

A reaction norm defines a genotype-specific function that translates environmental inputs into a phenotype. The G × E occurs when the reaction norms are not parallel, i.e., they intersect, diverge or converge (compare Fig. 1a and b with Fig. 1c–f). The occurrence of G × E forces phenotypic prediction models to become more elaborate and to contain genotype-specific parameters; intercepts, slopes and curvatures. These genotype-specific parameters are called sensitivity and adaptability parameters in the plant breeding literature and they facilitate the modeling of non-parallelism of reaction norms to account for G × E (Finlay and Wilkinson, 1963; Bänziger et al., 1997; Bradshaw, 2006; Sadras and Lawson, 2011; Slafer et al., 2014). Sensitivity applies to situations with single and well- identified explicit environmental gradients (drought stress index, temperature), adaptability to less concrete and non-explicit environmental gradients (environmental index based on average performance of all genotypes in a trial).

Stability and Risk

Observations on realized phenotypes in field trials will vary around the expected reaction norm for individual genotypes. This variation can be genotype-specific again and then it becomes another expression of G × E, which is captured by stability parameters (Eberhart and Russell, 1966; Wricke, 1966; Shukla, 1972; Fischer and Maurer, 1978; Lin et al., 1986; Lin and Binns, 1988; Piepho, 1998). Stability is called static (Lin et al., 1986), when the reaction norm model does not contain genotype-specific intercept or sensitivity terms and because of that does not correct for the general performance of the genotype or the production level of the environment. When stability is called dynamic, it is defined in terms of the variance of the residuals from more elaborate reaction norm models with genotype-specific parameters. In estimating procedures for dynamic stability variances, we need to account for the environmental factors shaping the reaction norms. For assessments of managed stress on a phenotyping platform, reaction norms are reduced to performance under stress versus control conditions and stability variances express variation within and between runs on the phenotyping platform. To increase the usefulness of stability measures, predictable or repeatable forms of G × E need to be distinguished from non-predictable and non-repeatable forms. For example, repeatable genotype × management interactions need to be distinguished from non-repeatable genotype × time and genotype × management × time interactions. For classical multi-environment trials across locations and years, a similar distinction between repeatable and non-repeatable types of G × E can be made (Chapman et al., 2000a, 2000b). Lin and Binns (1988) proposed to consider genotype × location interactions as predictable and to fit a reaction norm model to

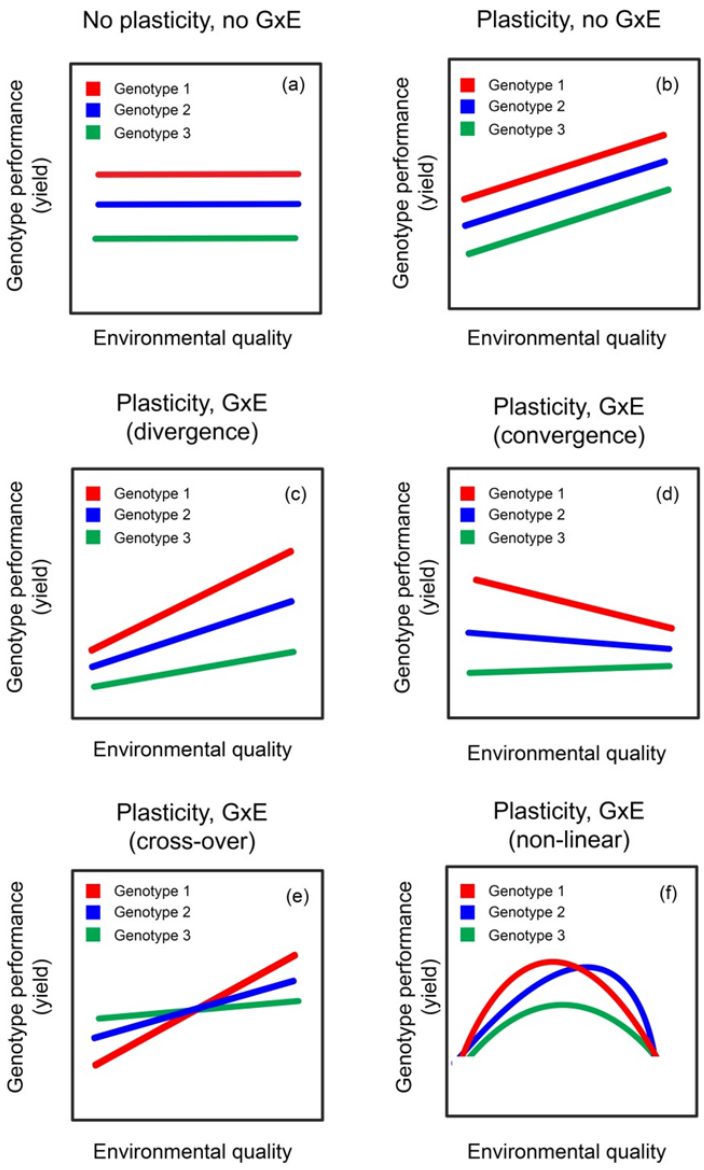

Fig. 1. Reaction norms for three genotypes that illustrate various forms of plasticity and Genotype × Environment interaction (G × E). No plasticity in (a) versus plasticity in (b) to (f), no G × E in (a) and (b) versus various forms of G × E in (c) till (f).

such interactions, whereas genotype × location × year interactions were considered to be unpredictable and genotype-specific stability variances were defined on the basis of the latter three-way G × E interactions.

Stability parameters are estimated mainly for yield and yield related traits, and to a lesser extent for quality traits. The concept of stability overlaps with the concepts of homeostasis, (no) plasticity, and resilience, i.e., reverting to equilibrium when perturbed (Lerner, 1954; Hanson, 1970; Govindaraju and Dancik, 1987; Debat and David, 2001; Sadras et al., 2009; Nicotra et al., 2010).

When genotypes differ in both reaction norm and stability, it may be worthwhile to combine these two concepts in a risk concept: the probability to exceed a

threshold yield level for part or the whole of the TPE (Eskridge and Mumm, 1992).

Environments

With respect to G × E, the major breeding questions for genotypes concern those related to adaptedness, adaptability and stability. For the environments, the major question is whether a set of trials, experiments, and conditions can be considered to form a sample from one particular target population of environments, or, alternatively, whether to attribute the set of trials, experiments and conditions to two or more target populations of environments. When trials (experiments) are grouped based on phenotypic or environmental information from the trials themselves, such groups are called mega-environments, or, sometimes, adaptation zones and ecological zones (Gauch and Zobel, 1988; Cornelius et al., 1992; Cooper et al., 1993; Gauch and Zobel, 1997; Atlin et al., 2000, 2011; Yan et al., 2000; Löffler et al., 2005; Chenu et al., 2011, 2013; Zhang et al., 2015). Mega-environments consist of trials and conditions that elicit comparable phenotypic responses in certain groups of genotypes. The question to the identification of mega-environments is the environmental counterpart of the question to genotypic adaptation. Several statistical approaches to G × E problems, like bilinear models (Gauch and Zobel, 1997) and bi-clustering (Corsten and Denis, 1990) approaches, simultaneously identify groups of genotypes with similar adaptedness patterns and groups of environments with similar conditions.

The identification of mega-environments is relevant for collections of trials that are assumed to originate from two or more discrete TPEs. However, even when assuming that trials stem from different TPEs, it may still be the case that genetic correlation exists between the genotypic performances in these trials. In the latter case, it seems better to speak of a subdivided TPE than of different TPEs. When the subdivision of the TPE is geographical, we often speak of regions. When trials are strongly heterogeneous, the central breeding question is whether to focus on specific adaptation or wide adaptation (Cooper and Woodruff, 1993; Braun et al., 1996; Ceccarelli et al., 1998; Atlin et al., 2000; Yan et al., 2000; Trethowan et al., 2001; Piepho and Möhring, 2005; Navabi et al., 2006; Chenu et al., 2011; Windhausen et al., 2012). If the decision is to go for wide adaptation, the most frequent question is how to efficiently combine the information from different regions of the TPE such as to maximize genetic gain across the undivided TPE (Atlin et al., 2000; Piepho and Möhring, 2005; Windhausen et al., 2012).

Another common question on the environmental side of G × E problems is that on the identification of trial locations that best represent the TPE (Heslot et al., 2013; Cooper et al., 2014a, 2014b). Representative locations should have a high correlation with the expectation across the TPE. Of course, this expectation is often unknown, because we are uncertain about which trials to assign to the TPE and which not. Another complicating factor is that genotypes may be strongly unbalanced across trials. When there is certainty about whether trials are from the TPE, a possible estimator for the expectation across the TPE is the mean of a set of standard or probe genotypes across trials (Mathews et al., 2011). Trials that are chosen to represent the TPE should also have good discriminative properties (high heritability).

Statistical Aspects of a G × E Course for Plant Breeders

Linear, Bilinear, and Mixed Models

In our G × E course of a few days to a week, the emphasis should be on data analysis, which means model formulation and model building, interpretation of statistical parameters in breeding terms and prediction of phenotypic responses to answer breeding questions, all as much as possible on real data. The learning objectives involve the expression of breeding questions in terms of suitable statistical models and parameters, the successful fitting of such models to breeding data, the interpretation of the results of the statistical analysis in breeding terms, and the reporting of the main findings in clear non-statistical language with insightful tables and figures. Statistical and theoretical issues related to parametrization, and estimation and testing procedures are of secondary importance. We believe that the main breeding questions above (adaptation, adaptability, stability, identification of TPE, choice of selection environments) can effectively be addressed by a combination of linear, bilinear, and mixed models. Essentially, the G × E course discusses models that contain linear and bilinear terms that can be fixed or random: linear-bilinear mixed models. The estimation procedures are typically least squares and REstricted Maximum Likelihood (REML) in software packages like SAS (SAS Institute, 2015), Genstat (VSN International, 2015a), Breeding View (Integrated Breeding Platform, 2015), ASReml (VSN International, 2015b), and ASReml for R (VSN International, 2015c). Testing of fixed effects takes place by Wald tests or F-tests, while variance components and correlations are tested by likelihood ratio tests (Smith et al., 2005; van Eeuwijk et al., 2010; Gumedze and Dunne, 2011). For us, testing is predominantly an activity within a model building strategy to identify a prediction model.

Linear and linear-bilinear models with only fixed terms are of limited use in the analysis of G × E data. Mixed models have facilities for modeling heterogeneity of genetic variances and correlations between environments as well as for modeling design features and spatial trends in individual trials. Furthermore, prediction of phenotypic traits across environments and estimation of

quantitative genetic parameters as genetic variance and correlation, heritability, and responses to direct, indirect and index selection are natural within the context of mixed models, whereas they become contrived in the context of models with only fixed terms.

Nevertheless, for didactical purposes, it is appropriate to start a course on G × E with fixed analysis of variance (ANOVA) models. Inferential procedures by standard F- and t-tests may be more familiar than Wald tests and likelihood ratio tests. ANOVA can also be used to repeat the basics of least squares estimation and testing theory. The ANOVA framework is convenient to introduce model building and different forms of interaction. Later, generalizations to bilinear models and mixed models are easily made.

One-and Two-Stage G × E Analysis with ANOVA

A good starting point for a G × E course is an ANOVA model for a response at plot level in a multi-environment trial (MET) with randomized complete blocks. The model has the form: $y_{ijk} = \mu + g_i + e_j + B_{k(j)} + ge_{ij} + \varepsilon_{ijk}$ with y_{ijk} the phenotypic response in block k for genotype i in environment (trial) j, g_i the fixed genotypic main effect, e_j the fixed environmental main effect, $B_{k(j)}$ the fixed effect of block k in environment j, ge_{ij} the fixed G × E term, and ε_{ijk} a random normally distributed error. This model is fitted by least squares. The students will be using specialized software to fit statistical models, so they do not need to be bothered with details of the fitting process. Still, they need to be able to choose least squares for fitting fixed regression and ANOVA models and REML for mixed models and they should be familiar with the expressions for estimators of statistics that are used to answer breeding questions. Familiarity then means that students recognize how the data are used to develop the estimator and how the estimator is related to a statistical model for the data.

In the fixed ANOVA above, the genotypic main effect and the G × E term will be tested against the residual mean square, with expectation $var(\varepsilon_{ijk}) = \sigma^2_\varepsilon$. The standard errors for contrasts on the genotypic main effects and the G × E effect are functions of the residual mean square. It is a good exercise for students to formulate and fit the fixed ANOVA two-way model with randomized complete blocks for a concrete MET dataset and perform a statistical analysis, including diagnostic checks and reporting of results. The analysis can initially concentrate on the test for G × E and ways of quantifying the amount of G × E in comparison to the genotypic main effect, for example, by comparing the magnitude of the sums of squares for those terms. The use of contrasts to investigate the structure of G × E can wait until models for G × E have been discussed (see below). We recommend following the protocol described by Welham et al. (2014) for the preparation of data, ANOVA analysis, identification of a predictive model, interpretation of results, and reporting.

For an analysis of MET data in a two-stage approach, in the first stage perform randomized complete block analyses per trial. After checking diagnostics for normality, homogeneity of variance and outliers, vectors of genotypic means are formed together with vectors of weights for subsequent G × E analysis. These weights are functions of the standard errors of means (Möhring and Piepho, 2009; Welham et al., 2010; Piepho et al., 2012). For the students, the two-stage approach offers an opportunity to rehearse the principles of ANOVA and mixed models on single trials. In the second stage, they fit a two-way ANOVA model without replication to the table of G × E means: $y_{ij} = \mu + g_i + e_j + ge_{ij} + \varepsilon_{ij}$. Estimation will follow least squares principles again, as in all fixed ANOVA models. The error term, ε_{ij}, has a normal distribution, $\varepsilon_{ij} \sim N(0,\sigma^2_\varepsilon)$, and is confounded with the two-way interaction. An independent error term can be obtained from the errors of the single trial analyses and will allow testing of the G × E term.

In an introductory G × E course, simple examples of one-stage analyses of G × E data can be presented. However, it is convenient to dedicate most of the time to two-stage modeling. The reason is that first, the statistical differences between one-stage and two-stage analyses are small in most cases and certainly when appropriate weighting schemes are used (Welham et al., 2010; Piepho et al., 2012), and second, the logistics of the two-stage approach are more intuitive. In the two-stage approach, the students first concentrate on the phenotypic analysis of single trials, taking into account design and spatial trends. Subsequently, the students can focus exclusively on the modeling of G × E. The simultaneous handling of large numbers of trials at plot level can be confusing, both at the statistical input and output level. Fewer things can go wrong in a two-stage approach, as compared with a one-stage approach.

Mixed Models

The fixed ANOVA model for the individual trials with randomized complete blocks in a MET can be turned into a mixed model by treating blocks as random in the one-stage model. For this kind of design, this will hardly change interpretations on the comparisons of genotypic means within individual trials, but students will have to realize that the mixed model imposes additional assumptions on the parameters in the model that will require checking. In this case, the extra assumption is that the block effects come from a normal distribution with a particular variance. Because block effects contribute to the variance of observations, the standard error for a genotypic mean will be larger in the model with random blocks. In contrast, for the standard error of a genotypic difference the block effects cancel out, so that this standard error is the same in the fixed and mixed model. Of course, for incomplete block designs, we expect the standard error of a difference to be smaller for random

blocks than fixed blocks. These points can be brought to the attention of the students by asking them to perform analyses on the individual trials of a MET with fixed and random blocks and to compare the results. Another point that deserves attention is the inspection of ANOVA tables of degrees of freedom, sums of squares, mean squares and F-tests, versus the inspection of tables of Wald tests for fixed effects and likelihood ratio tests for variance components. As a learning objective, students should be able to interpret these tables and use them to build prediction models for G × E data. We return to model building in a later section.

A more drastic change occurs when the genotypes are taken random, in both the one-stage and the two-stage model. All basic quantitative genetic concepts can be introduced and discussed: Best Linear Unbiased Predictor (BLUP), genetic variance, G × E variance, genetic correlation, heritability, response to selection, direct response to selection, correlated response to selection. We remark that we will use the word 'genetic' in a rather loose sense, mostly for indicating variance and correlation across genotypes, but also sometimes in a more restricted sense for exclusively additive genetic variance or correlation, as in mixed models with structured variance-covariance matrices.

The one-stage model is easier to work with for the presentation of estimators for all of these parameters. The students should be made familiar with expressions for quantitative genetic parameters. They should learn what the influence is of changing the number of environments and replicates on the heritability. A further objective is to learn how to evaluate the efficiency of selection in one particular environment (e.g., stress) to that in another (e.g., non-stress). For the two-stage modeling, the mixed model for a MET can be defined as $\gamma_{ij} = \mu_j + G_{ij} + \varepsilon_{ij}$, with μ_j the fixed intercept for environment j, and G_{ij} the random environment-specific genetic effect for genotype i in environment j. The matrix of random effects G_{ij} will have a multivariate normal matrix distribution with mean zero and variance-covariance matrix (VCOV) $\boldsymbol{\Sigma}$: $\{G_{ij}\} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$ (Smith et al., 2005). The VCOV, $\boldsymbol{\Sigma}$, is factorized in a VCOV for the genotypes that defines the correlations between the genotypes following from kinship and pedigree ($\boldsymbol{\Sigma}^G$) and another VCOV for the environments that expresses heterogeneity of genetic variance and correlations across environments ($\boldsymbol{\Sigma}^E$) as $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^G \otimes \boldsymbol{\Sigma}^E$, with $\otimes$ a Kronecker product. A major learning objective for a modern course on G × E is the formulation of VCOV structures for the genotypes, $\boldsymbol{\Sigma}^G$, and for the environments, $\boldsymbol{\Sigma}^E$. For the genotypes, common choices are $\boldsymbol{\Sigma}^G = \boldsymbol{I}$, with $\boldsymbol{I}$ an identity matrix for segregating biparental offspring populations, and $\boldsymbol{\Sigma}^E = \boldsymbol{A}$, with $\boldsymbol{A}$ the matrix of additive genetic relationships for association panels. For biparental offspring populations, compound symmetry, with a common covariance between pairs of genotypes, is another appropriate formulation. For association panels and for genomic prediction models, $\boldsymbol{\Sigma}^G$ can be based on pedigree information, marker information, or a combination of these two sources of information (Burgueño et al., 2012; Crossa et al., 2010a). The specification of $\boldsymbol{\Sigma}^G$ should be based on investigation of the genetic relationships, whereas the specification of $\boldsymbol{\Sigma}^E$ should be based on inspection of patterns in genotypic variances and correlations across environments. We discuss below a number of tools for determining such patterns in the context of the use of bilinar models. For $\boldsymbol{\Sigma}^E$ a number of well-known formulations exist. The most simple model has a common genotypic variance $\{\boldsymbol{\Sigma}^E\}_{jj} = \sigma^2_G$ and a common covariance $\{\boldsymbol{\Sigma}^E\}_{jj'} = \sigma_{GG'}$ across all environments, where $\{\boldsymbol{\Theta}\}_{kk'}$ denotes the entry of symmetric matrix $\boldsymbol{\Theta}$ for the kth row and k'th column. The heterogeneous compound symmetry model has environment-specific genotypic variances, $\{\boldsymbol{\Sigma}^E\}_{jj} = \sigma^2_j$ and a common covariance, while the uniform correlation model has also environment-specific variances, but a common correlation between environments, $\{\boldsymbol{\Sigma}^E\}_{jj'} / \sqrt{\{\boldsymbol{\Sigma}^E\}_{jj}\{\boldsymbol{\Sigma}^E\}_{j'j'}} = \rho$. The unstructured model has unique genotypic variances and covariances, $\{\boldsymbol{\Sigma}^E\}_{jj} = \sigma^2_j$, and $\{\boldsymbol{\Sigma}^E\}_{jj'} = \sigma_{jj'}$. As the latter model requires many parameters for estimation, factor analytic models are used as parsimonious approximations to unstructured models. Covariances are written as products of environmental scores, $\{\boldsymbol{\Sigma}^E\}_{jj'} = \lambda_j \lambda_{j'}$, while variances are composed of sums of squares of environmental scores and environment-specific variance terms,$\{\boldsymbol{\Sigma}^E\}_{jj} = \lambda^2_j + \delta^2_j$. More than one product term can be used for the covariances and variances, for example, $\{\boldsymbol{\Sigma}^E\}_{jj'} = \lambda_{1j}\lambda_{1j'} + \lambda_{2j}\lambda_{2j'}$ and $\{\boldsymbol{\Sigma}^E\}_{jj} = \lambda^2_{1j} + \lambda^2_{2j} + \delta^2_j$. In principle, the residual term, ε_{ij}, is confounded with G_{ij}, but the genetic and residual random terms can be separated by imposing a structure on $\boldsymbol{\Sigma}$, like $\boldsymbol{\Sigma}^G = \boldsymbol{A}$ or by using a factor analytic structure on $\boldsymbol{\Sigma}^E$ and interpreting the terms δ^2_j as non-genetic residuals. As an alternative to imposing structure on $\boldsymbol{\Sigma}$, σ^2_ε can be obtained from single trial analyses. For a single stage MET analysis, the problem of having to separate the genetic variance from the residual variance will not occur. The students should learn how to identify a suitable model for the genetic VCOV, $\boldsymbol{\Sigma}$, preferentially by log-likelihood ratio tests (Gumedze and Dunne, 2011). If these tests cannot be applied, because the VCOV models are not nested, then information criteria like AIC or BIC may be used (Verbeke and Molenberghs, 2009; Müller et al., 2013).

The students need to interpret the structure of $\boldsymbol{\Sigma}^E$ for conclusions about the heterogeneity of the sample of environments included in the MET. This is an excellent moment to discuss the concept of TPE and investigate the question of whether there are indications that the MET contains trials from more than one TPE, or, whether different mega-environments can be distinguished. To identify different mega-environments in the mixed model context, various types of cluster analysis can be applied to the estimate for $\boldsymbol{\Sigma}^E$, as explained in Cullis et al. (2010). Mixed model theory as developed by Piepho and Möhring (2005)

will help to establish the gain of subdividing the TPE in different groups or regions. We recommend the analysis of not just yield or biomass, but, if possible, also traits related to phenology. As yield is the primary trait of interest, mega-environments should foremost address yield. Nevertheless, inspecting VCOV models for yield and yield components will contribute to the physiological interpretation of solutions to mega-environment identification problems.

Closely related to the delineation of mega-environments, investigating the structure of $\mathbf{\Sigma}^E$ also helps in answering questions about the presence of crossover interactions (Crossa et al., 2004; Yang, 2007; Burgueño et al., 2008). Environments or trials with high genetic correlations will have few crossover interactions.

From the identified model for the MET data, predictions for genotypic performance can be made in each of the environments. Subsequently, adaptation of genotypes can be studied in the very limited sense of which genotypes do best in individual environments. Environments in which the same genotype or subset of genotypes do best may be members of a common mega environment. Questions on adaptability and stability cannot be answered straightforwardly from models without implicit (bilinear) or explicit (factorial regression) genotypic and environmental covariables.

Interaction in Linear Models

The two-way ANOVA can serve to introduce the *concept of interaction in linear models*. A plot of estimates of G × E means, $\hat{\mu}_{ij}$ versus environmental main effects, $\hat{e}_j$, or, equivalently, environmental means, $\hat{\mu}_j$, can help in diagnosing non-parallelism of genotypic responses across environments and thereby visualize the presence of G × E. Actually, G × E can equally be inspected via a plot of the genotype × environment means, $\hat{\mu}_{ij}$, versus the genotypic means, $\hat{\mu}_i$, although it is more natural to plot G × E means versus environmental means, as this is the basis for the Finlay Wilkinson regression (Yates and Cochran, 1938; Finlay and Wilkinson, 1963). Plots of G × E means versus an environmental characterization as the environmental mean can be used to come back to the concept of the reaction norm within the context of ANOVA. When reaction norms are non-constant, genotypes show plasticity, see Fig. 1a versus Fig. 1b–f (Allard and Bradshaw, 1964; DeWitt and Scheiner, 2004; Sadras and Lawson, 2011; Slafer et al., 2014). To elaborate this point statistically, plasticity points to the existence of an environmental main effect. When reaction norms show plasticity and are non-parallel, we have G × E, see Fig. 1c–f. When reaction norms are non-parallel, but do not intersect, we speak of *quantitative interaction* or *non-crossover interaction*, see Fig. 1c and d. When the reaction norms do intersect, we speak of *qualitative interaction* or *crossover interaction*, see Fig. 1e and f. This latter type of G × E has more severe consequences for breeders, as it will change the rank order of genotypes as a function of the environmental conditions.

A formal test for G × E interaction in a two-way ANOVA on G × E means requires an estimate for the error, σ^2_ε. Such an error estimate can come from the analysis of single trials. Independent of any formal testing, the ANOVA partitioning of the phenotypic variation into genotypic and environmental main effect and G × E gives a rough indication of the importance of G × E, judged by sum of squares, mean squares or F-values. We prefer the use of variance components in random and mixed models over sums of squares partitioning in fixed ANOVAs to quantify the contributions of genotype, environment and G × E to the phenotypic variation, even more so when genotypes and environments contain crossed and/or nested factorial structure. The variance components can be expressed on the scale of coefficients of variation to facilitate their interpretation (Gelman, 2005).

Estimation and inference for two-way ANOVA is straightforward for complete G × E tables. For ANOVA interaction parameters, under sum to zero constraints the classical estimator for G × E interaction residuals is $\widehat{ge}_{ij} = \overline{y}_{ij} - \overline{y}_i - \overline{y}_j + \overline{y}$, where the bar indicates averaging over subscripts that are omitted. Such simple looking estimators do not exist for non-orthogonal G × E data. A number of adaptation and stability parameters have been proposed that are functions of these interaction residuals (Wricke, 1962). However, the biological reality and usefulness of such estimators depends on the extent to which the sum to zero identification constraints can be biologically justified (Denis, 1991).

Crossover Interactions

In a course on G × E, attention should be given to methods for detecting crossover interactions. The importance of crossovers depends on the magnitude of these interactions and the genotypes and environments involved. The identification of crossovers is a first step in the exploration of the genetic and physiological factors underlying genotypic differences that are conditioned by the environment. Crossover tests make sense for G × E tables with limited numbers of cultivars of contrasting adaptation (tolerant or resistant versus susceptible or sensitive) that are tested under contrasting environmental conditions (e.g., stress versus non-stress). However, for these situations more powerful a priori crossover contrasts can be defined and tested by standard *t* test procedures. To emphasize the continuity between ANOVA and mixed models, tests for a priori crossover contrasts should also be demonstrated for mixed models. Students will need to learn to first identify the right fixed or mixed linear model and then answer specific breeding questions by imposing the appropriate contrasts on the levels of genotypic and environmental factors. When a suitable mixed model is

identified and specified, mixed model software programs like ASReml, Genstat, and SAS offer the possibility to define and test contrasts that then will automatically have the correct standard errors.

The literature on G × E shows a certain obsession with the identification of crossover interactions. The two-way ANOVA model provides a plain setting to address the issues of the identification of crossover interactions and the assessment of their importance and interpretation. A simple post hoc procedure for testing individual crossover interactions first identifies pairs of genotypes, i and i', and pairs of environments, j and j', for which either the condition $(\gamma_{ij} - \gamma_{i'j}) > 0$ and $(\gamma_{ij'} - \gamma_{i'j'}) < 0$ is fulfilled, or the condition $(\gamma_{ij} - \gamma_{i'j}) < 0$ and $(\gamma_{ij'} - \gamma_{i'j'}) > 0$. Subsequently, the interaction sum of squares for the contrast $\gamma_{ij} - \gamma_{i'j} - \gamma_{ij'} + \gamma_{i'j'}$ is calculated with a conservative correction for multiple a posteriori testing according to Scheffé's test for simultaneous inference (Kuehl, 2000). More powerful tests for testing crossovers are the Azzalini–Cox test (Azzalini and Cox, 1984) and the Gail–Simon test (Gail and Simon, 1985). See Baker (1988) for their application to plant breeding examples.

Model Building

Two-way G × E models are generalized to multi-way models when either or both of the genotypic and environmental factor contains itself factorial structure. For example, the genotypic factor may be the factorial product of an earliness classification and a stress tolerance classification. The environmental factor can be a product of location and year, or of management and year. In addition, multi-way factorial structure on the environments is easily imaginable when multiple management factors (drought/irrigated × high/low nitrogen) are crossed with years and/or locations. For analyzing multi-way G × E tables in ANOVA (all factors fixed) a recommended strategy first tests the highest order interaction (Welham et al., 2014, section 8.3.1). When this interaction is significant, the model for prediction will include all model terms up to the highest interaction. We then need to inspect the corresponding multi-way table of means and calculate contrasts on that table for interpreting the interactions. When the highest order interaction is not significant, the corresponding interaction term will not be included in the prediction model and we continue testing the next highest order interactions (Welham et al., 2014, section 11.2.3).

For mixed models, the strategy described by Welham et al. (2014) carries over to the table of fixed terms given that the random structure is kept constant, i.e., the random model is the same for all fixed terms. Testing is done by Wald tests or their F-test equivalents. For mixed models, the situation is complicated by the fact that for individual terms a decision needs to be motivated whether they are fixed or random. This is a subject for which the literature is not very clear, and students (and teachers) get easily confused. Choices for whether genotypic, environmental and G × E terms should be fixed or random depend on the question that requires an answer, but also on the assumptions that one is willing to make. Assuming that a term is random imposes additional constraints on the VCOV for the observations and requires verification of the distribution of the corresponding effects. A number of characteristic situations can be presented with the question to formulate reasons for choosing individual model terms as fixed or random and discuss the consequences of choosing particular terms as fixed or random. For METs with segregating populations that will be used for QTL mapping or association mapping, we prefer to take the main effects for genotypes and G × E terms random. If the genotypes consist of a limited set of genotypes late in the breeding program or of a set of candidate cultivars in an official variety testing cycle, then the genotypic main effects best are chosen fixed. For the environments, managed stress treatments can be treated as fixed when they refer to (hopefully) repeatable environmental conditions. Similarly, repeatable G × E for a small number of genotypes under various levels of a well-defined managed stress factor will be fixed. For locations and years, we prefer to take the main effects, which are a kind of intercept terms, fixed, where the year main effects may be taken random when it concerns many years. For the G × E interactions, the genotype × location interaction is fixed for repeatable locations and selected genotypes (Annicchiarico, 1997; Annicchiarico et al., 2005). The G × E interactions with years are all random. Testing of fixed genotype × management and genotype × location interactions will happen against a background of random genotype × management × year and genotype × location × year interactions. Therefore, tests for these repeatable G × E interactions in mixed models will differ from similar tests in an ANOVA as will standard errors of differences for contrasts. Students will need a series of practical examples and exercises to learn how to define appropriate mixed models and arrive at sound breeding conclusions.

Partitioning G × E in Pattern and Noise by Grouping Genotypes and Environments

Imposing a categorization or grouping on genotypes and environments allows a partitioning of the initial G × E into an interaction between genotype groups and environment groups and a G × E residual, a deviation from the grouping term: $\gamma_{i(k)j(l)} = \mu + g_i + e_j + rc_{kl} + \delta_{i(k)j(l)} + \varepsilon_{i(k)j(l)}$ with rc_{kl} the interaction between the genotype group k and the environment group l (rows and columns, respectively of the two-way G × E table). The term $\delta_{i(k)j(l)}$ is a deviation from the row by column group interaction model that can be considered to represent a random normal variable with zero mean and proper variance. The random term $\varepsilon_{i(k)j(l)}$

is confounded with $\delta_{i(k)j(l)}$ in a two-stage analysis of G × E means, but an estimate for its variance can be obtained from intra-trial analyses. The grouping of genotypes and environments is tested against the mean square for the deviations of the interaction model by an F-test (ANOVA) or Wald test (mixed model) (van Eeuwijk et al., 1996; Kuehl, 2000; Welham et al., 2014). The mean square for the deviations is tested against the estimate for the error, σ^2_ε. The choice of testing the groupings against the deviations from the groupings, changes an ANOVA model implicitly in a mixed model. Groupings of the levels of environmental factors can also be included in more complex mixed models with multiple environmental factors, in which some factors and G × E interactions are fixed and others random. For example, in a MET with fixed genotypes and multiple trials across fixed locations and random years, we may want to cluster the locations for the fixed genotype × location interaction, but not the years in the random genotype × year interaction. As a learning objective, students should identify and specify relevant groupings of genotypes and environments and test these groupings in both ANOVAs and mixed models. Correct specification in ANOVA and mixed model software will lead to the desired F-tests and Wald tests, respectively.

The success of groupings of genotypes and environments depends on the amount of variation that is described by the interaction between the groupings, the significance of the test on the groupings and the test on the deviations from the groupings. Strong significance of the test for groupings and no significance for the deviations indicate that the pattern in the G × E is captured by the genotype and environment groupings, leaving noise for the deviations from the groupings. The part of the G × E that is covered by genotype and environment groupings is the part that qualifies for being repeatable. Predictions from models with genotype and environment groups are delivered at the level of the groups and not at the level of individual G × E combinations. Partitioning the G × E into a part described by groupings and another part by deviations from grouping reflects a very general principle to reduce the complexity of G × E interactions. Many types of genotypic and environmental groupings can be proposed. Students should become capable of defining and constructing promising groupings from physiological and breeding knowledge. Groupings obtained from exploratory analyses like bi-clustering (Corsten and Denis, 1990) and the application of additive main effects and multiplicative interaction models (Gauch, 1992, 2013) can be inserted in ANOVA models and mixed models to test their contribution to the G × E. Some correction for data snooping and multiple testing is necessary. Again, Scheffé's simultaneous test for a posteriori contrasts is a conservative possibility.

Partitioning G × E in Pattern and Noise by Genotypic and Environmental Covariables

Grouping of genotypes and environments can be interpreted as the introduction of qualitative covariables on the levels of genotypes and environment. Similarly, quantitative covariables on genotypes and environments can be included to investigate the nature of the G × E. We consider the skillful introduction of covariables for investigating G × E patterns to constitute a major learning objective in a course on G × E. Inclusion of quantitative covariables in models for G × E is equivalent to the definition of contrasts to study interactions in classical ANOVA and mixed models. Various statistical textbooks contain insightful chapters on the use of contrasts to study factorial interactions. We recommend Kuehl (2000) and Welham et al. (2014). In the plant breeding literature, models for G × E using covariables are called factorial regression models. Some papers giving theory and applications of factorial regression are Denis (1988), van Eeuwijk et al. (1996), Vargas et al. (1999), Voltas et al. (1999a), Malosetti et al. (2004), and van Eeuwijk et al. (2005). Some nice recent examples of factorial regression can be found in Crossa et al. (2015) and Vargas et al. (2015).

Factorial Regression: A Single Environmental Covariable

For a factorial regression model that contains a single environmental covariable, to describe G × E, the fixed G × E term in ANOVA and mixed models, ge_{ij}, is partitioned into a regression part, $\beta_i z_j$, with a genotype-specific slope or sensitivity, β_i, to the environmental covariable, z_j, and a residual or deviation from the regression model, δ_{ij}: $ge_{ij} = \beta_i z_j + \delta_{ij}$. Thus, the double indexed G × E term ge_{ij}, which does not permit predictions to new environments, is replaced by a separable formulation, $\beta_i z_j$, with single indexed values for the genotypic parameter and the environmental covariable, that does offer the possibility for predictions to new environments for genotypes whose sensitivities have been estimated.

It should be emphasized that it is the heterogeneity of the slopes that is of importance for G × E. A sensitivity to an environmental covariable that is constant across genotypes partitions a fixed environmental main effect, $e_j = \beta z_j + \delta_j$, but not a fixed ge_{ij} term, so the environmental covariable would be responsible for plasticity without G × E. A formal test for whether the covariable z_j 'explains' a significant proportion of ge_{ij} can be constructed from testing the mean square for the heterogeneity of the genotypic sensivities over the mean square for the deviations from the factorial regression. In ANOVA, the degrees of freedom for the regressions are $I - 1$ for the heterogeneity of slopes and $(I - 1)(J - 2)$ for the deviations from the regression.

Factorial Regression with Multiple Environmental Covariables

Above, we introduced the concept of factorial regression for a single environmental covariable. However, typically more than one environmental covariable or characterization is required to arrive at an acceptable description of the pattern in the G × E. The factorial regression model can contain multiple environmental covariables that are elements of the environmental covariable set EC, $ge_{ij} = \Sigma_{k \in EC} \beta_{ki} z_{kj} + \delta_{ij}$, or polynomial forms of environmental covariables, $ge_{ij} = \beta_{1i} z_j + \beta_{2i} z_j^2 + \delta_{ij}$. Further elaborations of the factorial regression approach include standard nonlinear curves as the logistic, Richards and Gompertz function (Butler and Brain, 1993). For fixed G × E terms, the maximum number of environmental covariables that can be included is one fewer than the number of environments, or $J - 1$. When more than $J - 1$ environmental covariables are included, some form of penalization or dimension reduction is required, as in reduced rank regression (Denis, 1988; van Eeuwijk, 1992) and partial least squares (Aastveit and Martens, 1986; Vargas et al., 1999, 2015). Recently, interesting suggestions were made to include large numbers of environmental covariables in mixed model factorial regressions where these covariables define an environmental relationship matrix that imposes structure on random G_{ij} terms. As before, we assume that the matrix of random G_{ij} effects has a multivariate normal matrix distribution: $\{G_{ij}\} \sim MVN(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbf{\Sigma}^G \otimes \mathbf{\Sigma}^E$, but now environmental covariables define the environmental VCOV $\mathbf{\Sigma}^E$ (Jarquín et al., 2013; Bustos-Korts et al., 2016; Malosetti et al., 2016; Pérez-Rodríguez et al., 2015).

Selecting Environmental Covariables

Classical examples of environmental covariables are soil and meteorological variables. To choose environmental covariables in factorial regression models, standard variable subset selection (forward and stepwise regression) can be demonstrated to students. Nowadays, automatic environmental monitoring protocols produce measurements at short time intervals. It is not immediately obvious how to select from a multitude of short interval measurements the most relevant ones. Routine variable selection procedures will not work well with large numbers of variables to select from. Alternative statistical variable selection methods like penalized and sparse regression methods for high dimensional regression like the Lasso (Tibshirani, 1996; Meinshausen, 2007; Taylor et al., 2012) look interesting, but have not been extended yet to screen large sets of covariables for interaction terms. However, an attractive alternative to statistical selection methods seem integrations over time of multiple environmental variables by crop growth models (Chapman, 2008; Chenu et al., 2011, 2013; Heslot et al., 2013) to produce a limited set of environmental characterizations known to be biologically relevant. Students in a G × E course will benefit from an introduction to the concept of integration of environmental information over time with crop growth models. Instead of running a simple crop growth model to obtain environmental characterizations, students may also be provided with environmental characterizations from earlier crop growth model runs.

Factorial Regression with Genotypic Covariables

Factorial regression under inclusion of genotypic covariables is another useful approach to identify the patterns driving G × E and to search for separable models. The formulation of the model for multiple genotypic covariables, of the type x_{li}, that are all elements of the genotypic covariable set GC, in the case of fixed G × E is $ge_{ij} = \Sigma_{l \in GC} x_{li} \alpha_{li} + \delta_{ij}$, with the number of genotypic covariables limited to maximally $I - 1$. Again, more genotypic covariables can be included, but analogous to the situation for environmental covariables above some form of dimension reduction or penalization should be imposed. Genotypic covariables for yield can be measurements made under managed conditions, including phenotyping platforms, on traits like disease resistances and biotic stress tolerances. The parameters α_{li} then express the severity of the disease or abiotic stress in the experiment j. When the genotypic covariables are functions of marker genotypes, the factorial regression model above immediately becomes a model for multiple quantitative trait loci (QTLs) with the QTLs interacting with the environment (Malosetti et al., 2004, 2013; Boer et al., 2007; van Eeuwijk et al., 2010). For genomic prediction under G × E, the full set of markers can be used to define a genomic relationship matrix on the genotypes. This genomic relationship matrix can be combined with an environmental relationship matrix to produce genomic predictions for new genotypes in new environments (Bustos-Korts et al., 2016; Malosetti et al., 2016).

Linear-Bilinear Models: An Overview

Where factorial regression models contain multiplicative formulations for G × E that use explicit genotypic and environmental covariables, linear-bilinear models use implicit covariables. For our purposes, the essence of analyses by linear-bilinear models is that they generate ideas for groupings and covariables that can be further tested in ANOVA and mixed models. For recent papers on linear-bilinear models, see Gauch et al. (2008), Crossa et al. (2010b), Crossa, (2012) and Gauch (2013). Linear-bilinear models have their name from the fact that in addition to linear terms for genotype and/or environmental main effects they contain bilinear terms for G × E or for combinations of a main effect and G × E. Bilinear terms are separable products of parameters, like $r_i c_j$, with r_i for the sensitivity of genotype i and c_j for the characterization of environment j, that both

need to be estimated from the phenotypic data. The name comes from the fact that fixing the genotypic parameter r_i in $r_i c_j$ makes the model linear in c_j and vice versa. A well-known member of the linear-bilinear class of models is first, the Finlay Wilkinson model (Yates and Cochran, 1938; Finlay and Wilkinson, 1963). A second linear-bilinear model is the additive main effects and multiplicative interaction effects (AMMI) model (Gauch, 1988, 1992). Third, we mention the genotype main effects and genotype × environment interaction effects model (GGE model; Yan and Kang, 2002). This model is also known in the G × E literature as the site regression model (Crossa and Cornelius, 1997). Actually, the GGE model is equivalent to principal components analysis of the G × E two-way table of means and all theory on principal components carries over to GGE analysis (Gabriel, 1971; Jolliffe, 2013). More linear-bilinear models have been developed and are used for G × E analysis, but would fall outside a G × E course of 3 to 5 d, see for example Cornelius and Seyedsadr (1997) and Crossa (2012).

Finlay Wilkinson Model

For a table of two-way G × E means the fixed Finlay Wilkinson model is $y_{ij} = \mu_i + r_i c_j + \delta_{ij} + \varepsilon_{ij}$. In comparison with the two-way ANOVA model we rewrite the linear terms $e_j + ge_{ij}$ as $r_i c_j + \delta_{ij}$. In the Finlay Wilkinson model, a single environmental characterization is used, c_j, that either is equal to the average performance of all genotypes in an environment, as in the paper by Finlay and Wilkinson (1963), or is very close to it, as in Mandel (1969). For the latter case, a genotypic intercept term and the genotypic and environmental scores, r_i and c_j, respectively, are found from minimizing by least squares the expression: $\Sigma_{i,j}^{I,J}[y_{ij} - (\mu_i + r_i c_j)]^2$. Estimates for the scores are obtained from a singular value decomposition (SVD) of the matrix of G × E means corrected for the genotypic main effect: SVD($y_{ij} - \mu_i$) (Gabriel, 1978). Thus, genotypes are characterized by an intercept parameter, μ_i, for general performance and a slope or sensitivity or adaptability parameter, r_i, with unit mean. A large value for the intercept and a sensitivity value close to 1 together point to a widely adapted genotype. The recommended way for interpreting the results of a Finlay Wilkinson analysis is first to check whether a significant and relevant part of the G × E is explained by the heterogeneity of the slopes and whether the residuals from the model do not show shortcomings of the model. Next the predicted values can be calculated, $\hat{y}_{ij} = \hat{\mu}_i + \hat{r}_i \hat{c}_j$, and the fitted regression lines plotted to identify the superior genotypes for specific environments. Predictions for new environments are possible in so far new environments can be recognized as being similar to trials already included in the set of trials to build the model. The Finlay Wilkinson model is a candidate model for relatively simple environmental configurations in which the environments are homogeneous and differ in a single dimension. For example, the environments may represent optimal conditions except for the level of a limiting factor. When the environments have a more complicated factorial structure, Finlay Wilkinson terms can be embedded in mixed models. For example with a fixed location and a random year factor, a fixed bilinear Finlay Wilkinson term can be proposed that regresses genotype × location means on location means, to study adaptability, while the genotype × year and the genotype × location × year interactions are chosen random. The deviations from the Finlay Wilkinson regression will then be tested over the three-way genotype × location × year interaction. Another generalization of fixed Finlay Wilkinson models are mixed bilinear models in which random bilinear terms are included to describe G × E (Gogel et al., 1995; Nabugoomu et al., 1999): $y_{ij} = \mu_i + G_{ij} + \varepsilon_{ij}$ and $\{G_{ij}\} \sim MVN(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbf{\Sigma}^G \otimes \mathbf{\Sigma}^E$ and $\mathbf{\Sigma}^E$ having a factor analytic structure; the variance for environment j is $\lambda_j^2 + \delta_j^2$ and the covariance between environments j and j' is $\lambda_j \lambda_{j'}$ with λ_j a score for environment j that defines covariances with other environments and the basis for a shared part of the genetic variance, while δ_j^2 stands for an environment-specific genetic variance part.

The Finlay Wilkinson model is the most frequently used model for the analysis of G × E according to a citation analysis using the Web of Science (see later paragraph). For that reason, it deserves some time dedicated to it in any course on G × E. However, its applicability is limited. Within the class of fixed bilinear models more flexible models are available that allow more than just one bilinear term to be included for a description of G × E. For larger numbers of genotypes, mixed bilinear models seem a more viable modeling option than fixed bilinear models as the Finlay Wilkinson model.

The original Finlay Wilkinson model that aims at an analysis of adaptation and adaptability was extended by Eberhart and Russell (1966) with a stability analysis. To that end, they defined genotype-specific stability variances based on the deviations from the Finlay Wilkinson regression lines. With Gauch (2013), we share some doubts about the utility of stability analyses. First, when calculating stability statistics across a set of trials, it should be verified that these trials belong to a single TPE or mega environment. Second, the analysis of G × E should emphasize the identification of adequate models for the reaction norms and minimization of the deviations from the reaction norms, i.e., minimization of stability variances. When reaction norm models show a good fit to the data with clear genotypic differences for the reaction norm parameters, it is unlikely that simultaneously biologically relevant variation for stability variances will be present (see Kraakman et al., 2004). For historical and didactical reasons, in a G × E course, Finlay Wilkinson regression

and Eberhart Russell stability should be presented as early models that connected statistical parameters with breeding concepts such as adaptation, adaptability and stability. A learning objective for students is how to evaluate the merits of classical Finlay Wilkinson and Eberhart Russell approaches in comparison to recent more elaborate fixed and mixed bilinear models.

AMMI Models

Fixed bilinear models are and have been very popular for the analysis of G × E. With the Finlay Wilkinson model, the AMMI model, (Gauch, 1988; Gauch, 1992) deserves considerable attention in any G × E course. In the AMMI model, we write the G × E as $ge_{ij} = \Sigma_{a=1}^{A} r_{ai} c_{ai} + \delta_{ij}$, with r_{ai} denoting genotypic scores or sensitivities and c_{aj} denoting environmental scores or characterizations. The full AMMI model is $y_{ij} = \mu + g_i + e_j + \Sigma_{a=1}^{A} r_{ai} c_{ai} + \delta_{ij} + \varepsilon_{ij}$. For fixed two-way G × E tables, estimates for genotypic and environmental scores are obtained from an SVD of the ANOVA interaction residuals, ge_{ij}. The SVD of the interaction is equivalent to finding the environmental characterizations that best discriminate between genotypes following a least squares criterion. In an introductory course on G × E it is enough to mention that the multiplicative scores are obtained by minimizing $\Sigma_{i=1,j=1}^{I,J} (ge_{ij} - \Sigma_{a=1}^{A} r_{ai} c_{aj})^2$. The number of terms, A, for 'adequate' description of the G × E can be assessed in a number of ways (Cornelius, 1993; Bro et al., 2008; Josse and Husson, 2012), but for an introductory course, the F-test approximation by Gollob (1968) will do. This approach allocates $(I + J) - (1 + 2a)$ degrees of freedom to the sums of squares explained by a-th term to convert it into a mean square. Analogous to earlier partitions of the G × E, the mean square corresponding to the AMMI model can be tested against deviations from that model, while the deviations are tested against an independent estimate for the error from within trial analyses.

After assessing the dimension of the AMMI model, i.e., establishing the number of multiplicative terms to retain, A, adaptation can be investigated by plotting and comparing the environment-centered predictions for individual genotypes, $\hat{\gamma}_{ij} = \hat{\mu} + \hat{g}_i + \Sigma_{a=1}^{A} \hat{r}_{ai} \hat{c}_{aj}$. The predicted reaction norms may be used to find out which environments are similar or part of a mega-environment by comparing the best genotypes (Gauch et al., 2008). The genotypes can further be characterized by the genotypic main effect, g_i, a measure of wide adaptation, and the genotypic sensitivities, r_{1i}, r_{2i}, …, r_{Ai}. These genotypic sensitivities can be used to identify groups of genotypes with similar G × E.

The sum of squares of the genotypic scores, $S_i = \Sigma_{a=1}^{A} r_{ai}^2$, is an approximation to the sum of squares for interaction in a fixed two-way ANOVA, which is equivalent to Wricke's stability statistic (Wricke, 1962): $W_i = \Sigma_{i=i,j=1}^{I,J} ge_{ij}^2$. The AMMI genotypic scores need to be scaled appropriately for this approximation to work (Gauch et al., 2008). Still, although stability parameters can be defined in fixed ANOVA, we prefer to estimate stability parameters as genotype-specific variances in mixed models.

In the AMMI model, environments have as characteristics the general quality, e_j, and the specific qualities c_{1j}, c_{2j}, …, c_{Aj}. Comparable to the genotypes, environments can be grouped on the basis of their environmental scores. To test the contributions of these groupings, contrasts in ANOVA or mixed models can be defined. Another possibility is to apply tests for multiplicative interactions as described by Milliken and Johnson, (1989).

AMMI approaches are remarkably popular. Good predictive properties have been attributed to AMMI models (Gauch, 2006; Gauch et al., 2008). Surely, AMMI predictions will be useful for many breeding purposes, but if bilinear models for G × E can be embedded in mixed models, we would create predictions from the latter. Another asset of AMMI models, and bilinear models in general, is their possibilities to display G × E patterns graphically in biplots (Gauch and Zobel, 1997; Yan et al., 2000; Yan and Rajcan, 2002). For $A = 2$, an AMMI biplot contains genotypes with coordinates (r_{1i}, r_{2i}) and environments with (c_{1j}, c_{2j}). AMMI biplots are highly useful tools to explore G × E patterns. Outlying genotypes and environments can readily be detected, as can groupings of genotypes and groupings of environments. Even adaptation can be investigated in AMMI biplots (Voltas et al., 1999b). A G × E course will need to reserve time for students to learn the interpretation rules for various types of biplots, but AMMI models should be used primarily for exploration of G × E patterns, with formal testing of G × E structure in mixed models. Recent developments in Bayesian bilinear models (Crossa et al., 2011; Perez-Elizalde et al., 2012; Josse et al., 2014) alleviate the inferential restrictions on bilinear models, but Bayesian bilinear models would fall outside the scope of an introductory G × E course.

GGE Models

In the GGE model, the ANOVA terms $g_i + ge_{ij}$ are written as $\Sigma_{p=1}^{P} r_{pi} c_{pj} + \delta_{ij}$, with P bilinear terms, leading to the model $y_{ij} = \mu_j + \Sigma_{p=1}^{P} r_{pi} c_{pj} + \delta_{ij} + \varepsilon_{ij}$ (Yan and Kang, 2002). With Finlay Wilkinson models and AMMI models, GGE models enjoy a large popularity in the applied literature on G × E. A first reason is that the GGE model produces biplots that cover both the genotypic main effect and the G × E, while AMMI biplots focus on G × E solely. As a second reason, we mention that a particular feature of GGE biplots appears to allow the identification of mega-environments (Yan et al., 2000). The GGE biplot is mostly shown in two dimensions, as more dimensions are difficult to work with. The quality of the two dimensional GGE biplot for inference on genotypic adaptation (predicted values of GGE

model, equivalent to lengths of projections of genotypic vectors on environmental vectors), genotypic sensitivities (genotypic scores), genetic variances (squared lengths of environmental vectors), and genetic correlations (angles between environmental vectors) depends on the amount of variation that is represented by the first two dimensions of the GGE model. As the first axis of GGE models tends to mimic the genotypic main effect, which is not present in an additive form in the GGE model, a two dimensional GGE biplot will cover about as much G × E variation as an AMMI model with one bilinear term. Therefore, in two dimensional GGE biplots, less G × E pattern will be shown than in two dimensional AMMI biplots. In a GGE biplot, an average environment axis can be constructed as the average of the vector representations of the environments. Genotypic projections on this average environment axis approximate the genotypic main effect. The distance between a genotype representation in the GGE biplot (r_{1i}, r_{2i}), and its projection on the average environment axis is assumed to give an estimate for stability (Yan and Kang, 2002). This type of stability in the GGE biplot will be close to the stability based on the first genotypic score in an AMMI biplot. Various discussions have been published about the comparison of AMMI and GGE model analyses (Gauch et al., 2008; Yang et al., 2009; Gauch, 2013), without the authors getting to an agreement. We feel that both AMMI and GGE are very useful techniques to explore G × E interactions. There is little reason to prefer one to the other. GGE biplots present a view on the totality of genotype related variation, $g_i + ge_{ij}$, whereas AMMI biplots show more detail for the G × E part of the phenotypic variation, ge_{ij}.

GGE or principal component biplots are excellent ways to explore genetic variances and correlations. The biplot with optimal scaling for the environments can be used alongside with scatterplots matrices for the phenotypic responses across environments to develop ideas about the variance-covariance structure of the data. Hypotheses with respect to patterns in the genetic variances and correlations can be tested in mixed models with likelihood ratio tests.

Delineation of mega-environments based on winning genotypes in individual environments (Yan et al., 2000) in GGE models seems risky (Yang et al., 2009). Suggestions for defining mega-environments, or subdivisions of the TPE, can be obtained in many ways, but we would like to insist on testing the efficiency of such subdivisions in a suitable mixed model, following protocols as described by Piepho and Möhring (2005).

Crop Growth Models, Multi-Trait Reaction Norms and Networks

The phenotypes that are analyzed to answer questions with respect to G × E are mostly complex traits, as yield itself. From a statistical perspective, complex traits are traits for which multiple QTLs can be identified with additive, dominance and epistatic effects that can interact with the environment. For complex traits, in a G × E course students learn how to fit suitable models with genotypic and environmental covariables to describe G × E. In our statistical approaches to G × E little attention is given to the fact that phenotypes are products of genetic, physiological and environmental interactions over time. A good way to introduce the developmental aspects of phenotypes is via crop growth models. In this section, we present a statistical-physiological framework for better understanding G × E. Let us label the phenotype we want to predict as the target trait, or focus trait, y_{ij}^{f}. The target trait is now a response trait in a crop growth model, with as inputs a vector of genotype dependent component traits, $\mathbf{y}_i^c$, and a vector of environmental variables, $\mathbf{z}_j$ (Yin et al., 2000a; Chapman et al., 2002; Snape et al., 2007; Chenu et al., 2009; Malosetti et al., 2016; Technow et al., 2015; van Eeuwijk, 2015). Component traits are related to resource capture (e.g., leaf area index, root architecture), conversion efficiency (e.g., light use efficiency, water use efficiency) and biomass allocation to yield (e.g., harvest index), while environmental variables represent the amount of resource (e.g., light, water, nutrients) and conditions as temperature and CO_2 (Ceccarelli et al., 1991; Slafer and Andrade, 1993; Cooper and Hammer, 1996; Slafer et al., 2014, 1996; Yin et al., 2000b; Reynolds et al., 2009, 2011; Nicotra and Davidson, 2010; Foulkes et al., 2011; Parry et al., 2011; Sadras and Calderini, 2014). The component traits are integrated over time with the environmental inputs to form the target trait: $y_{ij}^{f} = \int f(\mathbf{y}_i^c;\mathbf{z}_j)dt + \varepsilon_{ij}$ (Chenu et al., 2009; Yin and Struik, 2010; Bustos-Korts et al., 2016), where $\int f(\mathbf{y}_i^c;\mathbf{z}_j)dt$ represents the integral over time of the function $f(\mathbf{y}_i^c;\mathbf{z}_j)$ that converts the inputs $\mathbf{y}_i^c$ and $\mathbf{z}_j$ into the target trait, y_{ij}^{f}, and ε_{ij} is an error term.

A crop growth model is a complex reaction norm model that describes how to convert component traits and environmental inputs into the target trait yield. Equally, crop growth models can be seen as devices for the dynamic modeling of multiple traits. An interesting extension of classical physiological crop growth models includes DNA marker variation as the cause underlying phenotypic variation in component traits. Effectively this means that the phenotypic values for the component traits are replaced by predictions from QTL models or genomic prediction models (Yin et al., 2000a, 2003, 2005; Reymond et al., 2003; Tardieu et al., 2005; van Eeuwijk et al., 2005; Chenu et al., 2009; Bogard et al., 2014; Malosetti et al., 2016; Technow et al., 2015).

Alternative approaches for modeling target traits as functions of component traits in their joint development over time are given by Sun and Wu (2015). These authors propose a differential equation framework for modeling the

dynamics of multiple traits in systems genetics, where the constants in the differential equations are themselves modelled as linear functions of underlying QTL genotypes.

The prediction of phenotypes can be improved by modeling intermediate levels of biological variation in between the DNA level (SNPs and sequence information) and the final phenotype (target trait): gene expression, proteins, metabolites, methylation, etc.. Network models are a popular type of model for combining the variation of different types of traits at multiple levels of biological organization, including target phenotypic traits at the highest level (Welch et al., 2003, 2005; Neto et al., 2008, 2010; Scutari et al., 2014; Wang and van Eeuwijk, 2014; Wang et al., 2015). Network models show the behavior of multiple traits in their dependence on each other. Variation in genetic correlations between traits across environmental conditions is an important form of G × E (Malosetti et al., 2008; Alimi et al., 2013). Network models can make such changes visible in a biologically meaningful way.

We see the multi-trait and dynamical modeling perspective offered by crop growth and differential equation modeling with genetic, genomic and environmental inputs as a benchmark for biologically meaningful modeling of G × E. In this framework G × E arises as an emerging property of the model system as all inputs and/or parameters are exclusively indexed by either genotypes or environments, but not by both (Cooper et al., 2002; Hammer et al., 2005, 2006). This exclusive dependence of the phenotype on either genotype or environment is referred to as separability (Gregorius and Namkoong, 1986; Cornelius et al., 1992). In our statistical approaches to G × E, we aim at identifying predictive models that approach as close as possible this ideal of separability.

Reaction norms as crop growth models and differential equations systems are biologically realistic by the emphasis they place on the multi-trait and dynamic aspects of the phenotype. In a 3- to 5-d course on statistical approaches to G × E, little attention can be given to explicit dynamic (Malosetti et al., 2006; Wu and Lin, 2006) and multi-trait modeling approaches. However, simultaneous univariate analyses of traits, especially yield and phenology, can shed light on trait dependencies, while the dynamic behavior of traits can be represented by slope and curvature parameters of reaction norms (van Eeuwijk et al., 2007, 2010; Hurtado-Lopez et al., 2015).

G × E Models in the Literature

Designing a G × E Citation Index

Previous sections gave an overview of our choices for statistical approaches to model G × E as to be presented in a course on G × E. We recommended an approach departing from concrete breeding questions and think mixed models with linear, bilinear and factorial regression terms are most suitable for a three to five day course. As a closing section of our paper, we compare our choices for particular methods to the popularity of methods for G × E analysis in the literature. We analyzed the Web of Science citation reports between 1965 and 2015 (Thomson Reuters, 2015). Our search for keywords related to G × E showed a total of 2275 references. We focused on two groups of references; those that had 20 or more citations (highly cited), and those that were published between 2013 and 2015 (recent literature). Based on our subjective judgement, we found that out of the 447 highly cited papers, 302 corresponded to applications and 175 to methodological papers, discussion papers and reviews. Papers could belong to more than one category. Methodological papers were, again subjectively, classified in 11 categories, depending on the model used (Fig. 2). Most of the 11 categories corresponded to models discussed in previous sections; e.g., mixed models, AMMI, GGE, ANOVA, stability measures. However, other categories were added, like, for example, papers that contained informal G × E analyses without fitting mixed models with different G × E terms, but that simply compared results of single environment QTL analyses. Of course, such an approach allows one to obtain an impression about environment-dependent QTL effects, but ideally one would like to fit a mixed model for simultaneous multi-environment QTL detection (Boer et al., 2007; Alimi et al., 2013).

For Bayesian models and approaches, these papers look at the estimation of genotypic and environment-specific variance components by Bayesian methods (Yang et al., 2007), or propose Bayesian methods to quantify uncertainty for genotypic or environmental scores in AMMI analysis (Josse et al., 2014). Thanks to their heavy use in genomic prediction methods, Bayesian methods are rapidly gaining in popularity in single-environment genetic analysis (Crossa et al., 2010a; Jia and Jannink,

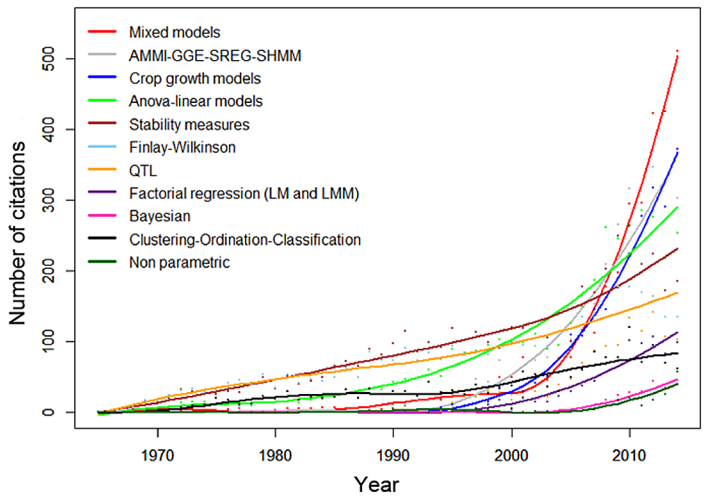

Fig. 2. Number of citations per year between 1965 and 2015 of 175 highly cited methodological papers on Genotype × Environment interaction (G × E), classified by the model that was used. Lines show the fit of a quadratic smoothing spline with three knots.

2012; Spindel et al., 2015). For G × E analysis, Bayesian methods are still less frequent (Fig. 1).

Results of G × E Citation Analysis

When evaluating the use of model categories over time, the number of citations obtained by G × E papers was relatively stable between 1965 and 2003, and it rapidly increased afterward (Fig. 1). Papers using stability measures, fixed Finlay Wilkinson models, ANOVA or simple linear regression models showed the largest number of citations during the last part of the 20th Century. However, the ranking of model use changed with the increase in G × E research starting in 2003. Between approximately 2000 and 2005, citations were still dominated by stability measures, Finlay Wilkinson and ANOVA, but alternative linear-bilinear models like AMMI, GGE, SHMM, and SREG models (Crossa, 2012) increased in popularity. After 2005, the impact of mixed models and crop growth models rapidly grew, to become the dominant category in recent years.

Recently published papers are less likely to have a large number of citations, making it difficult to predict their impact. However, the number of methodological papers that proposes a certain class of models gives an indication of the direction current G × E research is taking and about the possible future impact of these model types. The rapid increase in the number of citations obtained by mixed models and crop growth models (Fig. 2), together with the large number of papers proposing these methods that were published in the last 3 yr (Fig. 3), suggest that in the near future G × E methods will rely less on linear-bilinear models and more on mixed models, crop growth models and Bayesian models. Therefore, our choice in G × E courses for mixed models with bilinear and factorial regression terms for G × E seems to be a good reflection of the current trends in the literature.

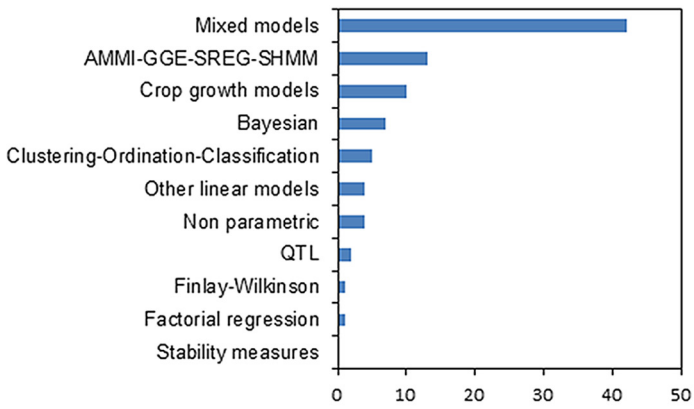

Fig. 3. Number of papers published between 2013 and 2015 for several types of analysis on Genotype × Environment interaction.

Acknowledgments

This paper was written as a contribution to the Integrated Breeding Platform, https://www.integratedbreeding.net/, and as part of the EU DROPS project, FP7-KBBE-2009-3, contract number 244374.

References

Aastveit, A.H., and H. Martens. 1986. ANOVA interactions interpreted by partial least squares regression. Biometrics 42:829–844. doi:10.2307/2530697

Albrecht, T., H.-J. Auinger, V. Wimmer, J. Ogutu, C. Knaak, M. Ouzunova, H.-P. Piepho, and C.-C. Schön. 2014. Genome-based prediction of maize hybrid performance across genetic groups, testers, locations, and years. Theor. Appl. Genet. 127:1–12.

Alimi, N.A., M.C.A.M. Bink, J.A. Dieleman, J.J. Magán, A.M. Wubs, A. Palloix, and F.A. van Eeuwijk. 2013. Multi-trait and multi-environment QTL analyses of yield and a set of physiological traits in pepper. Theor. Appl. Genet. 126(10):2597–2625. doi:10.1007/s00122-013-2160-3

Allard, R.W., and A.D. Bradshaw. 1964. Implications of genotype-environmental interactions in applied plant breeding. Crop Sci. 4(5):503–508. doi:10.2135/cropsci1964.0011183X000400050021x

Annicchiarico, P. 1997. Additive main effects and multiplicative interaction (AMMI) analysis of genotype-location interaction in variety trials repeated over years. Theor. Appl. Genet. 94(8):1072–1077. doi:10.1007/s001220050517

Annicchiarico, P., F. Bellah, and T. Chiari. 2005. Defining Subregions and Estimating Benefits for a Specific-Adaptation Strategy by Breeding Programs. Crop Sci. 45:1741–1749. doi:10.2135/cropsci2004.0524

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: The new crop breeding frontier. Trends Plant Sci. 19(1):52–61. doi:10.1016/j.tplants.2013.09.008

Araus, J.L., G.A. Slafer, C. Royo, and M.D. Serret. 2008. Breeding for yield potential and stress adaptation in cereals. Crit. Rev. Plant Sci. 27(6):377–412. doi:10.1080/07352680802467736

Atlin, G.N., R.J. Baker, K.B. McRae, and X. Lu. 2000. Selection response in subdivided target regions. Crop Sci. 40(1):7–13. doi:10.2135/cropsci2000.4017

Atlin, G.N., K. Kleinknecht, G.P. Singh, and H.P. Piepho. 2011. Managing genotype by environment interaction in plant breeding programs: A selection theory approach. J. Indian Soc. Agric. Stat. 65(2):237–247.

Azzalini, A., and D.R. Cox. 1984. Two new tests associated with analysis of variance. J. R. Stat. Soc. B 46(2):335–343.

Baker, R.J. 1988. Tests for crossover genotype-environmental interactions. Can. J. Plant Sci. 68(2):405–410. doi:10.4141/cjps88-051

Bänziger, M., F.J. Betrán, and H.R. Lafitte. 1997. Efficiency of high-nitrogen selection environments for improving maize for low-nitrogen target environments. Crop Sci. 37:1103–1109. doi:10.2135/cropsci1997.0011183X003700040012x

Boer, M.P., D. Wright, L. Feng, D.W. Podlich, L. Luo, M. Cooper, and F.A. van Eeuwijk. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL by environment interactions, with an example in maize. Genetics 177(3):1801–1813. doi:10.1534/genetics.107.071068

Bogard, M., C. Ravel, E. Paux, J. Bordes, F. Balfourier, S.C. Chapman, J. Le Gouis, and V. Allard. 2014. Predictions of heading date in bread wheat (Triticum aestivum L.) using QTL-based parameters of an ecophysiological model. J. Exp. Bot. 65(20):5849–5865. doi:10.1093/jxb/eru328

Bradshaw, A.D. 2006. Unravelling phenotypic plasticity– why should we bother? New Phytol. 170(4):644–648. doi:10.1111/j.1469-8137.2006.01761.x

Braun, H.-J., S. Rajaram, and M. van Ginkel. 1996. CIMMYT's approach to breeding for wide adaptation. Euphytica 92(1–2):175–183. doi:10.1007/BF00022843

Bro, R., K. Kjeldahl, A.K. Smilde, and H.A.L. Kiers. 2008. Cross-validation of component models: A critical look at current methods. Anal. Bioanal. Chem. 390(5):1241–1251. doi:10.1007/s00216-007-1790-1

Burgueño, J., J. Crossa, P.L. Cornelius, and R.-C. Yang. 2008. Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. Crop Sci. 48(4):1291–1305. doi:10.2135/cropsci2007.11.0632

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype× environment interaction using pedigree and dense molecular markers. Crop Sci. 52(2):707–719. doi:10.2135/cropsci2011.06.0299

Butler, R.C., and P. Brain. 1993. Nonlinear contrasts in ANOVA. Genstat Newsletter 29:20–27.

Bustos-Korts, D., M. Malosetti, S. Chapman, and F.A. van Eeuwijk. 2016. Modelling of genotype by environment interaction and prediction of complex traits across multiple environments as a synthesis of crop growth modelling, genetics and statistics. In: X. Yin and P.C. Struik, editors, Crop systems biology- Narrowing the gaps between crop modelling and genetics. Springer, New York. p. 55–82. doi:10.1007/978-3-319-20562-5_3

Ceccarelli, S. 1989. Wide adaptation: How wide? Euphytica 40(3):197–205.

Ceccarelli, S. 1996. Adaptation to low/high input cultivation. Euphytica 92(1–2):203–214. doi:10.1007/BF00022846

Ceccarelli, S., E. Acevedo, and S. Grando. 1991. Breeding for yield stability in unpredictable environments: Single traits, interaction between traits, and architecture of genotypes. Euphytica 56(2):169–185. doi:10.1007/BF00042061

Ceccarelli, S., and S. Grando. 1996. Drought as a challenge for the plant breeder. Plant Growth Regul. 20(2):149–155. doi:10.1007/BF00024011

Ceccarelli, S., S. Grando, and A. Impiglia. 1998. Choice of selection strategy in breeding barley for stress environments. Euphytica 103(3):307–318. doi:10.1023/A:1018647001429

Chapman, S. 2008. Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. Euphytica 161(1–2):195–208. doi:10.1007/s10681-007-9623-z

Chapman, S.C., S. Chakraborty, M.F. Dreccer, and S.M. Howden. 2012. Plant adaptation to climate change—opportunities and priorities in breeding. Crop Pasture Sci. 63(3):251–268. doi:10.1071/CP11303

Chapman, S.C., M. Cooper, D.G. Butler, and R.G. Henzell. 2000a. Genotype by environment interactions affecting grain sorghum. I. Characteristics that confound interpretation of hybrid yield. Aust. J. Agric. Res. 51(2):197–208. doi:10.1071/AR99020

Chapman, S.C., M. Cooper, and G.L. Hammer. 2002. Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments. Aust. J. Agric. Res. 53(4):379–389. doi:10.1071/AR01070

Chapman, S.C., G.L. Hammer, D.G. Butler, and M. Cooper. 2000b. Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population of environments. Aust. J. Agric. Res. 51(2):223–234. doi:10.1071/AR99022

Chenu, K., S.C. Chapman, F. Tardieu, G. McLean, C. Welcker, and G.L. Hammer. 2009. Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: A "gene-to-phenotype" modeling approach. Genetics 183(4):1507–1523. doi:10.1534/genetics.109.105429

Chenu, K., M. Cooper, G.L. Hammer, K.L. Mathews, M.F. Dreccer, and S.C. Chapman. 2011. Environment characterization as an aid to wheat improvement: Interpreting genotype–environment interactions by modelling water-deficit patterns in North-Eastern Australia. J. Exp. Bot. 62(6):1743–1755. doi:10.1093/jxb/erq459

Chenu, K., R. Deihimfard, and S.C. Chapman. 2013. Large-scale characterization of drought pattern: A continent-wide modelling approach applied to the Australian wheatbelt– spatial and temporal trends. New Phytol. 198(3):801–820. doi:10.1111/nph.12192

Cobb, J.N., G. Declerck, A. Greenberg, R. Clark, and S. McCouch. 2013. Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126(4):867–887. doi:10.1007/s00122-013-2066-0

Comstock, R. 1977. Quantitative genetics and the design of breeding programme. In: Proceedings of the International Conference on Quantitative Genetics. Iowa State Univ., Iowa State Press, Ames, IA. p. 705–718.

Cooper, M. 1999. Concepts and strategies for plant adaptation research in rainfed lowland rice. Field Crops Res. 64(1–2):13–34. doi:10.1016/S0378-4290(99)00048-9

Cooper, M., D.E. Byth, and I.H. DeLacy. 1993. A procedure to assess the relative merit of classification strategies for grouping environments to assist selection in plant breeding regional evaluation trials. Field Crops Res. 35(1):63–74. doi:10.1016/0378-4290(93)90137-C

Cooper, M., S.C. Chapman, D.W. Podlich, and G.L. Hammer. 2002. The GP problem: Quantifying gene-to-phenotype relationships. In Silico Biol. 2(2):151–164.

Cooper, M., C. Gho, R. Leafgren, T. Tang, and C. Messina. 2014a. Breeding drought-tolerant maize hybrids for the US corn-belt: Discovery to product. J. Exp. Bot. 65:6191–6204. doi:10.1093/jxb/eru064

Cooper, M., and G.L. Hammer. 1996. Plant adaptation and crop improvement. CAB International, Wallingford, UK.

Cooper, M., C.D. Messina, D. Podlich, L.R. Totir, A. Baumgarten, N.J. Hausmann, D. Wright, and G. Graham. 2014b. Predicting the future of plant breeding: Complementing empirical evaluation with genetic prediction. Crop Pasture Sci. 65(4):311–336. doi:10.1071/CP14007

Cooper, M., and D.R. Woodruff. 1993. Predicting grain yield in Australian environments using data from CIMMYT international wheat performance trials 3. Testing predicted correlated response to selection. Field Crops Res. 35(3):191–204. doi:10.1016/0378-4290(93)90153-E

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(6):1186–1193. doi:10.2135/cropsci1993.0011183X003300060016x

Cornelius, P.L., and M.S. Seyedsadr. 1997. Estimation of general linear-bilinear models for two-way tables. J. Stat. Comput. Simul. 58(4):287–322. doi:10.1080/00949659708811837

Cornelius, P.L., M. Seyedsadr, and J. Crossa. 1992. Using the shifted multiplicative model to search for "separability" in crop cultivar trials. Theor. Appl. Genet. 84(1–2):161–172.

Corsten, L.C.A., and J.B. Denis. 1990. Structuring interaction in two-way tables by clustering. Biometrics 46(1):207–215. doi:10.2307/2531644

Crossa, J. 2012. From genotype by environment interaction to gene by environment interaction. Curr. Genomics 13(3):225–244. doi:10.2174/138920212800543066

Crossa, J., and P.L. Cornelius. 1997. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. Crop Sci. 37(2):406–415. doi:10.2135/cropsci1997.0011183X003700020017x

Crossa, J., G. de los Campos, P. Perez, D. Gianola, J. Burgueño, J.L. Araus, D. Makumbi, R.P. Singh, S. Dreisigacker, J.B. Yan, V. Arief, M. Banziger, and H.J. Braun. 2010a. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186(2):713–724. doi:10.1534/genetics.110.118521

Crossa, J., S. Perez-Elizalde, D. Jarquin, J.M. Cotes, K. Viele, G. Liu, and P.L. Cornelius. 2011. Bayesian estimation of the additive main effects and multiplicative interaction model. Crop Sci. 51:1458–1469. doi:10.2135/cropsci2010.06.0343

Crossa, J., M. Vargas, C.M. Cossani, G. Alvarado, J. Burgueño, K.L. Mathews, and M.P. Reynolds. 2015. Evaluation and Interpretation of Interactions. Agron. J. 107:736–747. doi:10.2134/agronj2012.0491

Crossa, J., M. Vargas, and A.K. Joshi. 2010b. Linear, bilinear, and linear-bilinear fixed and mixed models for analyzing genotype × environment interaction in plant breeding and agronomy. Can. J. Plant Sci. 90(5):561–574. doi:10.4141/CJPS10003

Crossa, J., R.-C. Yang, and P. Cornelius. 2004. Studying crossover genotype × environment interaction using linear-bilinear models and mixed models. J. Agric. Biol. Environ. Sci. 9(3):362–380. doi:10.1198/108571104X4423

Cullis, B.R., A.B. Smith, C.P. Beeck, and W.A. Cowling. 2010. Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome 53(11):1002–1016. doi:10.1139/G10-080

Dawson, J.C., J.B. Endelman, N. Heslot, J. Crossa, J. Poland, S. Dreisigacker, Y. Manès, M.E. Sorrells, and J.-L. Jannink. 2013. The use of unbalanced historical data for genomic selection in an international wheat breeding program. Field Crops Res. 154:12–22. doi:10.1016/j.fcr.2013.07.020

Debat, V., and P. David. 2001. Mapping phenotypes: Canalization, plasticity and developmental stability. Trends Ecol. Evol. 16(10):555–561. doi:10.1016/S0169-5347(01)02266-2

Denis, J.B. 1988. Two way analysis using covariates. Statistics 19(1):123–132. doi:10.1080/02331888808802080

Denis, J.B. 1991. Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. Rev. Stat. Appl. 39(2):5–24.

DeWitt, T.J., and S.M. Scheiner. 2004. Phenotypic plasticity: Functional and conceptual approaches. Oxford Univ. Press, Oxford.

Eberhart, S.A., and W.A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. 6:36–40. doi:10.2135/cropsci1966.0011183X000600010011x

Eskridge, K.M., and R.F. Mumm. 1992. Choosing plant cultivars based on the probability of outperforming a check. Theor. Appl. Genet. 84(3–4):494–500.

Finlay, K.W., and G.N. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14(6):742–754. doi:10.1071/AR9630742

Fischer, R.A., and R. Maurer. 1978. Drought resistance in spring wheat cultivars. I. Grain yield responses. Aust. J. Agric. Res. 29(5):897–912. doi:10.1071/AR9780897

Foulkes, M.J., G.A. Slafer, W.J. Davies, P.M. Berry, R. Sylvester-Bradley, P. Martre, D.F. Calderini, S. Griffiths, and M.P. Reynolds. 2011. Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. J. Exp. Bot. 62(2):469–486. doi:10.1093/jxb/erq300

Gabriel, K.R. 1971. The biplot graphic display of matrices with application to principal component analysis. Biometrika 58(3):453–467. doi:10.1093/biomet/58.3.453

Gabriel, K.R. 1978. Least squares approximation of matrices by additive and multiplicative models. J. R. Stat. Soc. B 40(2):186–196.

Gail, M., and R. Simon. 1985. Testing for qualitative interactions between treatment effects and patient subsets. Biometrics 41(2):361–372. doi:10.2307/2530862

Gauch, H.G. 1992. Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier, London.

Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 46(4):1488–1500. doi:10.2135/cropsci2005.07-0193

Gauch, H.G. 2013. A simple protocol for AMMI analysis of yield trials. Crop Sci. 53(5):1860–1869. doi:10.2135/cropsci2013.04.0241

Gauch, H.G. 1988. Model selection and validation for yield trials with interaction. Biometrics 44(3):705–715. doi:10.2307/2531585

Gauch, H.G., and R.W. Zobel. 1988. Predictive and postdictive success of statistical analyses of yield trials. Theor. Appl. Genet. 76(1):1–10. doi:10.1007/BF00288824

Gauch, H.G., H.-P. Piepho, and P. Annicchiarico. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. Crop Sci. 48(3):866–889.

Gauch, H.G., and R.W. Zobel. 1997. Identifying mega-environments and targeting genotypes. Crop Sci. 37(2):311–326. doi:10.2135/cropsci1997.0011183X003700020002x

Gelman, A. 2005. Analysis of variance–why it is more important than ever. Ann. Stat. 33(1):1–53. doi:10.1214/009053604000001048

Gogel, B.J., B.R. Cullis, and A.P. Verbyla. 1995. REML estimation of multiplicative effects in multienvironment variety trials. Biometrics 51(2):744–749. doi:10.2307/2532960

Gollob, H. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33(1):73–115. doi:10.1007/BF02289676

Govindaraju, D.R., and B.P. Dancik. 1987. Allozyme heterozygosity and homeostasis in germinating seeds of jack pine. Heredity 59(2):279–283. doi:10.1038/hdy.1987.124

Gregorius, H.R., and G. Namkoong. 1986. Joint analysis of genotypic and environmental effects. Theor. Appl. Genet. 72(3):413–422. doi:10.1007/BF00288581

Gumedze, F.N., and T.T. Dunne. 2011. Parameter estimation and inference in the linear mixed model. Linear Algebra Appl. 435(8):1920–1944. doi:10.1016/j.laa.2011.04.015

Hammer, G.L., S. Chapman, E. van Oosterom, and D.W. Podlich. 2005. Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. Aust. J. Agric. Res. 56(9):947–960. doi:10.1071/AR05157

Hammer, G., M. Cooper, F. Tardieu, S. Welch, B. Walsh, F. van Eeuwijk, S. Chapman, and D. Podlich. 2006. Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci. 11(12):587–593. doi:10.1016/j.tplants.2006.10.006

Hanson, W.D. 1970. Genotypic stability. Theor. Appl. Genet. 40:226–231. doi:10.1007/BF00285245

Harrison, M.T., F. Tardieu, Z. Dong, C.D. Messina, and G.L. Hammer. 2014. Characterizing drought stress and trait influence on maize yield under current and future conditions. Glob. Change Biol. 20(3):867–878. doi:10.1111/gcb.12381

Heslot, N., D. Akdemir, M. Sorrells, and J.-L. Jannink. 2013. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 127:1–18.

Hurtado-Lopez, P.X., B.B. Tessema, S.K. Schnabel, C. Maliepaard, C.G. Van der Linden, P.H.C. Eilers, J. Jansen, F.A. van Eeuwijk, and R.G.F. Visser. 2015. Understanding the genetic basis of potato development using a multi-trait QTL analysis. Euphytica 204:229–241.

Integrated Breeding Platform. 2015. The IBP breeding management system. Version 3.0.8. Integrated Breeding Platform, CyVerse, CIMMYT, Houston, TX. https://www.integratedbreeding.net/ (verified 16 June 2016).

Jannink, J.L., A.J. Lorenz, and H. Iwata. 2010. Genomic selection in plant breeding: From theory to practice. Brief. Funct. Genomics 9(2):166–177. doi:10.1093/bfgp/elq001

Jarquín, D., J. Crossa, X. Lacaze, P. Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez, M. Calus, J. Burgueño, and G. Campos. 2013. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 3:1–13.

Jia, Y., and J.-L. Jannink. 2012. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics 192(4):1513–1522. doi:10.1534/genetics.112.144246

Jolliffe, I.T. 2013. Principal component analysis. Springer, New York.

Josse, J., F. van Eeuwijk, H.-P. Piepho, and J.-B. Denis. 2014. Another look at Bayesian analysis of AMMI models for genotype-environment data. J. Agric. Biol. Environ. Stat. 19(2):240–257.

Josse, J., and F. Husson. 2012. Selecting the number of components in principal component analysis using cross-validation approximations. Comput. Stat. Data Anal. 56(6):1869–1879. doi:10.1016/j.csda.2011.11.012

Kraakman, A.T.W., R.E. Niks, P.M.M.M. Van den Berg, P. Stam, and F.A. van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168(1):435–446. doi:10.1534/genetics.104.026831

Kuehl, R.O. 2000. Design of experiments: Statistical principles of research design and analysis. Thomson Learning, Duxbury.

Kuijken, R.P., F.A. van Eeuwijk, L.F.M. Marcelis, and H. Bouwmeester. 2015. Root phenotyping: From component trait in the lab to breeding. J. Exp. Bot. 66:5385–5387.

Lerner, I.M. 1954. Genetic homeostasis. Oliver and Boyd, London.

Lin, C.S., and M.R. Binns. 1988. A method of analyzing cultivar x location x year experiments: A new stability parameter. Theor. Appl. Genet. 76(3):425–430. doi:10.1007/BF00265344

Lin, C.S., M.R. Binns, and L.P. Lefkovitch. 1986. Stability analysis: Where do we stand? Crop Sci. 26(5):894–900. doi:10.2135/cropsci1986.0011183X002600050012x

Löffler, C.M., J. Wei, T. Fast, J. Gogerty, S. Langton, M. Bergman, B. Merrill, and M. Cooper. 2005. Classification of maize environments using crop simulation and geographic information systems. Crop Sci. 45:1708–1716. doi:10.2135/cropsci2004.0370

Malosetti, M., D. Bustos-Korts, M.P. Boer, and F.A. van Eeuwijk. 2016. Multi environment genomic prediction: Issues in relation to genotype by environment interaction. Crop Sci. doi:10.2135/cropsci2015.05.0311

Malosetti, M., J.-M. Ribaut, and F.A. van Eeuwijk. 2013. The statistical analysis of multi-environment data: Modeling genotype by environment interaction and its genetic basis. Front. Physiol. 4:1–17. doi:10.3389/fphys.2013.00044

Malosetti, M., J. Ribaut, M. Vargas, J. Crossa, and F. van Eeuwijk. 2008. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161(1–2):241–257. doi:10.1007/s10681-007-9594-0

Malosetti, M., R.G.F. Visser, C. Celis-Gamboa, and F.A. van Eeuwijk. 2006. QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato. Theor. Appl. Genet. 113(2):288–300. doi:10.1007/s00122-006-0294-2

Malosetti, M., J. Voltas, I. Romagosa, S.E. Ullrich, and F.A. van Eeuwijk. 2004. Mixed models including environmental covariables for studying QTL by environment interaction. Euphytica 137(1):139–145. doi:10.1023/B:EUPH.0000040511.46388.ef

Mandel, J. 1969. The partitioning of interaction in analysis of variance. J. Res. Natl. Bur. Stand. 73B:309–328. doi:10.6028/jres.073B.031

Mathews, K.L., R. Trethowan, A.W. Milgate, T. Payne, M. van Ginkel, J. Crossa, I. DeLacy, M. Cooper, and S.C. Chapman. 2011. Indirect selection using reference and probe genotype performance in multi-environment trials. Crop Pasture Sci. 62(4):313–327. doi:10.1071/CP10318

Meinshausen, N. 2007. Relaxed lasso. Comput. Stat. Data Anal. 52:374–393. doi:10.1016/j.csda.2006.12.019

Messina, C.D., D. Podlich, Z. Dong, M. Samples, and M. Cooper. 2011. Yield–trait performance landscapes: From theory to application in breeding maize for drought tolerance. J. Exp. Bot. 62(3):855–868. doi:10.1093/jxb/erq329

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4):1819–1829.

Milliken, G.A., and D.E. Johnson. 1989. Analysis of messy data: Nonreplicated experiments. Taylor & Francis, New York.

Möhring, J., and H.-P. Piepho. 2009. Comparison of weighting in two-stage analysis of plant breeding trials. Crop Sci. 49(6):1977–1988. doi:10.2135/cropsci2009.02.0083

Müller, S., J. L. Scealy, and A. H. Welsh. 2013. Model selection in linear mixed models. Stat. Sci. 28(2):135–167.

Nabugoomu, F., R.A. Kempton, and M. Talbot. 1999. Analysis of series of trials where varieties differ in sensitivity to locations. J. Agric. Biol. Environ. Stat. 4(3):310–325. doi:10.2307/1400388

Navabi, A., R.-C. Yang, J. Helm, and D.M. Spaner. 2006. Can spring wheat-growing megaenvironments in the Northern Great Plains be dissected for representative locations or niche-adapted genotypes? Crop Sci. 46(3):1107–1116. doi:10.2135/cropsci2005.06-0159

Neto, E.C., C.T. Ferrara, A.D. Attie, and B.S. Yandell. 2008. Inferring causal phenotype networks from segregating populations. Genetics 179(2):1089–1100. doi:10.1534/genetics.107.085167

Neto, E.C., M.P. Keller, A.D. Attie, and B.S. Yandell. 2010. Causal graphical models in systems genetics. A unified framework for joint inference of causal network and genetic architecture for correlated phenotypes. Ann. Appl. Stat. 4(1):320–339. doi:10.1214/09-AOAS288

Nicotra, A.B., O.K. Atkin, S.P. Bonser, A.M. Davidson, E.J. Finnegan, U. Mathesius, P. Poot, M.D. Purugganan, C.L. Richards, F. Valladares, and M. van Kleunen. 2010. Plant phenotypic plasticity in a changing climate. Trends Plant Sci. 15(12):684–692. doi:10.1016/j.tplants.2010.09.008

Nicotra, A.B., and A. Davidson. 2010. Adaptive phenotypic plasticity and plant water use. Funct. Plant Biol. 37(2):117–127. doi:10.1071/FP09139

Parry, M.A.J., M. Reynolds, M.E. Salvucci, C. Raines, P.J. Andralojc, X.-G. Zhu, G.D. Price, A.G. Condon, and R.T. Furbank. 2011. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. J. Exp. Bot. 62(2):453–467. doi:10.1093/jxb/erq304

Perez-Elizalde, S., D. Jarquin, and J. Crossa. 2012. A general Bayesian estimation method of linear–bilinear models applied to plant breeding trials with genotype × environment interaction. J. Agric. Biol. Environ. Stat. 17(1):15–37. doi:10.1007/s13253-011-0063-9

Pérez-Rodríguez, P., J. Crossa, K. Bondalapati, G. De Meyer, F. Pita, and G. de los Campos. 2015. A Pedigree-based reaction norm model for prediction of cotton yield in multi-environment trials. Crop Sci. 55:1143–1151. doi:10.2135/cropsci2014.08.0577

Piepho, H.-P. 1998. Methods for comparing the yield stability of cropping systems. J. Agron. Crop Sci. 180(4):193–213. doi:10.1111/j.1439-037X.1998.tb00526.x

Piepho, H.P., and J. Möhring. 2005. Best linear unbiased prediction of cultivar effects for subdivided target regions. Crop Sci. 45(3):1151–1159. doi:10.2135/cropsci2004.0398

Piepho, H.-P., J. Möhring, T. Schulz-Streeck, and J.O. Ogutu. 2012. A stage-wise approach for the analysis of multi-environment trials. Biometrical J. 54(6):844–860. doi:10.1002/bimj.201100219

Pigliucci, M. 2001. Phenotypic plasticity: Beyond nature and nurture. Johns Hopkins Univ. Press, Baltimore, MD.

Podlich, D.W., and M. Cooper. 1999. Modelling plant breeding programs as search strategies on a complex response surface. In: B. McKay, X. Yao, C. Newton, J.-H. Kim, and T. Furuhashi, editors, Simulated evolution and learning SE-23. Lecture notes in computer science. Springer, Berlin. p. 171–178. doi:10.1007/3-540-48873-1_23

Reymond, M., B. Muller, A. Leonardi, A. Charcosset, and F. Tardieu. 2003. Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiol. 131(2):664–675. doi:10.1104/pp.013839

Reynolds, M., D. Bonnett, S.C. Chapman, R.T. Furbank, Y. Manes, D.E. Mather, and M.A.J. Parry. 2011. Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. J. Exp. Bot. 62(2):439–452. doi:10.1093/jxb/erq311

Reynolds, M., M.J. Foulkes, G.A. Slafer, P. Berry, M.A.J. Parry, J.W. Snape, and W.J. Angus. 2009. Raising yield potential in wheat. J. Exp. Bot. 60(7):1899–1918. doi:10.1093/jxb/erp016

Sadras, V.O., and D. Calderini. 2014. Crop Physiology: Applications for genetic improvement and agronomy. Elsevier Science, London.

Sadras, V.O., and C. Lawson. 2011. Genetic gain in yield and associated changes in phenotype, trait plasticity and competitive ability of South Australian wheat varieties released between 1958 and 2007. Crop Pasture Sci. 62(7):533–549. doi:10.1071/CP11060

Sadras, V.O., and G.J. Rebetzke. 2013. Plasticity of wheat grain yield is associated with plasticity of ear number. Crop Pasture Sci. 64(3):234–243. doi:10.1071/CP13117

Sadras, V.O., M.P. Reynolds, A.J. de la Vega, P.R. Petrie, and R. Robinson. 2009. Phenotypic plasticity of yield and phenology in wheat, sunflower and grapevine. Field Crops Res. 110(3):242–250. doi:10.1016/j.fcr.2008.09.004

SAS Institute. 2015. SAS/STAT. SAS Inst., Cary, NC.

Scutari, M., P. Howell, D.J. Balding, and I. Mackay. 2014. Multiple quantitative trait analysis using Bayesian networks. Genetics 198(1):129–137. doi:10.1534/genetics.114.165704

Schlichting, C.D., and M. Pigliucci. 1998. Phenotypic evolution: A reaction norm perspective. Sinauer Associates Incorporated, Sunderland, MA.

Schulz-Streeck, T., J.O. Ogutu, Z. Karaman, C. Knaak, and H.P. Piepho. 2012. Genomic selection using multiple populations. Crop Sci. 52(6):2453–2461. doi:10.2135/cropsci2012.03.0160

Shukla, G.K. 1972. Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29(2):237–245. doi:10.1038/hdy.1972.87

Slafer, G.A., and F.H. Andrade. 1993. Physiological attributes related to the generation of grain yield in bread wheat cultivars released at different eras. Field Crops Res. 31(3–4):351–367. doi:10.1016/0378-4290(93)90073-V

Slafer, G.A., D.F. Calderini, and D.J. Miralles. 1996. Yield components and compensation in wheat: Opportunities for further increasing yield potential. In: M.P. Reynolds, S. Rajaram, and A. McNab, editors, Increasing yield potential in wheat: Breaking the barriers. CIMMYT, Sonora, Mexico.

Slafer, G.A., R. Savin, and V.O. Sadras. 2014. Coarse and fine regulation of wheat yield components in response to genotype and environment. Field Crops Res. 157(0):71–83. doi:10.1016/j.fcr.2013.12.004

Smith, A., B. Cullis, and A. Gilmour. 2001. Applications: The analysis of crop variety evaluation data in Australia. Aust. N. Z. J. Stat. 43(2):129–145. doi:10.1111/1467-842X.00163

Smith, A.B., B.R. Cullis, and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. J. Agric. Sci. 143:449–462. doi:10.1017/S0021859605005587

Snape, J.W., M.J. Foulkes, J. Simmonds, M. Leverington, L. Fish, Y. Wang, and M. Ciavarrella. 2007. Dissecting gene × environmental effects on wheat yields via QTL and physiological analysis. Euphytica 154(3):401–408. doi:10.1007/s10681-006-9208-2

Spindel, J., H. Begum, D. Akdemir, P. Virk, B. Collard, E. Redoña, G. Atlin, J.-L. Jannink, and S.R. McCouch. 2015. Genomic selection and association mapping in rice *(Oryza sativa)*: Effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet. 11(2):E1004982. doi:10.1371/journal.pgen.1004982

Sun, L., and R. Wu. 2015. Mapping complex traits as a dynamic system. Phys. Life Rev. 13:155–185.

Tardieu, F., M. Reymond, B. Muller, C. Granier, T. Simonneau, W. Sadok, and C. Welcker. 2005. Linking physiological and genetic analyses of the control of leaf growth under changing environmental conditions. Aust. J. Agric. Res. 56(9):937–946. doi:10.1071/AR05156

Tardieu, F., and R. Tuberosa. 2010. Dissection and modelling of abiotic stress tolerance in plants. Curr. Opin. Plant Biol. 13(2):206–212. doi:10.1016/j.pbi.2009.12.012

Taylor, J.D., A.P. Verbyla, C. Cavanagh, and M. Newberry. 2012. Variable selection in linear mixed models using an extended class of penalties. Aust. N. Z. J. Stat. 54(4):427–449. doi:10.1111/j.1467-842X.2012.00687.x

Technow, F., C.D. Messina, L.R. Totir, and M. Cooper. 2015. Integrating crop growth models with whole genome prediction through approximate Bayesian computation. PloS one 10(6):E0130855.

Thomson Reuters. 2015. Web of Science. http://apps.webofknowledge.com/ (verified 16 June 2016).

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. B (58):267–288.

Trethowan, R.M., J. Crossa, M. van Ginkel, and S. Rajaram. 2001. Relationships among bread wheat international yield testing locations in dry areas. Crop Sci. 41:1461–1469. doi:10.2135/cropsci2001.4151461x

van Eeuwijk, F.A. 1992. Interpreting genotype by environment interaction using redundancy analysis. Theor. Appl. Genet. 85(1):89–100. doi:10.1007/BF00223849

van Eeuwijk, F. 2015. How to dissect complex traits and how to choose suitable mapping resources for system genetics?: Comment on "Mapping complex traits as a dynamic system" by L. Sun and R. Wu. Phys. Life Rev. 13:186–189.

van Eeuwijk, F.A., M.C.A.M. Bink, K. Chenu, and S.C. Chapman. 2010. Detection and use of QTL for complex traits in multiple environments. Curr. Opin. Plant Biol. 13(2):193–205. doi:10.1016/j.pbi.2010.01.001

van Eeuwijk, F.A., J.B. Denis, and M.S. Kang. 1996. Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: M.S. Kang and H.G. Gauch, editors, Genotype by environment interaction. Taylor & Francis Group, Boca Raton. p. 15–49. doi:10.1201/9781420049374.ch2

van Eeuwijk, F.A., M. Malosetti, and M.P. Boer. 2007. Modelling the genetic basis of response curves underlying genotype by environment interaction. In: J.H.J. Spiertz, P.C. Struik and H.H. van Laar, editors, Scale and complexity in plant systems research: Gene-plant-crop relations. Spriner, Wageningen, The Netherlands. p. 113–124.

van Eeuwijk, F.A., M. Malosetti, X. Yin, P.C. Struik, and P. Stam. 2005. Statistical models for genotype by environment data: From conventional ANOVA models to eco-physiological QTL models. Aust. J. Agric. Res. 56(9):883–894. doi:10.1071/AR05153

Van Oijen, M., and M. Höglind. 2015. Toward a Bayesian procedure for using process-based models in plant breeding, with application to ideotype design. Euphytica 207(3):627–643.

Vargas, M., J. Crossa, F.A. van Eeuwijk, M.E. Ramírez, and K. Sayre. 1999. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype × environment interaction. Crop Sci. 39(4):955–967. doi:10.2135/cropsci1999.0011183X003900040002x

Verbeke, G., and G. Molenberghs. 2009. Linear mixed models for longitudinal data. Springer Science & Business Media, New York.

Vargas, M., B. Glaz, G. Alvarado, J. Pietragalla, A. Morgounov, Y. Zelenskiy, and J. Crossa. 2015. Analysis and interpretation of interactions in agricultural research. Agron. J. 107:748–762. doi:10.2134/agronj13.0405

Voltas, J., F.A. van Eeuwijk, J.L. Araus, and I. Romagosa. 1999a. Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley II.: Grain growth. Field Crops Res. 62(1):75–84. doi:10.1016/S0378-4290(99)00007-6

Voltas, J., I. Romagosa, A. Lafarga, A.P. Armesto, A. Sombrero, and J.L. Araus. 1999b. Genotype by environment interaction for grain yield and carbon isotope discrimination of barley in Mediterranean Spain. Aust. J. Agric. Res. 50(7):1263–1271. doi:10.1071/AR98137

VSN International. 2015a. GenStat. 17th ed. VSN International Limited, Hemel Hempstead, UK.

VSN International. 2015b. ASREML for R. Version 3.0. VSN International Limited, Hemel Hempstead, UK.

VSN International. 2015c. ASREML for Windows. Version 4.0. VSN International Limited, Hemel Hempstead, UK.

Wang, H., and F.A. van Eeuwijk. 2014. A new method to infer causal phenotype networks using QTL and phenotypic information. PLoS One 9(8):E103997. doi:10.1371/journal.pone.0103997

Wang, H., J. Paulo, W. Kruijer, M. Boer, H. Jansen, Y. Tikunov, B. Usadel, S. van Heusden, A. Bovy, and F. van Eeuwijk. 2015. Genotype-phenotype modeling considering intermediate level of biological variation: A case study involving sensory traits, metabolites and QTLs in ripe tomatoes. Mol. Biosyst. 11(11):3101–3110. doi:10.1039/C5MB00477B

Welch, S.M., Z. Dong, J.L. Roe, and S. Das. 2005. Flowering time control: Gene network modelling and the link to quantitative genetics. Aust. J. Agric. Res. 56(9):919–936. doi:10.1071/AR05155

Welch, S.M., J.L. Roe, and Z. Dong. 2003. A genetic neural network model of flowering time control in *Arabidopsis thaliana*. Agron. J. 95:71–81. doi:10.2134/agronj2003.0071

Welham, S.J., S.A. Gezan, S.J. Clark, and A. Mead. 2014. Statistical methods in biology: Design and analysis of experiments and regression. Taylor & Francis, Boca Raton, FL.

Welham, S.J., B.J. Gogel, A.B. Smith, R. Thompson, and B.R. Cullis. 2010. A comparison of analysis methods for late-stage variety evaluation trials. Aust. N. Z. J. Stat. 52(2):125–149. doi:10.1111/j.1467-842X.2010.00570.x

Windhausen, V.S., S. Wagener, C. Magorokosho, D. Makumbi, B. Vivek, H.-P. Piepho, A.E. Melchinger, and G.N. Atlin. 2012. Strategies to subdivide a target population of environments: Results from the CIMMYT-led maize hybrid testing programs in Africa. Crop Sci. 52(5):2143–2152. doi:10.2135/cropsci2012.02.0125

Woltereck, R. 1909. Weitere experimentelle Untersuchungen über Artveränderung, speziel über das Wesen quantitativer Artunterschiede bei Daphnien. Verhandlungen der Dtsch. Zool. Gesellschaft (19):110–173.

Wricke, G. 1962. Uber eine methode zur erfassung der okologischen streubreite in feldversuchen. Zeitschrift fur Pflanzenzuchtung 47:92–96.

Wricke, G. 1966. A biometric method of comprehending the environmental adaptation. (English translation.) Qualitas Plantarum et Materiae Vegetabiles 13(1–4):318–319. doi:10.1007/BF01103423

Wu, R., and M. Lin. 2006. Functional mapping-how to map and study the genetic architecture of dynamic complex traits. Nat. Rev. Genet. 7(3):229–237. doi:10.1038/nrg1804

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Sci. 40(3):597–605. doi:10.2135/cropsci2000.403597x

Yan, W., and M.S. Kang. 2002. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL. doi:10.1201/9781420040371

Yan, W., and I. Rajcan. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. Crop Sci. 42(1):11–20. doi:10.2135/cropsci2002.0011

Yang, R.-C. 2007.Mixed-model analysis of crossover genotype–environment interactions. Crop Sci. 47(3)): 1051–1062. doi:10.2135/cropsci2006.09.0611

Yang, R.-C., J. Crossa, P.L. Cornelius, and J. Burgueño. 2009. Biplot analysis of genotype × environment interaction: Proceed with caution. Crop Sci. 49(5):1564–1576. doi:10.2135/cropsci2008.11.0665

Yang, J., J. Zhu, and R.W. Williams. 2007. Mapping the genetic architecture of complex traits in experimental populations. Bioinformatics 23(12):1527–1536. doi:10.1093/bioinformatics/btm143

Yates, F., and W.G. Cochran. 1938. The analysis of groups of experiments. J. Agric. Sci. 28:556–580. doi:10.1017/S0021859600050978

Yin, X., S.D. Chasalow, C.J. Dourleijn, P. Stam, and M.J. Kropff. 2000a. Coupling estimated effects of QTLs for physiological traits to a crop growth model: Predicting yield variation among recombinant inbred lines in barley. Heredity 85(6):539–549. doi:10.1046/j.1365-2540.2000.00790.x

Yin, X., A.H.C.M. Schapendonk, M.J. Kropff, M. van Oijen, and P.S. Bindraban. 2000b. A generic equation for nitrogen-limited leaf area index and its application in crop growth models for predicting leaf senescence. Ann. Bot. 85(5):579–585. doi:10.1006/anbo.1999.1104

Yin, X., P. Stam, M.J. Kropff, and A.H.C.M. Schapendonk. 2003. Crop modeling, QTL mapping, and their complementary role in plant breeding. Agron. J. 95(1):90–98. doi:10.2134/agronj2003.0090

Yin, X., and P.C. Struik. 2010. Modelling the crop: From system dynamics to systems biology. J. Exp. Bot. 61(8):2171–2183. doi:10.1093/jxb/erp375

Yin, X., P.C. Struik, F.A. van Eeuwijk, P. Stam, and J. Tang. 2005. QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. J. Exp. Bot. 56(413):967–976. doi:10.1093/jxb/eri090

Zhang, X., P. Perez-Rodriguez, K. Semagn, Y. Beyene, R. Babu, M.A. Lopez-Cruz, F. San Vicente, M. Olsen, E. Buckler, J.-L. Jannink, B.M. Prasanna, and J. Crossa. 2015. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. Heredity 114(3):291–299. doi:10.1038/hdy.2014.99