44

Selection and G x E: Advanced Topics

One can hardly expect a single cultivar of a crop to flourish the world over under all environments and management practices. — Gauch and Zobel (1997)

Version 26 June 2014

While the covariance approach used in Chapter 38 (treating trait expression in different environments as correlated characters) provides a considerable amount of insight into selection when G x E is present, operationally it is not very useful in most real-world settings when more than two environments are considered. This chapter examines some of the tools that have been developed to examine at trait performance over multiple environments and how selection decisions can benefit from using these tools.

Three general topics are considered. The first is the measurement, and subsequent selection, of stability when genetic groups are scored over multiple environments. In the previous chapter, we examined the simplest case of just two environments where the difference in means provides a general measure of stability. With more than two environments, this metric no longer works. Indeed, with multiple environments there are several different types of stability that a breeder may be interested in, each of which has multiple measures.

Next, we develop a number of statistic tools that allow a breeder or experimentalist to look for structure when presented with a table of G x E interactions. The notion of structure is inherent in Falconer's view that G x E can be treated as a correlated traits problem. Another consequence of structure is the concept of mega-environments (Chapter 38). When comparing genotype performance within the set of environments that comprises a particular mega-environment, by definition little G x E is apparent. However, considerable G x E can occur when comparisons are made between different mega-environments. The foundation of methods that search for structure involve approximating the G x E interaction term by one (or a few) multiplicative terms comprised of variables that are constant over genotypes or environments (e.g., approximating GE_{ij} by $\alpha_i\gamma_j$ or, more generally, $\sum_k \alpha_{ki}\gamma_{kj}$). These are examples of **biadditive** (or **bilinear**) models, which includes the Finlay-Wilkinison regression, **AMMI** (**additive main effects, multiplicative interactions**) models, and **factorial regressions**. Graphic plots based on coefficients of biadditive models (often called **biplots**) provide a powerful visualization tool for detecting genotypes with similar response in particular environments as well as identifying mega-environments. Factorial regressions incorporate measured environmental factors (such as soil moisture or total days above a set temperature) and represent an attempt to bridge holistic black-box models of the environment with models that explicitly incorporate measured environmental variables. This introduces the notion of a reaction norm (the performance of a trait as a function of a measured environmental value), a topic examined in greater detail in Chapter 42.

Our final topic, intimately related with G x E structure, is the use of mixed models to estimate G x E effects. Recall that the covariance structure of the random effects is a critical feature of mixed-model analysis (Chapters 15, 35). As we will see, consideration of this covariance structure in models of G x E underpins various stability estimates. Thus, mixed-models provide a deep theoretical foundation for G x E analysis. On a more practical level, they also allow for the analysis of unbalanced data and (through BLUP) provide shrinkage

estimates which borrow information from the entire dataset to improve the estimation of genetic and G x E effects. We conclude by closing a loop left open in Chapter 38 by showing how structured G x E matrices (and their associated mixed-model estimates) are connected to the $\mathbf{G}$ matrix when using the multivariate breeders' equation to examine selection under G x E.

LAYING THE FOUNDATION: THE BASIC MODEL FOR G x E

The central theme in this chapter is detecting and characterizing structure in G x E interactions, which is also intimately connected to measures of stability. To do so, we need to build up a rich class of statistical models for examining G x E. We start here with the base model and then successively build on this throughout the chapter. Let z_{ijk} denote the value of the kth replicate of genotype i in environment j, where

$$z_{ijk} = \mu_{ij} + \epsilon_{ijk} \tag{39.1a}$$

Here μ_{ij} is the mean value of genotype i in environment j and ϵ_{ijk} is the residual, which has expected value zero. Residuals are initially assumed to be uncorrelated with constant variance σ_e^2, implying

$$z_{ijk} \sim (\mu_{ij}, \sigma_e^2) \tag{39.1b}$$

where the notation $x \sim (a, b)$ denotes that the random variable x has mean a and variance b. The importance of Equation 39.1 is the conceptual separation of modeling G x E effects (the structure of μ_{ij}) from modeling the residual effects (the covariance structure of the ϵ_{ijk}). Denote the sample mean for genotype i in environment j by z_{ij}, where

$$z_{ij} = \mu_{ij} + \overline{\epsilon}_{ij}, \quad \text{with} \quad \overline{\epsilon}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} \epsilon_{ijk} \tag{39.1d}$$

Assuming residuals are uncorrelated with constant variance σ_e^2,

$$z_{ij} \sim \left(\mu_{ij}, \sigma_e^2 / n_{ij}\right) \tag{39.1d}$$

Discussions of stability metrics for a genotype (i) involve measures of how the μ_{ij} varies over environments, as we wish to separate variation in performance over environments from variation due to the residual error.

The foundational model of G x E interactions is

$$\mu_{ij} = \mu + G_i + E_j + GE_{ij} \tag{39.2a}$$

where GE_{ij} denotes the interaction between genotype (or line or variety or family) i and environment j. Throughout this chapter we will build upon this basic model, namely by modeling GE_{ij} as the product of genotype (γ_i) and environmental (η_j) factors plus a residual δ_{ij}, so that Equation 39.2a has the form

$$\mu_{ij} = \mu + G_i + E_j + \left(\sum_{k=1}^{m} \gamma_{ki}\, \eta_{kj} + \delta_{ij} \right) \tag{39.2b}$$

While this, at first blush, may seem to be adding more complexity and confusion to the foundational model, as we develop throughout this chapter, such a bilinear model can provide considerable insight.

MEASURING AND SELECTING FOR STABILITY

Ideally, a breeder wants to select for consistently high performance over all environments, so that the desired goal is to improve *both* stability and mean performance. Perhaps a more realistic goal is to select for favorable genotypes that show some minimal level of performance over all environments while at the same time being able to response well to favorable environments (Verma and Chahal 1978). Both these goals lead to several subtle questions beyond the obvious issue of tradeoffs in stability vs. performance. The first is what exactly do we mean by consistency? For example, in some settings consistency is the identical performance of a genotype across all environments. This might be preferred when selecting a crop for a marginal region where the breeder is willing to sacrifice high performance in some years for consistent performance over all years. Conversely, over a large region, a breeder typically would like a crop to response well in favorable environments, but to do so in a predictable fashion. These conflicting goals lead to different notions of stability. The second issue is how do we best measure our particular notion of stability. Finally, how do we best select for our notion of stability and what tradeoffs may occur with performance? We examine these issues in turn.

Static vs. Dynamic Stability

Consistency is a function of predictability. If a feature is unpredictable, then consistency requires that performance remain relatively constant over different feature values. Conversely, if a feature is somewhat predictable, then consistency can refer to the reliability of the predicted values – line performance may change dramatically over environments, but if the performance changes in a highly *predictable* fashion we have consistency. Thus, we wish stability over unpredictable aspects of the environment (temporal features such as year to year variation in weather), but we also want genotypes that can *respond* to predictable environmental changes (spatial features such as different locations or soil types) in a consistent fashion. This schizophrenic view of desiring no change in some settings and (predictable) response in others leads to two very different basic notions of stability: static and dynamic (Becker and Léon 1988).

Under **static stability**, a genotype displays a small variance over environments. This is akin to the early notion of breeders (e.g., Römer 1917) that the performance of a line should remain relatively constant. Becker (1981) calls this **biological concept of stability**. Under **dynamic stability**, a genotype's preformance responds in a consistent fashion to changes in the environment. Becker suggests the term **agronomic concept of stability** for this, as the breeder does not want to penalize genotypes for responding to favorable environments, *provided* they respond in a predictable fashion.

The static stability of a line (i) is measured by its phenotypic variance over environments, which is given by the variance among the μ_{ij} over environments around their mean value of μ_i,

$$\sigma^2(\mu_i) = \sigma^2_E + \sigma^2_{G_i \times E} \tag{39.3a}$$

where $\sigma^2_{G_i \times E}$ is the variance in G x E for genotype i. From Equation 39.2a, the difference in performance for genotype i in two environments is giving by

$$\mu_{ij} - \mu_{ik} = (E_j - E_k) + (GE_{ij} - GE_{ik}) \tag{39.3b}$$

illustrating that both additive responses to change in the environment ($E_j - E_k$) and genotype x environment interactions generate between-environment variability and hence both σ^2_E and $\sigma^2_{G_i \times E}$ decrease the static stability of a genotype. Similarly, the change in the mean values of lines i and k in environment j is just

$$\mu_{ij} - \mu_{ik} = (G_i - G_j) + (GE_{ij} - GE_{kj}) \tag{39.3c}$$

In the absence of G x E, the difference between lines i and k is $G_i - G_j$. Thus, with additivity of G and E (i.e., no G x E so that $\mu_{ij} = \mu + G_i + E_j$) the *relative* performance of two lines is independent of the environment. Under dynamic stability, lines are not penalized for their additive responses to the environment, while departures from additivity ($\sigma^2_{G_i \times E}$) are penalized, with the most unstable genotypes being those with the largest interaction variance. As we will shortly see, the notion of dynamic stability can be expanded to consider those genotypes showing the least variation about the *predictable* responses to the environment to be the most stable.

The discerning reader will immediately note two issues. First, these measures are highly *context-specific*, as the fit of the various terms in Equation 39.2a is completely dependent upon the genotypes and environments chosen. Adding or subtracting a few key genotypes or environments can create (or destroy) a considerable amount of G x E. Thus, the stability of a genotype is always a *relative* concept, being a function of the genotypes and environments in our sample. The second issue is the environment E itself. As we have continually stressed, there are predicable and unpredictable features of an environment. Ideally, we wish genotypes that are responsive over predictive features (such as locations) but stable over unpredictable features (such as different years). Thus, a breeder may actually want genotypes with high interaction variances for location (to exploit the best environments) but low interaction variances for year and location-year effects. Again, there is a context-dependency to these measures, which will vary with the objectives of the breeder. A final consideration is the *scale* over which breeding take place. If most of the locations are fairly similar (as might occur when a small region is considered), then static stability may be appropriate, as most of the variation might be due to yearly effects, whose impacts the breeder wishes to minimize. Conversely, if the spatial scale of the breeding is vast, location effects can be considerable and dynamic stability is more appropriate.

Measuring Stability: Finlay-Wilkinson Regressions

In order to expand the notion of dynamic stability to allow for predictable changes in response, we start with the **Finlay-Wilkinson regression** (LW Chapter 22). Recall that this approach, also know as the **joint regression** or **regression on an environmental index**, fits the best linear regression of the performance of a genotype on some measure of the environment. Its roots trace back to Mooers (1921), Stringfeild and Salter (1934), and Yates and Cochran (1938), but it was popularized by Finlay and Wilkinson (1963) who suggested that a very natural measure of the "quality" of an environment is the mean value E_j over all genotypes in that environment. Here, the basic model for the mean value of genotype i in environment j is given by

$$\mu_{ij} = \mu + G_i + E_j(1 + \beta_i) + \delta_{ij} \tag{39.4a}$$

where β_i measures the **responsiveness** of genotype i to the environment and δ_{ij} is the residual (based on a seperate regressions for each genotype). An equivalent form of this model frequently appears in the literature,

$$\mu_{ij} = \mu + G_i + E_j\alpha_i + \delta_{ij} \tag{39.4b}$$

where $\alpha_i = 1 + \beta_i$ and hence departures of α from one indicate genotype-environment interactions. The literature is almost evenly split on these two alternative (but essentially equivalent) formulations. A source of confusion is that α_i in Equation 39.4b is often denoted by β_i in the literature, so a little care must be taken when reading a paper to see whether β refers to departures from zero (Equation 39.4a) or departures of the slope from one (Equation 39.4b).

The various statistical issues with fitting Equation 39.4a/b are reviewed in LW Chapter 22, as well as by Freeman and Perkins (1971), Fripp and Caten (1971), Wright (1976), Miezan et al. (1979), Skrøppa (1984), Crossa (1990), and Piepho (1998a). Comparing Equation 39.4a with our base model of G x E (Equation 39.4a), we see that

$$GE_{ij} = \beta_i E_j + \delta_{ij} \tag{39.4c}$$

The Finlay-Wilkinson regression attempts to approximate the G x E interaction term as the product of a genotypic-specific environmental sensitivity β_i and the relative mean performance of the environment E_j. The elegance of Equation 39.4c is that the $n_g + n_e$ terms (for the β_i and E_j, respectively) can potentially account for most of the variation in $n_g n_e$ interaction terms. The error in predicting GE from this simple multiplicative model is given by δ_{ij}. The model residual has expected value zero over the regression (i.e., when averaged over the E_j values for each genotype-specific regression). If the variance in the model residual is large, so that $\sigma^2_{\delta_i}$ is a large fraction of $\sigma^2(G_i E)$, then the linear approximation is a poor fit and most of the interaction variance remains.

Equation 39.4 shows that β_i is a measure of the **linear sensitivity** of the genotype to the environment while δ_{ij} is a measure of its **nonlinear sensitivity** (Fripp and Caten 1971). Specifically, β_i measures the linear **response** of genotype i to changes in the environment, while δ_{ij} measures departures from this response, which can arise from sampling error, improperly fitting a linear function to a true nonlinear response, or both. Finlay-Wilkinson regressions form a useful starting point for more elaborate statistical methods (such as bilinear models and biplots) that search for structure over a complex set of genotype x environment interactions.

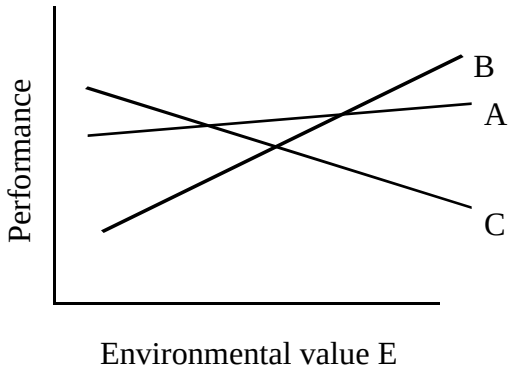


Figure 39.1. The Finlay-Wilkinson regression attempts to predict the performance of a genotype/line as a linear function of the environmental value E, measured (here) as the mean trait value for that environment (for example, the average yield over all genotypes in that environment). A flat slope indicates little change across environments (static stability), with the genotype having roughly equal performance in both poor (low yield) and good (high yield) environments. Genotype *A* shows this behavior, being outperformed by *C* in poor environments and by *B* in good environments, and yet has the highest mean performance over all environments. Conversely, a large (absolute) slope indicates a *lack* of static stability, with the genotype responding to the environment. A slope greater than one indicates over-performance (with respect to the other genotypes in the sample) in high environments (genotype *B*), while a slope less than one indicates over-performance in poor environments (genotype *C*). A Finlay-Wilkinson plot is a convenient way to visually summarize both mean performance and stability for a series of genotypes.

Figure 39.1 shows a hypothetical plot of the mean of three genotypes as a function of

the environmental values E. The slope of this regression for genotype i is $(1+\beta_i)$. A slope of one ($\beta_i = 0$) implies no linear trend in G x E. Such a genotype would *not* show static stability, as its value changes with the environment ($\mu_{ij} = \mu + G_i + E_j$). Static stability (a genotype whose mean values are relatively unchanged over environments, genotype A in Figure 39.1) requires a value of β_i near -1, giving a slope near zero. Note that this is *strong* genotype x environment interaction (relative to the other genotypes in the sample) so that static stability often requires strong interactions if other genotypes in the sample show responses over environments. Conversely, genotypes with $\beta_i > 0$ over-perform in good environments and under-perform in bad ones, with the converse being true for $\beta_i < 0$. Genotype C in Figure 39.1 out-performs the other two genotypes in poor environments, while genotype B is the best performer in good environments. However, genotype A gives the best average performance over all environments. Hence, A might be selected if a single line is to be chosen for wide adaptivity, while if specialized lines for low and high environments are desired, these would be C and B, respectively.

It is important to stress that the β_i from Equation 39.2 are obtained by the best fit over *all* genotypes and environments in the sample. The particular value of β_i obtained is not just a function of the genotype i and the environments in the sample, but also of the other genotypes as well (as these, in turn, influence the values of E_j). Further, the regression slope can be highly leveraged by one or two extreme values, so that the performance in the single best, or worst, environment can have a highly disproportionate impact on the regression slope. When the genotype x environment table is incomplete (not all genotypes were replicated over all environments), least-squares estimates of β_i can still be obtained using an iterative procedure (Digby 1979). As we show alter, mixed models easily account for missing data.

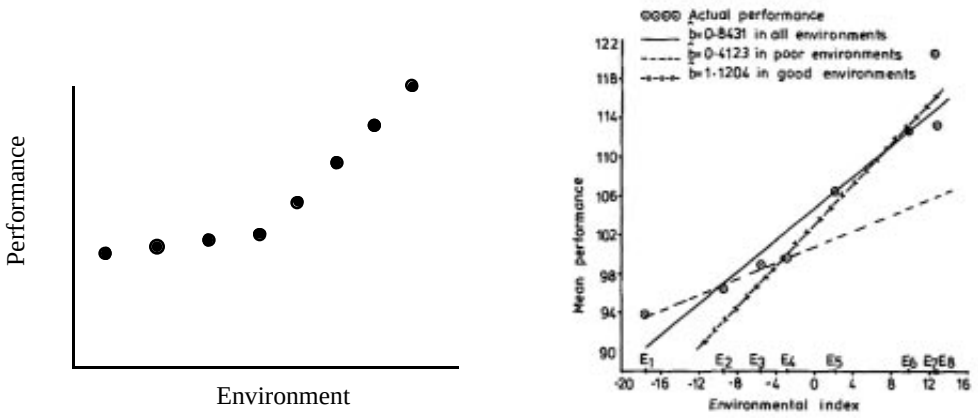


Figure 39.2. **Left.** The behavior of an ideal genotype, whose performance does not suffer in poor environments (low responsiveness) while at the same time responding favorably in good environments. **Right.** Verma and Chahal (1978) found that an inbred line of *Nicotiana rustica* showed close to this pattern over the eight environments in which it was scored. Open circles are the actual values, while the three regression lines correspond to the slope using either all, or only good, or only poor environments.

Splitting the Regression: The Varma-Chahal Modification

Ideally, a genotype should be insensitive to environmental change in a poor environment and responsive in a good one. The behavior of such a hypothetical genotype is shown in Figure

39.2 (left). Clearly, the response with environment value is nonlinear in this case. This results in the slope of a Finlay-Wilkinson regression being very misleading indicator of desirability, to the point that such genotypes could be discarded if selection is solely based on regression slopes. Verma and Chahal (1978) suggested that this problem can be avoided by splitting the regression — fitting separate regression for good and bad environments. The idea genotype would be very unresponsive in poor environments ($1+\beta_p \simeq 0$) and quite responsive in good environments ($\beta_g >> 0$). As shown in Figure 39.2 (right), Verma and Chahal found that one of the ten inbred lines of *Nicotiana rustica* they examined showed close to this ideal behavior. Fitting the response over all environments with a single regression resulted in a slope of 0.84, while the slope for the poor environment was 0.41 (showing below-average response) and for the good environments 1.12 (showing above-average response).

Mean Performance Versus Stability: Insights from Finlay-Wilkinson Regressions

Simmonds (1991) has argued that selection for improvement in high-yielding environments not only selects for lines with higher mean performance, but also for lines that are more responsive to high environments, thus having a larger Finlay-Wilkinson slope (and less stability). This increased responsiveness to high environments (from the larger slope) comes at the expense of under-performance in low environments. As a result, Simmonds argues that lines selected for high performance in high-yield environments may often under-perform other lines in low-yield environments.

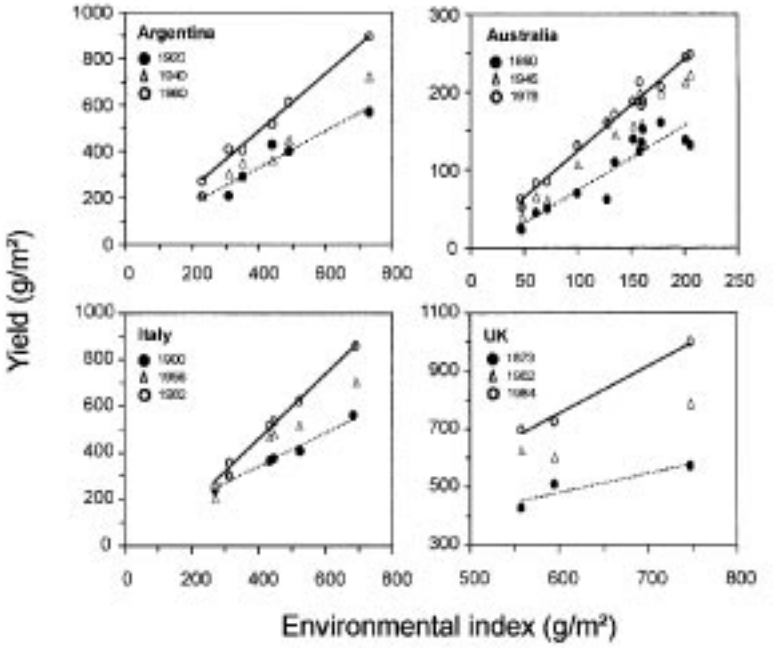


Figure 39.3. Finlay-Wilkinson plots of wheat lines from different eras for four different countries. While newer lines show higher yield across the range of environments examined, they also show decreased stability, as indicated by their greater slopes. After Calderini and Slafer (1999).

How much support is there for this idea? Dashiell et al. (1994) found a positive correla-

tion between Finlay-Wilkinson regression slopes and seed yield in a set of 18 soybeans tested in five locations in Nigeria, indicating that the higher-yielding lines were more responsive to environmental changes. Thus, within this set of lines, Simmonds' suggestion holds. Conversely, Léon (1986) did not find a significant correlation between yield and regression slope in winter wheat in Germany.

What patterns are seen over time as new cultivars are released? Campos et al. (2004) examined a set of 18 maize hybrids released for the US corn belt between 1953 and 2001 (using three hybrids per decade). Mean performance of the hybrids was an increasing function of year of release. As expected, for any particular line, yield under drought stress was less than under controlled water conditions. However, more recently released hybrids gave greater responses under drought than did older releases. Indeed, the mean number of ears per plant under control conditions for lines released in the 1950's was similar to performance under drought for lines released in the 1990's (and beyond). Hence, selection for increased overall yield also seemed to improve performance under drought.

A more direct assessment was offered by Calderini and Slafer (1999) who compared Finlay-Wilkinson regressions for wheat lines from four different countries (and presumably different cropping systems) released over different eras. Figure 39.3 shows their results. In all four countries (Argentina, Australia, Italy, and the UK), more recently released lines gave greater yields (typically out-performing their earlier counterparts over all environments), but also showed increased slopes, indicating that more recently released lines are less stable than their earlier counterparts. Thus, more recently released lines were more sensitive to environment conditions (good and bad) than their older counterparts. This stability trend has yet to become an issue because the more recently released lines, despite being less stable, still out performed older lines over the range of environments tested. The general trend, however, is potentially cause for concern.

Location-based Modifications of the Finlay-Wilkinson Regression

Suppose genotypes are measured for three years over five different locations. We can treat this as 15 environmental values (for each year-location combination) and use the mean performance (over genotypes) for each combination as the 15 E values for the joint regression. However, it is often profitable to decompose the environment into a location effect, a year/season effect, and a location x year interaction (Chapter 38). The motivation is that location effects (the average value for that location over a series of years) can be a rather predictable feature, reflecting persistent environmental factors such as soil type and general climate. Conversely, year (or seasonal) effects are much less predictable, reflecting variation in weather over different years/seasons. With this decomposition, the environmental effect for year/season k in location l becomes

$$E_{lk} = \ell_l + y_k + \ell y_{kl} \tag{39.5a}$$

Our basic model (Equation 39.2a) then becomes

$$\mu_{ilk} = \mu + G_i + (\ell_l + y_k + \ell y_{kl}) + [\, G\ell_{il} + Gy_{ik} + G\ell y_{ikl}\,] \tag{39.5b}$$

where μ_{ikl} is the mean performance of genotype i in location l for year/season k. The three terms in parentheses are the environmental factors and the three terms in the brackets are the corresponding genotype x environment interactions. In the Finlay-Wilkinson regression, we approximated each GE_{ij} interaction term by $\beta_i E_j$, a genotypic sensitivity/response coefficient β_i and the environmental effect E_j. This approximation can be used with any of the individual interaction terms in Equation 39.5b. Given that location effects are expected to be rather predictable, a Finlay-Wilkinson regression is often used on the genotype x location interaction terms $G\ell_{il}$,

$$G\ell_{il} = \gamma_i \ell_l + \delta_{il} \tag{39.5c}$$

Equation 39.5b becomes

$$\mu_{ilk} = \mu + G_i + (1 + \gamma_i)\ell_l + \delta_{il} + (y_k + \ell y_{kl}) + [Gy_{ik} + G\ell y_{ikl}] \qquad (39.5d)$$

While this (initially) looks more complex that Equation 39.4a, we can think of this as the location-genotype interactions plus all of the other environmental interactions — the interactions associated with the predictable, and unpredictable, features of the environment, respectively. As we will see in the next section, this decomposition has been suggested as the basis for a more general measure of dynamic stability.

Finlay-Wilkinson Regressions and Types 1 Through 4 Stability

Up to this point, we have focused on two basic concepts of stability (static versus dynamic). Lin et al. (1986) and Lin and Binns (1988) suggest that we can further refine dynamic stability into additional classes, eventually leading to four types of stability. Lin et al. suggest that Becker's biological concept of stability/static stability (common value over all environments) be called **Type 1 stability**, while his agronomic concept (genotype-environment variance of zero) be called **Type 2 stability**. For example, a Finlay-Wilkinson slope near zero (i.e., β around -1) implies little response over environments and hence Type 1 stability. A slope near one (β around 0) implies response to environments, but no differential *linear* response over genotypes. If there is little G x E in this case (i.e., the residual variance about the regression is small), then we have Type 2 stability. Lin et al. (1986) suggested a third type of stability, also motivated from the Finlay-Wilkinson regression, based on the residual G x E that is not predictable. Thus **Type 3 stability** allows for G x E, *provided* it is predictable, which is generally expression in terms of the Finlay-Wilkinson regression.

This is accomplished as follows. Equation 39.4c breaks the genotype x environment interaction into a predictable component $\beta_i E_j$, which measures how each genotype linearly responds to the environment, and a residual component δ_{ij}, which is the amount of G x E not predicted from this simple approximation. Lin et al. (following Eberhart and Russell 1966) suggest that the variance of the residual component ($\sigma^2_{\delta_i}$ denoting the residual variance associated with the regression for line i) is yet another measure of stability, as it is the amount left over after the predictable amount of G x E has been removed. In this framework Type 3 stability is indicated by a small value of $\sigma^2_{\delta_i}$, the variance of the residuals in the joint regression for line i. In theory, one could use more general non-linear models to predict values of G x E (e.g., the split regressions of Verma and Chahal, or AMMI$_m$ models, to be discussed shortly), with the variance of residuals from these more general models also being a measure of Type 3 stability.

These various stability criteria result from considering every finer divisions of the variance in genotype means across environments. Type 1 considers $\sigma^2_E + \sigma^2_{G_i \times E}$, Type 2 uses $\sigma^2_{G_i \times E}$, and Type 3 the residual component $\sigma^2_{\delta_i}$ of $\sigma^2_{G_i \times E}$. Within the Finlay-Wilkinson framework, Type 3 stability expands the notion of dynamic stability to allow for *additive* changes in mean performance as a function of the environment, with different genotypes potentially having different slopes $1 + \beta_i$, so that the expected value of genotype i in environment j is given by $\mu_{ij} = \mu + G_i + (1 + \beta_i)E_j$.

Lin and Binns (1988) suggested a refinement based on a focus on only part of the environmental value, namely location effects, in the joint regression. Their motivation is that location contains predictable features, while the years within a location contains unpredictable features. Genotype x location values are computed by averaging over years, with the notion that this captures the predicable x genotype effects. Conversely, years within location are the genotype x unpredictable effect interactions. A breeder often wants stability in the latter component (years), while having the former (locations) respond to predictable features.

The Lin-Binns measure starts with a Finlay-Wilkinson regression based on location effects (Equation 39.5d), with the years within locations mean squares taken as their measure

of stability, which they call **Type 4 stability**. Again, this is in keeping with the idea of removing predictable components of interactions and trying to find genotypes that minimize the residual components. In this case, we treat locations as being largely predictable effects under the control of the breeder, while yearly variations within a location are not. Thus, under Type 4 stability, we seek to find those lines showing the largest response to particular locations while minimizing their year-to-year variation at those locations.

Univariate Measures of Stability

With several definitions of stability, it is not surprising that a number of stability metrics have been proposed, and we review a few of the more widely used measures here. For more detailed treatments, the reader is directed to reviews by Freeman (1973), Hill (1975), Haufe and Geidel (1978), Lin et al. (1986), Hühn (1987), Becker and Léon (1988), Crossa (1990), Pritts and Luby (1990), Fernandez (1991), Lin and Binns (1994), Kang and Guach (1996), Piepho (1996, 1998a), Kang (1998), Flores et al. (1998), and Yan and Kang (2003). Piepho (1996) discusses statistical tests for some of these measures. Our focus is on univariate measures of stability, as they have the appeal of providing potentially convenient targets of selection. At least in theory, one (or more) stability measures could be incorporated into a selection index along with performance. Multivariate measures, which are largely graphical and based on clustering and ordination, have also been proposed, and these will be discussed shortly. As emphasized by Piepho (1998a), stability measures critical depend on the underlying statistical model. When G x E is taken as a random effect in a mixed mode, elements of the covariance structure are intimately related to measures of stability, as we show later in the chapter.

The foundation for the stability measures presented in this section follow from the standard two-way (fixed-effects) ANOVA (Equation 39.2a) and Finlay-Wilkinson (Equation 39.4a) models of $G \times E$. Ironically, there is considerable instability in the literature on the notation for the different stability statistics presented below. We have tried to adapt a consensus view where possible, but the reader of the primary literature should be aware that notational variability is widespread. Finally, while the reader may view stability measures as requiring pure (genetically-identical, or nearly so) lines, these measures can also be applied to families or other segregating genetic groups (such as different open-pollinated lines). Obviously, we must have some way of replicating group members over different environments to access group stability, but these replicate elements need not be genetically identical.

• **The phenotypic variance of a genotype** over environments is a classic measure (e.g., Römer 1917) of static (i.e., Type 1) stability. Letting z_{ij} be the mean value of genotype i in environment j, under a balanced design ($n_{ij} = n_i$) this is estimated by

$$s_i^2 = \frac{1}{n_e - 1} \sum_j (z_{ij} - z_{i\cdot})^2 \tag{39.6a}$$

where $z_{i\cdot}$ is the mean value of genotype i over the n_e sampled environments. Assuming the environment E_j and interaction GE_{ij} terms are uncorrelated and ignoring the variance contributed by the residuals ϵ_{ijk} (which is expected to be small when replication is modest to large) this has expected value

$$E[s_i^2] = \sigma_E^2 + \sigma_{G_i \times E}^2$$

Francis and Kannenberg (1978) suggest that the coefficient of variation, $s_i/z_{i\cdot}$ is often a better measure, as the variance can increase with the mean.

• **Wricke's Ecovalence**. Wricke (1962, 1964) suggested a measure W_i he called **ecovalence**,

which is essentially the G x E sum of squares associated with a particular genotype,

$$W_i = \sum_j (z_{ij} - z_{i\cdot} - z_{\cdot j} + z_{\cdot\cdot})^2 \tag{39.6b}$$

Note that this measure assumes a balanced design (n_{ij} is constant), otherwise terms would be differentially weighted. Related ideas were suggested earlier (and less formally) by Salmon (1951), Horner and Frey (1957), and Plaisted and Peterson (1959). The smaller W_i, the more stable a genotype. Ecovalence is a Type 2 measure of dynamic stability.

• **Shulka's stability variance.** Shulka (1972) proposed a **stability variance** σ_i^2, the amount of genotype-environmental variance associated with genotype i. Note that this is equivalent to what we have been denoting as $\sigma^2_{G_i \times E}$, but here we use the less-descriptive notation σ_i^2 to conform to the literature. With n_g genotypes, under a balanced design this is estimated

$$Q_i = \frac{n_g(n_g-1)W_i - \sum_i W_i}{(n_g-1)(n_g-2)(n_e-1)} \tag{39.6c}$$

showing that the stability variance is a linear function of the ecovalance (Wricke and Weber 1980, Kang et al 1987, Piepho 1995b). As with W_i, smaller values of Q_i imply stronger Type 2 stability. Ignoring any contribution from the residual variance, its expected value is $E[Q_i] = \sigma^2_{G_i \times E}$. Note that any method that partitions the $G \times E$ terms into components for each genotype faces the problem that such components are not statistically independent (Freeman and Perkins 1971, Fripp and Caten 1971, Shukla 1972, Crossa 1990). This observation reenforces our comment that any analysis of $G \times E$ is context-specific. While Shukla's paper is widely cited, Piepho (1998a) notes that Calinski (1960) appears to have been the first use this approach in breeding.

• **Finlay-Wilkinson regression coefficient.** The estimated slope for a Finlay-Wilkinson regression (Equation 39.4a), either measured as the departure of actual slope $(1+\beta_i)$ from one, or the deviation β_i from zero, is a measure of the *responsiveness* of a genotype to changes in the environment. Let b_i denote the estimate of β_i. As noted above, a slope near zero implies Type 1 stability (little change over environments), while a slope near one implies Type 2 stability (little G x E, *provided* the residual variance $\sigma^2_{\delta_i}$ for the regression involving genotype i is small). There is a fair bit of criticism of this measure in the literature, in that if the coefficient of determination r_i^2 for the fit of the Finlay-Wilkinson regression for genotype i is small, then b_i is not a useful measure of stability, as the model explains little of the G x E (Eberhart and Russell 1966, Lin et al. 1986). Westcott (1986) went a step further, noting that regression approaches "cannot be regarded as trust-worthy" because they are too heavily leveraged (influenced) by a few extreme values.

• **Residual means squares from the Finlay-Wilkinson regression**. Eberhart and Russell (1966) proposed a Type 3 measure of stability based on $\sigma^2_{\delta_i}$, the variance of residuals about the regression, whose estimator is denoted by $s^2_{\delta_i}$. A very similar measure was proposed by Perkins and Jinks (1968), while Tai (1971) proposed a slightly different measure, but still motivated by the same approach. Breese (1969) was a strong champion of this approach, noting that one wishes to separate predicable response of the environment from "unpredictable irregularities in the response to an environment". This measure is also not without critics. Lin et al (1986) noted that this was simply a measure of model fit, not necessarily an indication of stability, and hence quite problematic.

• **Coefficient of determination for the Finlay-Wilkinson regression**, r_i^2. If r_i^2 is near one, then there is little residual variance $\sigma^2_{\delta_i}$. Thus, Type 3 stability is indicated by an r_i^2 value close to one. Although often attributed to Pinthus (1973), this measured does not appear to

have been formally proposed in his paper. The advantage of using r_i^2 over $\sigma_{\delta_i}^2$ is that it is a scaled measure and hence independent of the units of measurements. However, it is still a measure of model fit and hence suffers from the same concerns as $\sigma_{\delta_i}^2$.

• **Nonparametic rank-based measures**. Hühn (1979), Nassar and Hühn (1987), and Huehn (1990a,b) proposed two stability measures based on the distribution of the ranks of performance over environments. The first is a measure of absolute pair-wise difference in ranks over environments,

$$S_i^{(1)} = \frac{2}{n_e(n_e-1)} \sum_{j<k} | \, r_{ij}^* - r_{ik}^* \, | \tag{39.7a}$$

while the second is the variance in ranks over environments

$$S_i^{(4)} = \frac{1}{n_e - 1} \sum_{j} (\, r_{ij}^* - r_{i\cdot}^* \,)^2 \tag{39.7b}$$

Here r_{ij}^* denotes the rank (within environment j) of $z_{ij} - z_{i\cdot}$, the mean value of genotype i in environment j adjusted for the mean overall value of that genotype. Considering the deviation in environment j from the mean genotypic value extracts the stability information without being confounded by overall genotypic performance across all environments (Nassar and Hühn 1987; Becker and Léon 1988; Huehn 1990a,b). Without such a correction, stability and performance are confounded. For example, a genotype with the best performance in all environments (and hence uncorrected rank-based stability statistics of zero) may still show considerable variation. Likewise, a genotype with a constant performance over all environments might be scored as rather unstable using the ranks of uncorrected performance, as the performance of other genotypes may cause its rank to change over environments. The significant advantage of rank-based methods is that they are rather insensitive to outliers and measurement errors.

Table 39.1. Summary of common univariate measures of stability. Static stability is implied by Type 1 stability, while dynamic stability is implied by Type 2 and Type 3 stability. Here b_i is the estimated departure of the Finlay-Wilkinson regression slope (for genotype i) from one. With the exception of r_i^2, smaller values of these statistics imply stronger stability. For r_i^2, strong stability is implied by values near one.

Method	Statistic	Stability	Stability class
Cross-environmental variance	s_i^2	Static	Type 1
Ecovalence	W_i	Dynamic	Type 2
Shulka's stability variance	Q_i	Dynamic	Type 2
Finlay-Wilkinson Regression slope	b_i	Static	Type 1 ($b_i = -1$)
		Dynamic	Type 2 ($b_i = 0$)
Deviation from the FW regression	$s_{\delta_i}^2$	Dynamic	Type 3
FW Regression coefficient of determination	r_i^2	Dynamic	Type 3
Mean corrected rank difference	$S_i^{(1)}$	Dynamic	Type 2
Variance of corrected ranks	$S_i^{(4)}$	Dynamic	Type 2

Table 39.1 summarizes these various stability measures, and the aspect (static versus dynamic, Type 1 to 3) of stability they attempt to measure. While using different stability metrics often results in a different ranking of genotypes, there are some relationships between the different metrics. Wricke's Ecovalence and Shulka's stability variance produce

identical ranks (although their numerical values can be rather different). Analytic relationships among the parametric measures were examined by Schell (1967), Utz (1972), Becker and Léon (1988), and Peipho (1995b). For example (again ignoring the residual variance), the variance of genotypes over environments can be expressed as

$$E[s_i^2] = (1 + \beta_i)^2 \sigma_E^2 + \sigma_{\delta_i}^2 \tag{39.8a}$$

demonstrating how the regression slope (β_i) and regression residual variance ($\sigma_{\delta_i}^2$) are connected to the between-line variance. Likewise, with n_g genotypes, Shulka's stability variance has expected value

$$E[Q_i] = \left(\frac{n_g}{n_g - 2}\right) \beta_i^2 \sigma_E^2 + \sigma_{\delta_i}^2 + \text{constant} \tag{39.8b}$$

As these expressions highlight, the correlation among the different measures in a particular setting depends on the relative values of their various components. Since these vary, this becomes an emphirical issue, although a few general trends seem to be emerging. For a data set involving German trials of maize, barley, and wheat, Becker (1981) and Becker and Léon (1988) observed that b_i and s_i^2 tend to be highly correlated, as do $s_{\delta_i}^2$, W_i and r_i^2. Peipho (1995b) also observed these same correlations in simulated data. Huehn (1990a,b) notes that $S_i^{(1)}$ and $S_i^{(4)}$ are highly correlated with $s_{\delta_i}^2$ and W_i in a German winter wheat data set. Piepho notes that the between-environmental variance s_i^2 and the ecovalance/stability variance W_i (or Q_i) might be expected to show a high correlation, as $E[s_i^2] = \sigma_E^2 + \sigma_{G_i \times E}^2 = \sigma_E^2 + E[Q_i]$. Since σ_E^2 should be consistent over genotypes, the expectation is both measures should be completely correlated. However, this relationship does not hold if E_j and GE_{ij} are correlated, as is discussed in detail in Example 39.12.

Repeatability and Breeding for Stability

How does one attempt to breed for stability? Theoretically, one would first decide the appropriate stability criteria given the breeding objective and then apply one of the above measures, with the most favorable lines/families chosen on the basis of these values. With multiple traits, a selection index approach could be used to combine (and weight) information on the stability of different traits. In this framework, selecting for stability is no different than selecting on any metric trait. Let's examine this approach a bit more critically.

First, the meaning of stability is critical, and different breeding objectives may have different stability requirements. For example, for genetic resistance to pests and diseases, Type 1 stability is highly desirable (constant performance over all environments). With yield, however, we would like a genotype to respond to favorable environmental changes, and hence Type 2, 3, or 4 stability is more appropriate. Thus, different traits within the same breeding system may require different stability measures, and the appropriate weighting of the stability of different traits can be problematic. One approach would be to construct an economic model for the total value of a line given different amounts of stability in the target traits and this could be used to generate index selection weights. However, to generate an optimal (Smith-Hazel) index, one would need to estimate the genetic variances and covariances for the stability metrics over the target traits, something rarely done (Ortiz et al 2001).

Provided that one can assess, and weight, the appropriate stability metrics for the breeding objective, the next critical issue is the repeatability of these metrics. If low, selection response will be very ineffective. Is there evidence that these metrics are indeed heritable? We have already seen (Chapters 12, 37) that one can select for individuals that display lower environmental variance, measured as micro-environmental variation within a particular macro-environment. Surprisingly, while one might think that the literature has numerous studies on the correlation between micro- and macro-environmental stability, we have been

unable to find any. One might logically think that a genotype/line that is fairly robust to micro-environmental changes (and hence has a small micro-environmental variance), might likely be similarly buffered over larger changes in the environment. This remains an unproven assertion. If between-environmental changes are small, then one might suspect that such a correlation would hold. However, if there are significant differences between macro-environments, these may operate over a different scale than the variation seen within a particular environment, and biological features that provide for environment robustness on one scale may not work over another.

While there does appear to be at least some selectable micro-environmental variance, the issue is less clear for macro-environmental variance. Indeed, a number of studies have found that the repeatability of the common stability metrics is often very low (Fatunla and Frey 1976, Snoad and Arthur 1976, Hühn and Léon 1985, Weber and Wricke 1987, Léon and Becker 1988, Pham and Kang 1988, Pritts and Luby 1990, Léon and Becker 1998). For example, Weber and Wricke (1987) and Léon and Becker (1988) found that $s^2_{\delta_i}$ and W_i have a low repeatability, but (for the crop-environment set they examined) that the Finlay-Wilkinson slope b_i and the between-environment variance s^2_i often have good repeatability.

Example 39.1. Léon and Becker (1988) examined the year-to-year repeatability of various stability statistics using eight multi-year trails involving spring and winter wheat, spring barley, and oats in Germany from 1975 to 1985. In addition to yield (as a benchmark trait), two measures of response parameters (regression slope b_i and between-environmental variance s_i, which are often correlated) were considered. Four stability statistics were also examined, the residual variance $s^2_{\delta_i}$ from the FW regression, the coefficient of determination r^2_i of this regression, Wricke's Ecovalence W_i, and Hühn's variance of ranks statistics $S^{(4)}_i$. The average correlation between single year- and multiple-year results over eight experiments (each ranging from three to five years) involving four crops (winter wheat experiments and single experiments for spring wheat, spring barley and oats) were as follows:

	Yield	b_i	s^2_i	$s^2_{\delta_i}$	r^2_i	W_i	$S^{(4)}_i$
Mean correlation	0.80	0.56	0.54	0.45	0.37	0.44	0.37

While the year to year correlation in yield is high, the response parameters are less consistent from year to year, and stability measures even less so. Thus, using estimates from a single year will be very ineffective. A high year x location variance likely accounts for the observed low repeatability of the stability metrics.

A more direct assessment of genetics of stability parameters was offered by Lin and Binns (1991), who examined the stability values in a diallele analysis of smooth bromegrass genotypes in Canada. The stability estimate of each line was compared with that for the F_1 from their cross, and a standard test for additivity (e.g., LW Chapter 9) was performed. If there is some additive-genetic basis for these parameters, then the additivity test should be significant. They found that is was for the variance across environments s^2_i and for their Type 4 stability estimate (variance of years within location). However, neither W_i and $s^2_{\delta_i}$ showed a significant additive effect. The authors conclude (for this particular data set) that Type 1 and 4 stability estimates may have some additive variance, while Type 2 and 3 do not. Of course, if we are selecting among pure lines, it is the total genetic, as opposed to additive, variance that is critical. Nevertheless, this analysis suggests caution when trying to select for Type 2 and 3 stability in outbreeding populations. A second diallel study was performed by Ortiz et al (2001), who examined eight lines of bread wheat grown for several

seasons in two locations in Uganda. They did find significant, but small, heritabilities for the Finlay-Wilkinison regression slope ($h^2 = 0.15$), the residual variance $s^2_{\delta_i}$ from this regression ($h^2 = 0.18$), and the coefficient of determination r^2_i ($h^2 = 0.10$). As a point of reference, heritability for yield was 0.49.

Given that the heritability (in either the narrow or the broad sense) of a stability metric is often low, response to selection based on it is by no means guaranteed. This suggests than an alternative approach is to instead select on more heritable traits that are correlated with measures of stability. One such example was noted by Russell and Eberhart (1968). They observed that yield in maize lines with two ears was more stable than in single-ear lines. In stressful environments, single-eared lines were more likely to be *barren* (showing no ears, and hence no yield) than two-eared lines. Thus, features of two-eared genotypes appear to temper the effect of a stressful environment. In general, how might such traits be found? Plant and animal physiology models might suggest traits that can buffer environmental variation, as well as those traits that can respond to environmental changes (Edmeades et al. 2004, Yin et al. 2004). Clearly, this is an important area for future research. We return to this topic in Chapter 42 when we discuss selection on norms of reaction.

Joint Selection for Stability and Performance

Although one could select solely on stability using any of the above metrics, in real-world applications, selection on stability almost always incorporates some measure of performance. Again, one could place this in an index selection framework, choosing genetic groups on the basis of some index of a performance score plus a stability score. However, this is generally not done (outside of just using ranks, see below). The principal reason is the difficulty in obtaining appropriate weights for performance and stability. Instead, selection is usually based on some composite measure of both performance and selection, which are often generalizations of stability metrics. A number of such metrics appear in the literature:

• **Kang's Rank-sum Index**. Kang (1988) suggests that the problematic issue with an index (the appropriate scaling of its components) can be avoiding by using ranks in place of the mean performance and stability statistic values. In the simplest index, a genotype is assigned an index value that is the sum of its rank of its mean performance (best = rank one) plus the rank associated with some appropriate stability statistic (such as, but not limited to, W_i or $s^2_{\delta_i}$). Again the genotype/line with the best stability is assigned a rank of one. The best lines/genotypes are those with the smallest sum of ranks. Likewise, one could assign different weights to the two ranks. For example, the index $I = 2 * r_p + r_s$ assigns twice the weight to performance rank r_p as it does to stability rank r_s. Again, selection is for individuals with the smallest index values.

• **Varma-Chahal regression coefficient.** Building on Verma and Chahal's (1978) idea that an ideal genotype/line should be insensitive to poor environmental conditions while being responsive to good conditions (Figure 39.2), Pritts and Luby (1990) suggest basing selection on the difference of the Finlay-Wilkinson regression slopes estimated in the good ($b_{g,i}$) versus poor ($b_{p,i}$) environments,

$$b_{VC} = b_{g,i} - b_{p,i} \tag{39.9}$$

Ideally $b_{p,i}$ should be near zero (insensitivity to unfavorable environmental condition), while $b_{g,i}$ is much greater than one (response to favorable environmental conitions. Hence, selection is for larger values of b_{VC}. Note that the measure selects lines on the basis of their responsiveness, rather than performance.

• **Superiority measures**. Jensen (1976) and Lin and Binns (1988) suggested a simple superiority measure P_i as the basis for selection of the best line over a series of environments. If n_e

is the number of environments, then

$$P_i = \frac{1}{2n_e}\sum_j (z_{ij} - m_j)^2, \quad \text{where} \quad m_j = \max_i z_{ij} \tag{39.10a}$$

This measures the mean square distance between the line's response and the maximal response for that environment, averaged over all locations. Selection is for smaller values of P_i. Lin and Binns note the P_i must be used with care, as a genotype poor in general adaptability, but good in specific adaptation, can easily be discarded. Lin and Binns show that P_i can also be expressed as

$$P_i = \frac{1}{2}(z_{i\cdot} - m_\cdot)^2 + \frac{1}{2n_e}\sum_{j=1}^{n_e}(z_{ij} - z_{i\cdot} - m_j + m_\cdot)^2 \tag{39.10b}$$

This decomposition partitions P_i into average performance (first term) and stability (the second term). Note by comparison with Equation 39.6b that the second term is similar to Wricke's ecovalance, but with average performance over an environment replaced by the maximal performance in that environment (m_j in place of $z_{\cdot j}$) and the overall mean is replaced by the mean of the environment-specific maximums ($m_\cdot$ in place of $z_{\cdot\cdot}$). While Equation 39.10b appears to suggest that P_i might weight mean performance and stability roughly equally, in reality, P_i tends to be more strongly correlated with performance than stability (Léon 1986, Kang and Pham 1991), reflecting the relative magnitudes of the two terms in Equation 39.10b.

• **Nonparametic rank-based measures**. Using the *uncorrected* rank of a genotype with an environment, Hühn (1979; Huehn 1990a,b) suggested two measures of joint performance and stability:

$$S_i^{(3)} = \frac{1}{r_{i\cdot}}\sum_{j<k} |\, r_{ij} - r_{ik}\,| \tag{39.11a}$$

and

$$S_i^{(6)} = \frac{1}{r_{i\cdot}}\sum_j (\, r_{ij} - r_{i\cdot}\,)^2 \tag{39.11b}$$

For both measures, r_{ij} is the rank (in environment j) of the z_{ij} value. In contrast to the rank-based stability statistics $S_i^{(1)}$ and $S_i^{(4)}$, ranks used in Equation 39.11 are based on *unadjusted* values, as we are concerned with *both* mean performance and stability. Further, they are expressed in terms of units of the mean rank of the ith genotype. The numerator of both metrics is a measure of stability, while the denominator ($r_{i\cdot}$) is a measure of performance. While both $S_i^{(3)}$ and $S_i^{(6)}$ show significant correlations with the stability measures W_i and $s^2_{\delta i}$ on one hand, and mean performance on the other, $S_i^{(6)}$ tends to be more strongly correlated with performance and less with stability that $S_i^{(3)}$(Léon 1986, Becker and Léon 1988, Kahn and Pham 1991). Thus the choice of $S_i^{(6)}$ places more emphasis on mean performance, while $S_i^{(3)}$ places more weight on stability. Kang's index giving equal weight to the performance and stability rank tends to be intermediate to $S_i^{(6)}$ and $S_i^{(3)}$ in terms of weighting performance versus stability (Khan and Pham 1991). Additional nonparametric measures combining performance and stability have also been proposed, see Sabaghnia et al. (2006).

As with stability measures, the above metrics may rank genotypes rather differently. For example, Dashiell et al. (1994) examined P_i, $S_i^{(3)}$, and Kang's rank-sum index on a set of soybean lines grown in five locations in Nigeria. One particular line was quite desirable based on its $S_i^{(3)}$ and Kang index value, but quite undesirable given its P_i value.

Probability Criteria for Selection: Risk Aversion and Safety-First

As the above discussion suggests, the appropriate weighting of mean performance and stability is often unclear, in part because the breeding objective is often unclear. One very reasonable set of breeding objectives, especially under marginal conditions, are measures of food security. In such low production environments, the central breeding goal may be risk aversion — avoiding crop failures. Under a risk aversion framework, the selection criteria is to choose lines that minimize adverse events, and a number of such schemes have been proposed (Barah et al. 1981; Mead et al. 1986; Eskridge 1990, 1991; Eskridge and Mumm 1992).

One simple approach is to compute **risk probabilities**, the probability of performance being less than some pre-set value λ, the minimum yield for subsistence (being able to feed ones family) or economic well being. The goal is estimate $\Pr(z_{ij} < \lambda)$ for each line, with the most desirable line having the lowest risk probability (Mead et al. 1986). A common assumption is that the distribution of z_{ij} (over environments) follows a normal distribution, in which case

$$\Pr(z_{ij} < \lambda) = \Pr\left(\frac{z_{ij} - \mu_i}{\sigma_i} < \frac{\lambda - \mu_i}{\sigma_i}\right) = \Pr\left(U < \frac{\lambda - \mu_i}{\sigma_i}\right) \tag{39.12}$$

where U is a unit normal random variable. Lines with the lowest value of Equation 39.12 are favored. Note that this measure incorporates both mean performance μ_i and stability σ_i^2 (the variance for line i over environments). The assumption of normality can be problematic, but is not an unreasonable starting point in the absence of additional information. Alternatively, a more heavily-tailed distribution (such as a t random variable with low degrees of freedom) could be used, as these give more weight than a normal to extreme events.

Example 39.2. Consider three lines (A, B, C) whose means and variances over a random set of environments are given in the left half of the table below. The right half of the table gives values of the risk probabilities (in percentage) for these genotypes for two different values of λ (8, 10) computed using Equation 39.12.

Line	A	B	C	λ	A	B	C
μ_i	12	15	17	10	7.9	5.7	2.2
σ_i^2	2	10	12	8	0.2	1.3	0.5

For a subsistence/economic viability threshold of $\lambda = 10$, the ordering of lines (best = smallest probability to worst = largest probability) is C, B, A. In this case, the line with the highest variance (and highest mean) is favored. Conversely, for $\lambda = 8$, the ordering becomes A, C, B, so that the line with the smallest mean, but also smallest variance, is now favored. Given this dependence of λ, Peipho (1998a) recommends graphing performance over different λ values. One metric used in risk analysis is the notion of **stochastic dominance** (Bawa 1975). First-order stochastic dominance of genotype i over k implies $\Pr(z_{ij} < \lambda) < \Pr(z_{kj} < \lambda)$ for all λ and j of interest. Note that if both 8 and 10 are in our set of interest for λ that none of the above lines exhibit stochastic dominance over the others.

Risk probabilities are an example of **safety-first** models — provide for safety of the food supply by selecting lines with little chance of producing poor yield, even if this occurs at the expensive of average yield (Eskridge 1990, 1991). As an alternative to specifying a threshold λ, Eskridge suggested computing the confidence intervals for performance and

then selecting those lines with the largest values for their lower confidence intervals. Again, confidence intervals be computed assuming normality of yield with the variance following from the stability metric chosen (consult Eskridge 1990 for details). Using the hypothetical data from Example 39.2, with a 95% confidence interval, the ranking of the three lines (best to worst) are C, A, B (with lower confidence limits of 10.2, 9.2, 8.8), while for a 99% confidence interval the ranking becomes A, C, B (8.4, 8.1, 6.8).

SEEKING STRUCTURE IN G x E: AMMI, THE SVD, AND BIPLOTS

Selection of lines/families over multiple environments is a complex multivariate problem. When the genetic covariances (the same trait measured in different environments) are the same over all combinations, them approaches such as given by Equation 38.10 based on the mean of a line over all sampled environments can be used. Covariances, however, are usually expected to vary, with trait expressions in some environments being very similar, while being quite different in others. With estimates of these covariances in hand, the machinery of index selection can be used. In reality, its application is very limited, as the Smith-Hazel index is not particularly robust to poor parameter estimation (Chapter 33).

Thus, more robust approaches are needed for the two board goals a breeder is usually concerned with: defining sets of environments where genotypes show roughly similar performance and defining sets of genotypes that show roughly similar performance over different environments. Both of these are issue of looking for structure in the G x E patterns. One approach to accomplishing this goal is to use various clustering methods (Lin et al. 1986, Westcott 1986), but the more widely used approach is to use various statistical models to approximate the full table of G x E interactions, which we now consider in some detail.

Let's return to the base model of G x E,

$$z_{ij} = \mu + G_i + E_j + GE_{ij} + \overline{\epsilon}_{ij} \tag{39.13}$$

With n_g genotypes and n_e environments, the $n_g n_e$ genotype-environment interactions can be thought of an elements in an $n_g \times n_e$ matrix **GE**, with rows corresponding to different genotypes and columns to different environments (this can equally be defined as an $n_e \times n_g$ matrix were columns now corresponds to different genotypes). A important issue, to be addressed shortly, is which of the terms in Equation 39.2a are take as fixed and which are taken as random. Here, we start by taking all terms (except for the residual ϵ_{ijk}) as fixed, which forces us to consider only balanced designs (all pairwise combinations are genotypes and environment are present and each has roughly the same number of replicates).

In the completely **unstructured** model, each of the interaction terms in Equation 39.13 is individually estimated, and no relationships are assumed among them. In **structured** models, we try to account for at least part of the variation inherent in the **GE** matrix with far fewer than $n_e n_g$ terms. The Finlay-Wilkinson regression is an example of a structured model, wherein the interaction term is approximated by $GE_{ij} = \beta_i E_j + \delta_{ij}$ (Equation 39.4c), the product of a genotypic-specific environmental sensitivity β_i and the relative mean performance of the environment E_j, plus a residual error δ_{ij}.

Provided $\beta_i E_j$ accounts for most of the variation in GE_{ij} ($\sigma^2_{\delta_i}$ is small relative to $\sigma^2_{G_i \times E}$), then the interaction of genotype i in environment j is largely determined by an overall environmental sensitivity β_i for that genotype and the mean performance E_j of all lines within that environment. An important consequence of this model is that E and GE are no longer uncorrelated, as

$$\sigma(GE_{ij}, E_j) = \sigma(\beta_i E_j + \delta_{ij}, E_j) = \beta_i \sigma(E_j, E_j) = \beta_i \sigma^2_E \tag{39.14}$$

The natural extension of Equation 39.4c is to replace the simple product approximation with a sum of m such products, e.g.,

$$GE_{ij} = \sum_{k=1}^{m} f_{ik} h_{jk} + \delta_{ij} \tag{39.15a}$$

where δ_{ij} is the interaction residual and f_{ik} and h_{jk} are the contributions from the ith genotype and jth environment value associated with the kth term in the sum. Recalling Equations 39.1a and 39.2a , we can write $z_{ijk} = \mu_{ij} + \epsilon_{ijk}$, where

$$\mu_{ij} = \mu + G_i + E_j + \sum_{k=1}^{m} f_{ik} h_{jk} + \delta_{ij} \tag{39.15b}$$

The interaction (δ_{ij}) and model (ϵ_{ijk}) residuals are often lumped into a single residual. This can be done if one assumes that the variances for both residuals are either homoscedatic (independent of the index) or if one (or both) varies with genotype i (giving the lumped residual as $\sigma^2_{\delta_i} + \sigma^2_{e_i}$).

A number of approaches (detailed below) are based on the basic idea suggested by Equation 39.15a. The use of a multiplicative term (or terms) to approximate an interaction is called a bilinear or biadditive model, and we examine three such models here. The first is the **AMMI** model (**additive main effects, multiplicative interactions**), which just uses the performance data and does not require that any specific environmental variables (such as rainfall or temperature) are measured. The second are **factorial regressions**, which attempt to use measured environmental variables to account for the interactions. Finally, **reduced rank factorial regressions** attempt to use both performance data and measured environmental values, combining features of both AMMI and factorial regressions. We discuss these approaches in turn. Excellent introductions to these classes of models (as applied to plant breeding) are given by Copper and DeLacy (1994), van Eeuwijk (1995), van Eeuwijk et al. (1996), Vargas et al. (1999), and Crossa and Cornelius (2002). Before proceeding further, however, we need to introduce some additional matrix machinery.

The Singular-Value Decomposition (SVD)

An $n \times p$ matrix $\mathbf{A}$ can always be decomposed as the product of three matrices: an $n \times p$ diagonal matrix $\boldsymbol{\Lambda}$ and two unitary matrices, $\mathbf{U}$ which is $n \times n$ and $\mathbf{V}$ which is $p \times p$. The resulting **singular value decomposition** (**SVD**) of $\mathbf{A}$ is given by

$$\mathbf{A}_{n\times p} = \mathbf{U}_{n\times n} \boldsymbol{\Lambda}_{n\times p} \mathbf{V}^T_{p\times p} \tag{39.16a}$$

We have indicated the dimensionality of each matrix to allow the reader to verify that each matrix multiplication conforms. The diagonal elements $\lambda_1, \cdots, \lambda_s$ of $\boldsymbol{\Lambda}$ correspond to the **singular values** of $\mathbf{A}$ and are ordered by decreasing magnitude. Returning to the unitary matrices $\mathbf{U}$ and $\mathbf{V}$, we can write each as a row vector of column vectors,

$$\mathbf{U} = (\mathbf{u}_1, \cdots, \mathbf{u}_i, \cdots \mathbf{u}_n), \qquad \mathbf{V} = (\mathbf{v}_1, \cdots, \mathbf{v}_i, \cdots \mathbf{v}_p) \tag{39.16b}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are n and p-dimensional column vectors (often called the **left** and **right singular vectors**, respectively). Since both $\mathbf{U}$ and $\mathbf{V}$ are unitary, by definition (Appendix 4) each column vector has length one and are mutually orthogonal (i.e., if $i \neq j$, $\mathbf{u}_i \mathbf{u}_j^T = \mathbf{v}_i \mathbf{v}_j^T = 0$). Since $\boldsymbol{\Lambda}$ is diagonal, it immediately follows from matrix multiplication that we can write any element in $\mathbf{A}$ as

$$A_{ij} = \sum_{k=1}^{s} \lambda_k \, u_{ik} \, v_{kj} \tag{39.16c}$$

where λ_k is the kth singular value and $s \leq \min(p, n)$ is the number of non-zero singular values. Since only the first s diagonal elements of $\boldsymbol{\Lambda}$ are nonzero, the SVD can also be written as

$$\mathbf{A}_{n\times p} = \mathbf{U}_{n\times s}\boldsymbol{\Lambda}_{s\times s}\mathbf{V}^T_{s\times p}, \quad \text{with} \quad \mathbf{U} = (\mathbf{u}_1, \cdots, \mathbf{u}_s), \quad \mathbf{V} = (\mathbf{v}_1, \cdots, \mathbf{v}_s) \tag{39.16d}$$

This is the **compact SVD**. Since $\boldsymbol{\Lambda}$ is diagonal, we can express Equation 39.16c as

$$A_{ij} = \mathbf{u}_i\boldsymbol{\Lambda}\mathbf{v}^T_j \tag{39.16e}$$

A final useful identity is that the total variance of $\mathbf{A}$ (the sum of all its squared values) equals the sum of its squared singular values,

$$\sum_{ij} A^2_{ij} = \sum_{k=1}^{s} \lambda^2_k \tag{39.16f}$$

The importance of the singular value decomposition in the analysis of G x E arises from the **Eckart-Young theorem** (1938), which relates the best approximation of a matrix by some lower-rank (say k) matrix with the SVD. Define as our measure of goodness of fit between a matrix $\mathbf{A}$ and a lower rank approximation $\widehat{\mathbf{A}}$ as the sum of squared differences over all elements,

$$\sum_{ij}(A_{ij} - \hat{A}_{ij})^2$$

Eckart and Young show that the best fitting approximation $\widehat{\mathbf{A}}$ of rank $m < s$ is given from the first m terms of the singular value decomposition (the **rank-m SVD**),

$$\hat{A}_{ij} = \sum_{k=1}^{m} \lambda_k\, u_{ik}\, v_{kj} \tag{39.17a}$$

For example, the best rank-2 approximation for the G x E interaction is given by

$$GE_{ij} \simeq \lambda_1\, u_{i1}\, v_{j1} + \lambda_2\, u_{i2}\, v_{j2} \tag{39.17b}$$

where λ_i is the ith singular value of the **GE** matrix, $\mathbf{u}$ and $\mathbf{v}$ are the associated singular vectors (see Example 39.3). The fraction of total variation of a matrix accounted for by taking the first m terms in its SVD is

$$\sum_{k=1}^{m} \lambda^2_k / \sum_{ij} A^2_{ij} = \frac{\lambda^2_1 + \cdots + \lambda^2_m}{\lambda^2_1 + \cdots + \lambda^2_s}$$

Example 39.3. Consider the follow performances of soybean lines in different environments in New York from a larger dataset given by Gauch (1992). Each entry represents the average of four replications. For environments, the first letter denotes the testing location, while the two numbers denote the year of the test. For ease of presentation, we have rounded entry means and the associated row and column means.

	Genotypes				
Environment	EVAN	WILK	CHIP	row mean	$\widehat{E}$
A77	2725	2471	2333	2510	494
V79	1111	578	1278	989	−1027
R81	2038	1386	2350	1925	−91
I85	1736	1607	1588	1644	−372
G85	3258	2961	2813	3011	995
column mean	2174	1801	2072		
$\widehat{G}$	158	−215	56		

Denote the entry means by z_{ij}, the row (environment) and column (genotype) means by $z_{\cdot j}$ and $z_{i\cdot}$ and the grand mean by $z_{\cdot\cdot}$ (here 2016). Under a fixed-effects model (with homoscedastic residual variances), the estimated main effects are given by $\widehat{\mu} = z_{\cdot\cdot}$, $\widehat{E}_j = z_{\cdot j} - z_{\cdot\cdot}$, and $\widehat{G}_i = z_{i\cdot} - z_{\cdot\cdot}$, and these are given in the above table. The interaction terms are estimated by

$$\widehat{GE}_{ij} = z_{ij} - (\widehat{\mu} + \widehat{G}_i + \widehat{E}_j)$$

For example, for EVAN in A77, $\widehat{GE} = 2725 - (2016 + 158 + 494) = 57$. The resulting GE table of interactions becomes

	Genotypes		
Environment	EVAN	WILK	CHIP
A77	57	176	−233
V79	−36	−196	233
R81	−45	−324	369
I85	−66	178	−112
G85	89	165	−254

Now let's examine singular value decomposition (SVD) approximations for this table of GE values, which write in matrix form (using columns equaling genotypes) as

$$\mathbf{GE} = \begin{pmatrix} 57 & 176 & -233 \\ -36 & -196 & 233 \\ -45 & -324 & 369 \\ -66 & 178 & -112 \\ 89 & 165 & -254 \end{pmatrix}$$

In **R**, the compact SVD (Equation 39.16d) of a matrix X is given by **`svd(X)`**, returning the SVD of $\mathbf{GE}$ as

$$\begin{pmatrix} 0.40 & 0.21 & 0.18 \\ -0.41 & 0.00 & 0.91 \\ -0.66 & 0.12 & -0.30 \\ 0.26 & -0.83 & 0.11 \\ 0.41 & 0.50 & 0.19 \end{pmatrix}$$

$$\begin{pmatrix} 0.40 & 0.21 & 0.18 \\ -0.41 & 0.00 & 0.91 \\ -0.66 & 0.12 & -0.30 \\ 0.26 & -0.83 & 0.11 \\ 0.41 & 0.50 & 0.19 \end{pmatrix} \begin{pmatrix} 746.10 & 0 & 0 \\ 0 & 131.36 & 0 \\ 0 & 0 & 0.53 \end{pmatrix} \begin{pmatrix} 0.12 & 0.64 & -0.76 \\ 0.81 & -0.51 & -0.30 \\ 0.58 & 0.58 & 0.58 \end{pmatrix}$$

The first singular value accounts for $746.10^2/(743.26^2 + 131.36^2 + 0.53^2)$ = 97.0% of the total variation of $\mathbf{GE}$, while the second singular value accounts for 3.0%, so that together they account for essentially all of the total variation. The rank-1 SVD approximation of $\mathbf{GE}$ is given by setting all of the diagonal elements of $\boldsymbol{\Lambda}$ except the first entry to zero,

$$\mathbf{GE}_1 = \begin{pmatrix} 0.40 & 0.21 & 0.18 \\ -0.41 & 0.00 & 0.91 \\ -0.66 & 0.12 & -0.30 \\ 0.26 & -0.83 & 0.11 \\ 0.41 & 0.50 & 0.19 \end{pmatrix} \begin{pmatrix} 746.10 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0.12 & 0.64 & -0.76 \\ 0.81 & -0.51 & -0.30 \\ 0.58 & 0.58 & 0.58 \end{pmatrix}$$

Similarly, the rank-2 SVD is given by setting all but the first two singular values to zero,

$$\mathbf{GE}_2 = \begin{pmatrix} 0.40 & 0.21 & 0.18 \\ -0.41 & 0.00 & 0.91 \\ -0.66 & 0.12 & -0.30 \\ 0.26 & -0.83 & 0.11 \\ 0.41 & 0.50 & 0.19 \end{pmatrix} \begin{pmatrix} 746.10 & 0 & 0 \\ 0 & 131.36 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0.12 & 0.64 & -0.76 \\ 0.81 & -0.51 & -0.30 \\ 0.58 & 0.58 & 0.58 \end{pmatrix}$$

The resulting rank-1 and rank-2 approximations for $\mathbf{GE}$ are

$$\mathbf{GE}_1 = \begin{pmatrix} 34.57 & 189.94 & -224.67 \\ -35.78 & -196.59 & 232.54 \\ -57.51 & -316.03 & 373.81 \\ 22.33 & 122.71 & -145.15 \\ 36.04 & 198.02 & -234.23 \end{pmatrix}, \quad \mathbf{GE}_2 = \begin{pmatrix} 56.95 & 175.95 & -233.05 \\ -36.28 & -196.28 & 232.72 \\ -44.91 & -323.91 & 369.09 \\ -66.03 & 177.97 & -112.03 \\ 88.94 & 164.94 & -254.06 \end{pmatrix}$$

Since there are only three non-zero singular values, the rank-3 SVD returns the matrix $\mathbf{GE}$, i.e., $\mathbf{GE}_3 = \mathbf{GE}$. For example, consider the G x E term for EVAN and A77. The rank-1 SVD value for this entry is 3457, the rank-2 VSD value is 56.95, and the rank-3 (full rank) just the original value of 57.

It is often useful to consider Equation 39.16c as an inner-product of two vectors, with $A_{ij} = \mathbf{y}_i \mathbf{x}_j^T$. To do so, we first need to absorb the singular value λ_k over the original $\mathbf{u}_k$ and $\mathbf{v}_k$ vectors. Writing

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}^c \boldsymbol{\Lambda}^{1-c} \tag{39.18a}$$

where the ith diagonal element of $\boldsymbol{\Lambda}^c$ is simply λ_i^c, Equation 39.16e becomes

$$A_{ij} = \mathbf{u}_i \boldsymbol{\Lambda} \mathbf{v}_j^T = \mathbf{u}_i \boldsymbol{\Lambda}^c \boldsymbol{\Lambda}^{1-c} \mathbf{v}_j^T = \left(\mathbf{u}_i \boldsymbol{\Lambda}^c\right) \left(\boldsymbol{\Lambda}^{1-c} \mathbf{v}_j^T\right) = \mathbf{y}_i \mathbf{x}_j^T \tag{39.18b}$$

The new variables x and y are given by fixing a choice of c and then taking

$$y_{ik} = \lambda_k^c \, u_{ik}, \qquad x_{kj} = \lambda_k^{1-c} \, v_{kj} \tag{39.18c}$$

When we discuss biplots, different choices of c allow for different scalings and hence different visual representations of the data. As detailed below, three common choices are $c = 1$ (λ is directly absorbed into u), $c = 0.5$ (λ is even divided between u and v), and finally c can be chosen so that x and y have the same range of values (Equation 39.23).

Note that using the SVD to approximate a matrix is a least-squares approach, and as such is sensitive to outliers. Hawkins et al. (2001) proposed a more robust SVD procedure when outliers are of concern.

AMMI Models

The SVD provides the foundation for AMMI (additive main effects, multiplicative interactions) models. AMMI uses the first m terms of the SVD of the matrix of G x E interactions to approximate the GE_{ij} interactions (Good 1969, Gabriel 1978), with

$$GE_{ij} = \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij} \tag{39.19a}$$

giving

$$\mu_{ij} = \mu + G_i + E_j + \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij} \tag{39.19b}$$

The first axis (given by the η_{1j}) corresponds to a hypothetical environmental variable that describes the largest amount of the environment interaction (the **primary** effect), we then fit the best-fitting **secondary** effect, and so on. Similarly the γ_{1i} describe the axis of genetic sensitivities accounting for largest amount of genetic interactions, and so on.The fraction

of total variation accounted for by this approximation is the sum of the first m squared singular values divided by the total sum. The Finlay-Wilkinson regression can be thought of as a special case of an AMMI model, where a single term $GE_{ij} \simeq \lambda_1 \gamma_i \eta_j$ is considered and the environmental scores η_j are constrained to be the environmental main effect E_j. Gauch (1988) motivates the use of AMMI models by suggesting that the first few terms in the SVD for the G x E interaction matrix correspond to the true signal of any structure, while higher order terms correspond to noise due to sampling error. If the singular values are all fairly similar, there is little structure in the interactions, while if the first few singular values dominate the others, there is considerable signal (and hence structure) in the data.

Gauch (1988) popularized the use of AMMI models in plant breeding, although their development is due to a number of earlier workers (Fisher and MacKenzie 1923, Williams 1952, Gollob 1968, Mandel 1969, Perkins 1972, Gabriel 1978, Kempton 1984). One way of thinking about AMMI models is that the main effects (G_i and E_j) are removed, and then a process akin to principal components (PCs) is performed on the residuals (i.e., the GE_{ij} interaction terms).

Example 39.4. This example shows the connection between the SVD and the PCs of the $\mathbf{GE}$ matrix, following dos S. Dias and Krzanowski (2003). While $\mathbf{GE}$ is non-square, the related matrix $\mathbf{GE}(\mathbf{GE})^T$ is a square symmetric $n \times n$ matrix, and as such we can decompose it into its eigenvalues and eigenvectors, giving

$$\mathbf{GE}(\mathbf{GE})^T = \mathbf{U}_{PC} \boldsymbol{\Lambda}_{PC} \mathbf{U}_{PC}^T$$

where $\mathbf{U}_{PC}$ is a unitary matrix of the eigenvectors of $\mathbf{GE}(\mathbf{GE})^T$, while $\boldsymbol{\Lambda}_{PC}$ is a (square) diagonal matrix of the eigenvalues of this matrix. Now consider using the SVD,

$$\mathbf{GE}(\mathbf{GE})^T = (\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T)(\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T)^T = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T\mathbf{V}\boldsymbol{\Lambda}^T\mathbf{U}^T = \mathbf{U}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^T\mathbf{U}^T$$

The second step follows by taking the transpose through, as $(\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T)^T = \mathbf{V}\boldsymbol{\Lambda}^T\mathbf{U}^T$, while the third step follows from $\mathbf{V}^T\mathbf{V} = \mathbf{I}$, as $\mathbf{V}$ is a unitary matrix. Note that $\boldsymbol{\Lambda}\boldsymbol{\Lambda}^T$ is an $n \times n$ diagonal matrix whose elements correspond to the square of the diagonal elements of $\boldsymbol{\Lambda}$. Hence, we see that the squares of the singular values of $\mathbf{GE}$ correspond to the eigenvalues of $\mathbf{GE}(\mathbf{GE})^T$, while the eigenvectors of this matrix are given by the matrix $\mathbf{U} = (\mathbf{u}_1, \cdots, \mathbf{u}_n)$ from the SVD. Conversely, $(\mathbf{GE})^T\mathbf{GE}$ is an $p \times p$ square matrix, and the SVD gives

$$(\mathbf{GE})^T\mathbf{GE} = (\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T)^T(\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T) = \mathbf{V}\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}\mathbf{V}^T$$

The matrix $\mathbf{V} = (\mathbf{v}_1, \cdots, \mathbf{v}_p)$ from the SVD corresponds to the eigenvectors of $(\mathbf{GE})^T\mathbf{GE}$, while $\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}$ is an $p \times p$ diagonal matrix whose elements again correspond to the square of the diagonal elements of $\boldsymbol{\Lambda}$, and these again correspond to the eigenvalues of $(\mathbf{GE})^T\mathbf{GE}$ as well of $(\mathbf{GE})\mathbf{GE}^T$.

As noted by Weber and Westermann (1994), the connection between an AMMI and a principal components analysis is that we can write Equation 39.19a as

$$E(GE_{ij}) = \sum_{k=1}^{m} \sqrt{\lambda_k^{(PC)}}\, u_{ki}\, v_{kj}$$

where $\lambda_k^{(PC)}$ is the kth eigenvalue of both $(\mathbf{GE})^T\mathbf{GE}$ and $\mathbf{GE}(\mathbf{GE})^T$, while u_{ki} is the ith element in the kth (normalized) eigenvector of $\mathbf{GE}(\mathbf{GE})^T$ and v_{ki} is the ith element in the kth (normalized) eigenvector of $(\mathbf{GE})^T\mathbf{GE}$.

AMMI represents a *family* of models, with AMMI_k indicating an AMMI model with k terms in the sum given by Equation 39.19a (the notation AMMI-k also appears in the literature). AMMI_0 represents a model with only main effects (no interactions), while the full model, AMMI_F uses all of the singular values. In this case, Equation 39.19a recovers the individual cell means where all $\delta_{ij} = 0$ (e.g., Example 39.3). A number of tests have been proposed for the obvious question of how many terms should be included in Equation 39.19a (reviewed by dos S. Dias and Krzanowski 2003). Gauch and Zobel (1998) note that two different frameworks for testing can be used, often resulting in different models (number of terms m) for the same dataset. Under a **postdictive** approach, one adds additional terms to Equation 39.19a until their addition does not significantly improve the fit of the model to the full data set. This is typically done using F-like statistics, but this approach is very distribution-dependent and therefore not very robust (Cornelius 1993; Cornelius et al. 1996; Piepho 1994, 1995a). Under a **predictive** approach, cross-validation is used, wherein one subset of the data is used to fit the model, and the fit (predictability) of this model is then tested on the remaining data. While much more robust, the problem with this approach is that too much data is left out of the initial analysis (being saved to be used in the validation set). dos S. Dias and Krzanowski (2003) suggest two different "leave-one-out" validation methods that provide the robustness of cross-validation approaches while still using most of the data for estimation.

Modifications of the Basic AMMI Family of Models

A variety of modifications of Equation 39.19 appear in the literature. Two common variations are the **sites regression model**, or **SREG** (Crossa and Cornelius 1997) wherein the genetic main effect (G_i) is absorbed into the interaction terms and hence the regression main effects are only over sites (E_j),

$$\mu_{ij} = \mu + E_j + \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij} \tag{39.20a}$$

and the **shifted multiplicative model**, or **SHMM**, Crossa and Cornelius 1997) where *both* the environment and genetic main effects are absorbed into the interaction terms,

$$\mu_{ij} = \mu + \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij} \tag{39.20b}$$

As with AMMI, these models are usually subscripted to indicate the number of multiplicative terms included, so that SREG_3 is Equation 39.20a with $m = 3$ terms. As we will shortly see, these variants can prove useful for joint visualization of both genetic main effects plus GE interactions. Other variants have also been proposed, again based on which terms are kept as main effects versus being absorbed into a general interaction term, see Cornelius and Crossa (1999) for details.

Finally, AMMI-type models can be selectively applied to partitions of the interaction terms. We previously applied the Finlay-Wilkinson regression approximation to the genotype x location terms, while treating the year and year by location terms as noise (Equation 39.3). This approach can be expanded by using AMMI approximations for genotype x location terms (Annicchiarico and Perenzin 1994; Annicchiarico 1997a,b).

Using AMMI to Predict Cell Means

In the terminology of experimental design, a particular genotype- environment combination is often called a **cell**, and an important issue is to estimate both the cell mean and the

interaction term associated with that cell. Assuming a fixed-effects model under a balanced design and homoscedastic residual variance, the least-squares (LS) estimate (often called BLUE, for Best Linear Unbiased Estimator) for the mean, genotypic, environmental, and interaction effects are

$$\text{BLUE}(\mu) = z_{..}, \quad \text{BLUE}(G_i) = z_{i.} - z_{..}, \quad \text{BLUE}(E_j) = z_{.j} - z_{..},$$
$$\text{BLUE}(GE_{ij}) = z_{ij} - \overline{z}_{.j} - \overline{z}_{i.} + \overline{z}_{..} \tag{39.21a}$$

Thus, the least-squares estimate of the mean of genotype i in environment j is

$$\text{BLUE}(\mu_{ij}) = \text{BLUE}(\mu + G_i + E_j + GE_{ij}) = z_{ij} \tag{39.21b}$$

which is simply the sample mean z_{ij} (see Equation 39.30b for a derivation). While an obvious estimator, it is also based on a small sample size (the number of replicates in a cell) relative to all of the data in the entire design. If the data has some structure (i.e., there are correlations across certain sets of cells), then by borrowing information from correlated cells we may possibly obtain a better estimate of the true mean of a cell than is given by using its sample mean. The same holds true for the interaction effect associated with that cell.

Gauch (1988) proposed that the LS estimates of μ_{ij} and GE_{ij} can be inferior to estimates using an AMMI model with a truncated number of terms. His motivation is as follows. The LS estimates given by Equations 39.21a/b is equivalent to using a SVD that retains all terms ($m = s$). However, as we have noted, the higher-order terms (those with small singular values) often represent noise while the first few terms may contain much more signal. Thus, by appropriate truncation to remove noisy terms, we can improve on the LS estimator by what amounts to borrowing information from other cells to separate the signal over the entire G x E array from the noise inherent in the sample mean from any particular cell. In this framework, if m is the number of terms retained in the AMMI analysis, then the estimate of the interaction term is given by

$$\text{AMMI}_m(GE_{ij}) = \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} \tag{39.21c}$$

while the AMMI estimate of the cell mean is given by

$$\text{AMMI}_m(\mu_{ij}) = \text{BLUE}(\mu) + \text{BLUE}(G_i) + \text{BLUE}(E_j) + \text{AMMI}_m(GE_{ij}) \tag{39.21d}$$

where the BLUE estimates are given by Equation 39.21a.

Example 39.5. Consider the data presented in Example 39.3. The observed mean value for genotype EVAN in environment I85 was $z_{ij} = 1736$. This is also the BLUE estimate for this genotype-environment combination. As given in this example, $\widehat{\mu} = 2016, \widehat{G}_{EVAN} = 158, \widehat{E}_{I85} = -372$. Thus, 1802 is the BLUE estimate of $\mu + G_{EVAN} + E_{I85}$. Likewise from Example 39.3, the AMMI_1 approximation of $GE_{EVAN,I85}$ is 22.33, while the AMMI_2 approximation is -66.03. Hence, the AMMI_1 estimate of this cell mean is 1802+22.33 = 1824.33, while the AMMI_2 estimate is 1802-66.03 = 1735.97.

Simulations by Cornelius (1993) showed that using AMMI indeed does result in a better improved (over least-squares) estimators of both cell means and interaction effects. Cornelius noted while the truncated AMMI approach does well, the singular values tend to be

overestimated, which in turn overestimates the interaction terms. This prompted Cornelius and Crossa (1999) to suggest a **shrinkage estimator**,

$$\text{SNK}(GE_{ij}) = \sum_{k=1}^{s} S_k \lambda_k \, \gamma_{ki} \, \eta_{kj} \tag{39.22}$$

where the kth shrinkage factor S_k lies within (0,1). One key feature of this estimator is that *all* singular values are included in the sum, so the issue of where to truncate the series is replaced by the issue of how to weight each contribution. Equation 39.22 simply replaces the estimated singular values λ_k with adjusted (shrunken) singular values $S_k\lambda_k$. All terms are kept, but higher order terms are increasingly down-weighted. The motivation for this (and other) shrinkage estimators is that one suspects that the observed mean for the largest variety likely overestimates its true mean, just as the observed mean for the smallest variety likely underestimates its true mean. Shrinkage is akin to regressing both back towards the mean. The use of a truncated series (the first m terms retained) can be thought of an special case of a shrinkage estimator, where the shrinkage factor is one for the retained terms and zero for the excluded terms. Different shrinkage factors have been proposed by Cornelius and Crossa (1999) and by Moreno-González et al. (2003a,b; 2004), who can be consulted for details. We will also return to shrinkage when we discuss BLUP estimation at the end of the chapter.

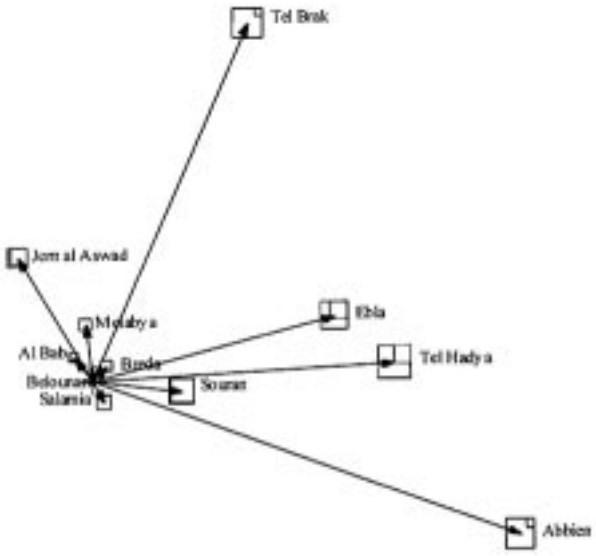


Figure 39.4. A phenotypic variance-covariance plot of Barley yield data from a series of locations in Syria. Each vector corresponds to the phenotypic variance for a particular environmental location. The length of the vector is the standard deviation for that environment, with the angle θ between two vectors a measure of their correlation, as $\rho = \cos(\theta)$. Note that yields in Abbien and Tel Brak show similar levels of variation, but are essentially uncorrelated (as the angle between their vectors is around 90 degrees). Conversely, Ebla and Tel Hadya are highly (positively) correlated, while both are negatively correlated (angle greater than 90 degrees) with Jern al Aswad. After van Eeuwijk et al. (2001).

Visualization of Interactions: Biplots

One advantage of AMMI (and other bilinear approximations of the interaction terms) is that they can reduce a complex multi-dimensional problem (the distribution of elements in the

GE matrix over genotypes and environments) into a simpler representation that may capture much of the variation in far fewer dimensions. In order to visualize the approximations of these interactions, the key is to represent objects of interest (such as genotypes or environments) as vectors, which we can then compare by using angles, lengths, and projections (Appendix 4). Our overview of this subject is somewhat brief, and additional reviews can be found in Kroonenberg (1995), Gower and Hand (1996), Yan and Kang (2003), and Yan and Tinker (2006). Laffont et al. (2007) discuss other graphic tools, such as dot plots, that can further facilitate the analysis of a biplot.

To introduce the basic ideas, consider Figure 39.4, which is a plot of the phenotypic variances (i.e., the variance across line means) for a series of environments. More variable environments result in longer vectors. The breeder is interested in such environments, as they provide potentially greater discrimination among genotypes (as the variation across mean genotypic performance is larger in these environments). Correlations between environments is indicated by the relative directions of the vectors, with the correlation ρ between two vectors a function of the angle θ between them, $\rho = \cos(\theta)$ (Appendix 4). What is immediately apparent is that the brain can fairly easily process the relationships among the elments in this high-dimensional date set by looking at the plot. For example, the testing location Souran is not very discriminating, but the closely related locations (at least for this particular set of genotypes) Tel Hadya and Abbien are, so that these might be used in place of Souran.

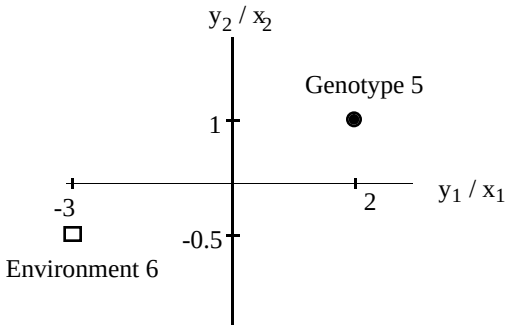


Figure 39.5. The building-blocks of a biplot: Plotting values for genotypes and environments. The two-term SVD generates a pair of coordinates (y_1, y_2) for each genotype and likewise a pair of coordinates (x_1, x_2) for each environment. Since the biplot displays two sets of coordinate systems (x_1, x_2 and y_1, y_2) simultaneously, an important issue is choosing the scaling constant c such that the y and x axes scale well together (Equations 39.18c, 39.23). Suppose the first two terms in the SVD for the sensitivity of genotype 5 are $(y_{15}, y_{25}) = (2, 1)$. These are plotted above with the filled circle. Likewise, suppose the SVD returns the coordinates for environment 6 as $(x_{16}, x_{26}) = (-3, -0.5)$. The rest of the genotypes and environments are plotted in a similar fashion.

Similar comparisons can made between the interactions of genotypes and environments by looking at **biplots** (Gabriel, 1971), a term arising because we plot both row (genotype) and column (environmental) data from the **GE** matrix on a single graph. The idea is to take the first two terms of an AMMI (or other bilinear) model and use these as the coordinates for genotype i and environment j. Figure 39.5 shows the basic idea. Writing

$$GE_{ij} \simeq \sum_{k=1}^{2} \lambda_k \, \gamma_{ki} \, \eta_{kj} = y_{1i} x_{1j} + y_{2i} x_{2j}$$

where we have used some scaling scheme (choice of c, see Equations 39.18c and 39.23)

to absorb the singular value, creating genotypic (y_{ki}) and environment (x_{kj}) scales. The approximation sign indicates we are predicting the full interaction with only the first two terms of the SVD. If these account for most of the variation (i.e., $\lambda_1^2 + \lambda_2^2$ is a very large fraction of the sum of all squared singular values, such as occured in Example 39.3), then this approximation is very good. We then plot each of the n_g genotypes on a graph using their y_1 and y_2 values from the SVD, e.g., the point for genotype 5 is given by (y_{15}, y_{25}). Similarly all n_e environments (typically locations or location-years) are plotted using their values on the x_1 and x_2 axes, so that environment 3 is given by (x_{13}, x_{23}). The accuracy of this 2-D plot as a representation of the whole G x E space depends on how well the first two axes account for the total variation. If the fit is less than desired, this same approach can be used by generating plots based on the first three axes (and using computer visualization software to rotate this space to get a better feel for it), or one could simply construct three biplots: axis 1 vs. axis 2, axis 1 vs. axis 3, and axis 2 vs. axis 3. Care must be taken when using biplots in that if the first few axis do not account for the vast majority of variation then the results they suggest can be very misleading. Biplots are a useful tool, but only if use properly.

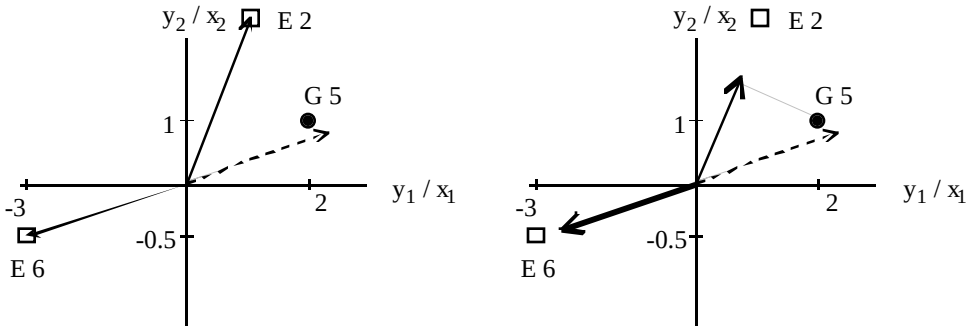


Figure 39.6. The building-blocks of a biplot: Projections determine strength and sign of interactions. The coordinates for genotype 5 and environment 6 are as in Figure 39.5. An additional environment (E2) is also plotted with $(x_{13}, x_{23}) = (1, 2)$. Using an AMMI$_2$ model, the predicted G5 x E6 interaction is $x_{15} \cdot y_{16} + x_{25} \cdot y_{26} = -3 \cdot 2 - 0.5 \cdot 1 = -6.5$, while the predicted G5 x E3 interaction is $1 \cdot 2 + 2 \cdot 1 = 4$. Thus G5 underperforms in E6 and overperforms in E2, with predicted mean values of $\mu + G_5 + E_6 - 6.5$ and $\mu + G_5 + E_3 + 4$, respectively. With a little practice, one can quickly visualize the strength and sign of interactions by considering the projection of the $\mathbf{y}$ vector (of coordinates for G) onto the $\mathbf{x}$ vector (the coordinates of E). **Left.** The first step is to consider the vectors to E2 and E6. Since the environment vector for E6 points in the opposite direction of G5, we extend it by reflecting it through the origin (Appendix 4), as shown by the dashed line. **Right.** The magnitude of the predicted interaction is given by the length of the projection times the length of the vector being projected upon. The projection of the G5 vector onto both E2 and E6 (thick arrows) indicates a modest positive interaction in E2 and a larger negative interaction on E6, the sign being negative because the projection is on a reflection about the origin (i.e. $\cos(\theta) < 0$).

Once we have plotted the genotypic and environment vectors, the real power of a biplot arises in that the interaction between genotype i and environment j is a function of the projection of the genotypic vector for i onto the environmental vector for j (Figure 39.6). The length of the resulting projection is proportional to the magnitude of the interaction and the angle between the genotypic and environmental vectors indicates its sign (positive if $-90 < \theta < 90$, negative otherwise). This directly follows as the 2-term SVD decomposition estimate of the interaction term can be written as the inner product of the genotype ($\mathbf{y}$) and

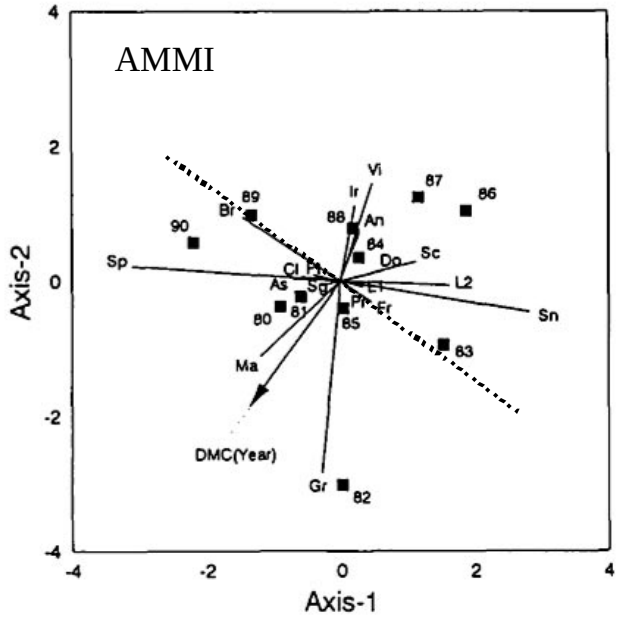
environment (**x**) coordinate vectors,

$$GE_{ij} \simeq y_{1i}x_{1j} + y_{2i}x_{2j} = \mathbf{y}_i\mathbf{x}_j^T = ||\mathbf{y}_i||\,||\mathbf{x}_j||\cos(\theta)$$

The last step follows from Equation A4.2, which relates the inner product of two vectors to their lengths and the angle θ between them (for a review of inner products and vector projection see Appendix 4). Likewise, Equation A4.3b shows that the absolute value of this expression is the lenght of the projection of $\mathbf{y}_i$ onto $\mathbf{x}_j$ times the lenght of $\mathbf{x}_j$.

Hence, if two genotypes have vectors with similar orientation (the angle θ between them is small), they will have similar types of interactions over environments, as (up to their length differences), the resulting projections onto any environmental vector will also be similar. The same argument holds when considering environmental vectors. Note that Equation 32.25 emphasizes two key features for predicting interactions: the lengths of the two vectors and the angle θ between then. Thus, if a genotype or environment has a vector of very short length, it will tend to have very little interactions over environments (or genotypes). Conversely, just because a genotype and an environment both have long vectors, this does not imply they will have a large interaction. If the vectors are nearly orthogonal, there is essentially no interaction (as $\cos(\theta) \simeq 0$). Likewise, if there is a large interaction, its sign is a function of the angle between the vectors. Example 39.6 highlights some of these points.

Example 39.6. van Eeuwiki et al. (1995) applied AMMI to examine G x E in the Dutch maize variety trails. These involved 18 maize varieties evaluated over a 10 year period at four different locations in the Netherlands. While location had a large impact over all varieties (i.e., the location main effects were highly significant), there was little variety-by-location interaction. By contrast, there was significant variety-by-year interaction. Hence, the analysis of G x E focused on the interaction between varieties and years. The biplot for the first two axis of the AMMI analysis is given below, with lines used to represent genotypes (given by two-letter codes), while squares indicate years (e.g., 80 = 1980, etc.).



The arrow ending in DMC (dry matter content) shows the direction of greatest change to

increase DMC. The dashed line perpendicular to this vector separates above vs. below average years for DMC yield, with 82 being the highest performing and 86 and 87 the lowest performing years, while 89 and 83 were close to average. Points near the origin correspond to lack of interactions, so that varieties and environments (years) at the greatest distance from the origin are those that show the most interaction (e.g., varieties Sp, Sn, and Gr; years 82, 83, 86, 87, 89, 90.) Conversely, points near the origin show little G x E. Hence, varieties such as Sg, Pr and Fr (all near the origin) have relatively stable performance over environments (i.e., $\mu + G_i + E_j$ is a good predictor of their expected yield). Projection of the variety vector onto the environmental vector indicates the strength (and direction) of any G x E. Hence, Gr and 82 had a large positive interaction (so that the yield of Gr in 82 significantly exceeded $\mu + G_{Gr} + E_{82}$), while both Sp and Sn are nearly orthogonal to 82, and hence showed very little interaction. This is especially interesting, because Sp, Sn, and 82 all in general show large interactions. Conversely in year 90, Sp had a large positive interaction, Sn a large negative interaction, and Gr a very small interaction. These various comparisons show the power of the biplot, as the investigator can simply look at a set of years or genotypes of interest and quickly assess the pattern of interactions between them. As an example for the reader to try, 87 was a bad year, but which of the varieties Vi, Sn, Sp, and Gr overperformed, underperformed, or showed no significant interactions?

The answer is that Vi overperforms (having a small angle with year 87), Sn and Sp show little interaction (being at close to right angles to 87), and Gr underperforms (having an angle in excess of 90 degrees with 87).

A final comment on scaling, i.e., the choice of c in Equation 39.18c used to generate the y and x coordinate axes from the SVD. It is immediately apparent from Equation 39.18a that the value of the inner product of the genotypic $\mathbf{y}_i$ and environmental $\mathbf{x}_j$ vectors is independent of the scaling value c chosen. The role of c is simply to help the viewer in extracting information displayed in a biplot. If all genotypic values are bunched up around the origin, while the environmental values show a much greater length divergence, it can be hard to visualize projections (and hence interactions) for different genotypes. One way to facilitate visualization is to choose the scaling to ensure that the range of the coordinate vectors for the environments is the same as for the genotypes. Consider the scaling on the primary axis first, with the SVD returning values of λ_1, γ_{1i} for $1 \leq i \leq n_e$, and η_{1j} for $1 \leq j \leq n_g$. Letting

$$r_\gamma = \max_i(\gamma_{1i}) - \min_i(\gamma_{1i}), \quad \text{and} \quad r_\eta = \max_j(\eta_{1j}) - \min_j(\eta_{1j}),$$

Yan et al. (2001) and Crossa et al. (2002) found that the scaling which gives y_1 and x_1 the same range is

$$y_{1,i} = \lambda_1^{c_1}\gamma_{1i}, \quad x_{1j} = \lambda_1^{1-c_1}\eta_{1j}, \quad \text{where} \quad c_1 = \frac{1}{2}\left(1 + \frac{\ln(r_\gamma/r_\eta)}{\ln(\lambda_1)}\right) \tag{39.23}$$

Similar scalings for both secondary axes follows by using the observed range of the γ_{2i} and η_{2j}.

GGE/GGL Biplots

Just as an appropriate scaling of the genotypic and environmental axes can facilitate the interpretation of a complex data set, so can appropriate **centering**. The plots based on the first two axis of the AMMI approximation of the G x E effects are often called **GE biplots**, as they are centered by removing the main effects ($\mu + G_i + E_j$) before plotting the interaction terms.

Hence, the origin on GE biplots corresponds to no interaction, and the closer a genotypic or environmental vector is to the origin, the smaller its interaction effect. While such a visualization is extremely useful, a GE biplot does not tell the breeder their real interest: what is μ_{ij}, the expected value of genotype i in environment j? A large positive G x E in a particular environment is *not* sufficient to ensure that the genotype has the best performance in that environment. To find the best genotype for a particular environment using a GE biplot, one would compute the inner product of the $\mathbf{y}_i$ and $\mathbf{x}_j$ coordinate vectors and then add the estimated main effect G_i in order to predict the deviation of this genotype from the environmental average.

A modification of the basic AMMI model generates a **GGE biplot** (also a **GGL biplot**, with L for location) where now the genotypic axis contains information on both the main effect of a genotype and its propensity for interactions with the environments measured. The inner product of such a vector with an environmental vector returns the expected yield produced by that genotype, expressed as a deviation from the mean yield for that environment. This is done by first absorbing the genotypic main effects into the interaction terms, such as occurs with the SREG or SHMM variants of AMMI discussed above (Equations 39.20a/b). A common approach is the **environment-centered** model based on SREG,

$$\mu_{ij} - (\mu + E_j) = \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij} \tag{39.24}$$

Here environmental effects are subtracted off so that the origin in the GGE biplot corresponds to the average yield for an environment. A longer genotypic vector indicates the potential for a significant deviation from the average yield, encompassing both the main effect of that genotype as well as its propensity for interacting with the environment.

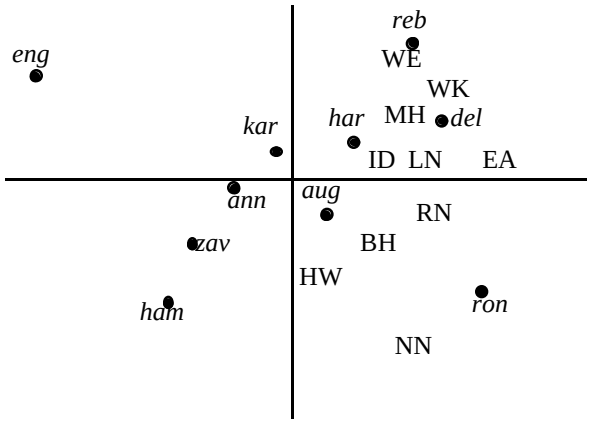


Figure 39.7. A GGL biplot for winter wheat yield in Ontario (for 1992). Cultivars are given by the small dots with lowercase letters, while location (environments) are labeled by a pair of capital letters. The primary axis accounts for 57% of the total variation, the secondary for 20%. After Yan et al. (2000).

There are numerous advantages to using GGE biplots. The first is that the investigator can quickly find the optimal genotypes for each environment by again looking at the projections of the genotypic vector onto the environmental vector. Under a GE biplot, the longest such resulting inner product would be the genotype with the largest G x E interaction in a particular environment. Under a GGE biplot, the resulting inner product now includes

both the main effect plus the specific interaction, and hence the longest projection vector corresponds to the best-performing genotype in that environment. For example, in Figure 39.7, *reb* is the best-performing genotype for location WE, while *ron* is the best-performer for location NN.

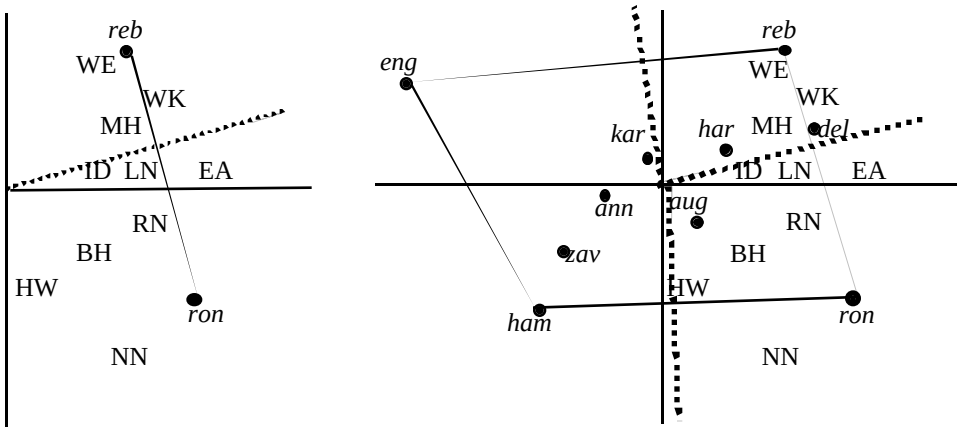


Figure 39.8. Further GGL analysis of Yan et al.'s winter wheat data from Figure 39.7. **Left:** For two genotypes (say *reb* and *ron*), to determine who wins where, draw a line between then (solid line in the figure) and then construct a line (the dotted line in the figure) that both runs through the origin and is perpendicular to the initial line. Environments on the side of the line with a genotype are where it outperforms its counterpart. Hence *reb* is better than *ron* in locations WE, WK, and MH. In all others, *ron* is better. **Right:** The GGL plot can also indicate mega-environments among the locations. The most extreme performing genotypes in the entire sample are the corner (or vertex) genotypes, which in this case are *reb*, *ron*, *ham*, and *eng*. One can then take adjacent vertex pairs, and construct the regions where each genotype performs best (sectors separated by dotted lines). Such regions correspond to mega-environments and the corner genotype in each such sector is the best-performing genotype for that particular mega-environment (*reb* in WE, WK, MH, *ron* everwhere else). After Yan et al. (2000).

Again using the notion that performance is based on the (scaled) projection of the genotypic vector into the environment vector, GGE biplots allows one to quickly compare the performance of two genotypes over all the environments to see **who wins where**. As shown in Figure 39.8, this is done by first connecting the two genotypes with a line, and then constructing a second line that runs through the origin and is perpendicular to the connecting line. Environments on the same side of the perpendicular line as the genotype are where it out-performs the other. For example, *reb* outperforms *ron* in environments WE, WK, and MH (Figure 39.7). The logical basis for this simple method again follows from the yield being the projection onto a vector. Since the environmental vector is the same for both genotypes, the winner is the one with the largest projection. Simple geometry indicates this occurs when the environment falls above the perpendicular line.

Figure 39.8 also highlights another important feature, namely separation of the locations/environments into similar clusters (or mega-environments) and determining the best-performing genotype for such a cluster. Inspection of the GGE biplot gives the **corner** (or **vertex**) genotypes with the most extreme values. These are connected by lines to form a polygon. One then takes adjacent pairs and uses the previous method to construct the regions where each genotype wins (i.e., who wins where). The resulting sectors correspond to

the mega-environments and the corner genotype is the best performing of all in that set of environments.

Yan et al. (2001) and Yan and Rajcan (2002) stress an additional insight provided by GGE biplots. *Provided* that the genotypic main effects variance is much larger than the G x E variance, the primary axis (often called PC1) of the GGE plot is largely based on the genotypic effects, and hence is average yield, while the secondary axis (PC2) is a measure of sensitivity/tolerance, with small (absolute) values showing high tolerance (lower sensitivity) and hence more (Type 2) stability across environments. Thus, when PC1 is very highly correlated with the main effects, the best cultivars for broad adaptability are those that show a large PC1 value and a small PC2 value. Likewise, if PC1 is very highly correlated with genotypic main effects, then the ideal testing environments are also indicated on the GGE biplot, namely those that also have a large PC1, and small PC2, score. The large PC1 score implies that the environment is more discriminatory in terms of separating genotypes based on their main effects, while a small PC2 score implies that it is more representative of the overall environment.

Yan et al. (2001) note that this strong correlation between genotypic PC1 score and main effects breaks down when there is a complex G x E structure. In this case, they proposed a modification of the environmental-centered SREG model that forces the first PC to have a perfect correlation with genotypic main effects. They do so by using Mandel's (1961) model for including a single multiplicative term for interactions, by taking $G_i\alpha_j$ as the first term, where G_i is the genotypic main effect and α_j is a measure of environmental variability. Note that this is similar to Finlay-Wilkinson except that the genotypic main effect is included in place of the environmental main effect (i.e., it is based on G_i rather than E_j). The second term is then fitted using an environmental-centered SREG model, giving

$$\mu_{ij} - (\mu + E_j) = G_i\alpha_j + y_{2i}\, x_{2j} \tag{39.25}$$

Yan et al. refer to Equation 39.25 as the SREG_{M+1} model, for an environmental-centered SREG model with the first multiplicative term given from Mandel's approximation (hence the M subscript) plus an additional term from SREG. Yan et al. found that while this does not account for as much variation as a standard SREG_2 model, they felt that the ability to quickly find the high-yield, high stability lines and ideal test environments is often worth the cost of accounting for slightly less variation.

An additional feature that is often plotted on a GGE biplot is the **average-environment axis** (or **AEA**), given by the vector whose two components correspond to the average values of the environments, $(\sum_j x_{1j}/n_e, \sum_j x_{2j}/n_e)$. This vector corresponds to the average environment over the sample. If desired, it cab also be computed seperately for each mega-environment. Environments close to this vector are the most representative of the average environment, and hence a refinement of the above approach for finding the most representive environment is to replace a small PC2 score with a small distance from the AEA. Likewise, one measure of the stability of a genotype is how close it lies to this axis. Based on this, Yan and Kang (2003) and Fan et al. (2007) suggest the idea of selecting on **GGE distance**, the distance from an ideal genotype (taken as the largest observed value with complete stability – a hypothetical point obtained by taking the largest observed PC1 score for any line in the sample and placing this distance on the average-environment axis). Of course, if the target is for locally-adapted lines, the approach shown in Figure 39.8 to finding the best genotype for a given mega-environment can be used.

GGE vs. AMMI Biplots

The GGE approach for finding mega-environments and who-wins-where is not the only biplot-based scheme to do so, nor was it the first. Gauch and Zobel (1997) recommended

a plot with the genotypic main effect on one axis and the value for $AMMI_1$ on the other. Recently, a fair amount of ink and angst has been devoted to contrasting the AMMI and GGE approaches (Gauch 2006, Yan et al. 2007, Gauch et al. 2008), and whether the reader finds these discussions endearing or tedious has a high G x E component. The basic argument for AMMI-based plots is that these two axes (main effect + first AMMI component) explain more of the G + G x E variation than do the first two axes of a GGE biplot. The counter-argument is that GGE gives a "true" biplot in that the inner product rule holds, with the projection between two vector indicating the nature and amount of their interactions. This is not the case for main effects by $AMMI_1$ plots, making conclusions using GGE plots much easier to visualize. In reality, both approaches offer somewhat different information, and a careful breeder would do well to employ both to generate a more holistic picture of their data.

A major caveat to *both* approaches is that a biplot is a visual tool, not a rigorous statistical test (Yang et al. 2009). If the majority of the variation is *not* accounted for by the first two singular values for the model being considered, then a very misleading picture can arise. As more axes (i.e., singular values) are added to a model, typically more mega-environments that emerge. Likewise, even when the first two axis account for a very significant fraction of the total variation, the coordinates on a biplot are point estimates and do not express the uncertainty in their actual values. While Denis and Gower (1996, 1999) develop large-sample confidence intervals for coefficients from biadditive models, and show how these can be used to enhance biplot representations, these are based on asymptotic normality assumptions. This can be especially problematic when relatively small sample sizes are used. Yang et al. (2009) suggest bootstrapping approaches. While these can be more robust, recall from our discussion in Chapter 32 that the resampling unit in a bootstrap analysis is critical, and given the potential (and unknown) correlation structure in a sample of genotypes and environments, exactly how such units should be chosen in unclear. The point to be stressed is that biplots are a very useful tool, but should be used with caution.

SEEKING STRUCTURE IN G x E: INCORPORATING ENVIRONMENTAL FACTORS

AMMI models (which includes Finlay-Wilkinson as a special case) simply use the average performance of all genotypes to provide the measure of any given environment. In some sense, this is the most natural measure, as the interest is in the impact of the environment on the trait. Hence, AMMI does not require the investigator to measure any environmental variable (such as mean temperature or total rainfall). An alternative, or supplementary, approach to treating G x E is to incorporate measured environmental factors into the analysis and determine which genotypes (if any) show interactions with any of the measured environmental values. This is not a new approach (Abou-El-Fittouh 1969, Perkins 1972, Fripp 1972, Hardwick and Wood 1972, Wood 1976), but it saw a resurgence with the development and use of bilinear regressions (e.g., Baril 1992, van Eeuwijk 1992, van Eeuwijk and Elgersma 1993). Good introductions to the statistical machinery and its application to G x E problems are given by van Eeuwijk (1995), van Eeuwijk et al. (1996), and Vargas et al. (1999).

Factorial Regressions

While AMMI models attempt to extract information about how G x E interactions are related across sets of genotypes and environments, **factorial regressions** incorporate *direct* measures of environmental factors in an attempt to account for the observed pattern of G x E. Obviously, the power of this approach is that if we can determine which genotypes are more (or less) sensitive to which environmental features, the breeder may be able to more finely tailor a line to a particular environment without necessarily requiring trials in the target environment.

The basic approach is as follows. Suppose m features are measured in each environment, with x_{kj} corresponding to the value of environmental feature k in environment j. One can then consider the G x E interaction terms arising from genotypes differentially interacting with the set of environmental features. In this case the interaction term is modeled by

$$GE_{ij} = \sum_{k=1}^{m} \zeta_{ki}\, x_{kj} + \delta_{ij} \tag{39.26}$$

where x_{kj} is the observed value of environmental feature k in environment j, ζ_{ki} the sensitivity of genotype i on environmental feature k, and δ the model residual. The Finlay-Wilkinson regression is a special case of a factorial regression using only a single environment variable, the observed mean performance of genotypes in that environment, $x_j = E_j$.

While the equations appear nearly identical, the difference between AMMI models and factorial regression is that under an AMMI model, we need not measure any environmental variables, with the SVD decomposition of the interactions array returning the environmental vectors. Hence, the environmental factors η_{kj} in an AMMI are *estimated* from the SVD. By contrast, with factorial regression, we *directly measure* environmental values x_{kj} of interest. While this may potentially provide much more biological insight, it also usually accounts for a smaller fraction of the G x E variance (which is not surprising because AMMI is designed to maximally capture this).

Example 39.7. Baril et al. (1995) used factorial regression to examine 64 varieties of potatoes over 26 environments (classified by soil type and year), in an experiment running 16 years. Soil type was classified as either sandy or clay, and four other environmental variables were included: number of frost days in the first (and second) half of April, mean temperate over the April to September growing season, and total solar radiation over the growing season. Each of these five environmental variables accounted for between 5 and 8% of the G x E interaction sums of squares, together accounting for roughly 1/3 of all of the interaction variance. Factorial regression models can also include genotypic covariates. Baril et al. initially examined three such covariates: classification of lines according to starch type and early maturity (both on a 1-10 scale), and leaf development (on a 1 to 9 scale), which was not significant and discarded from future analysis. The resulting factorial regression model becomes

$$\mu_{ij} - \mu - G_i - E_j = GE_{ij} = \sum_{k=1}^{5} a_{ki} x_{kj} + \sum_{k=1}^{2} y_{ki} b_{kj} + \delta_{ij}$$

Here, the x_{kj} correspond to the observed values of the five environmental variables (soil type and the four climatic summary statistics), with a_{ki} measuring the sensitivity of genotype i to environmental factor k. The trick to keeping track of all the terms in such equations is to remember that terms with an i subscript involve features associated with line/genotype i, while terms with a j subscript involve features of environment j. Likewise, y_{ki} corresponds to the value for the kth genotypic covariate for genotype i and b_{kj} are the environmental weightings for the kth genotypic covariate in environment j. Note that the x and y are directly observed while a and b are estimated. Interaction terms involving early maturity accounted for 11% of the interaction variance and starch accounted for an addition 3%, for a total of 45% of the interaction variance (when combined with the five environmental variables). As a reference, an AMMI_5 model accounted for 58% of the interaction variance, but at a cost of more degrees of freedom. When compared at comparable degrees of freedom (i.e., a plot of total percent of interaction variance accounted for versus percentage of degrees of freedom used), AMMI and factor-analytic models showed comparable performance, with factorial regression being slightly better at a low percent of the total degrees of the freedom and AMMI better at a higher

percentage. In a similar study comparing AMMI and factorial regression for yield in winter wheat, Brancourt-Hulmel and Lecomte (2003) found that AMMI and factorial regressions explained similar amounts of variation (77.4% vs. 74.0%), but that factorial regression used fewer degrees of freedom and was more biologically informative, providing a description of the genotypic sensitivities with respect to the observed environmental variables.

Example 39.8. Epinat-Le Signor et al. (2001) offer an interesting example showing the potential power for factorial regressions in providing biological insights. They examined grain yield from 132 maize hybrids in 30 locations over a 12 year period (given the unbalancedness of the data, this resulted in a total of 229 environments). Two genetic covariates (date of flowering and date of maturity of each line) and four environmental covariates together accounted for around 40% of the interaction variance, with the genetic and environmental covariates roughly accounting for equal amounts. A major contributor was the interaction between a line's date of flowering and water supply, with early varieties becoming more favorable as the water supply decreases. A second contributor was between late maturing varieties and solar radiation around time of flowering, with the interaction effect increasing (becoming more favorable) for early maturing varieties as solar radiation increases. It is certainly not unexpected that different varieties have different reactions to the environment, and factorial regression in this case has indicated favorable interactions between particular genetic subgroups and particular environmental factors.

Reduced Rank Factorial Regressions

The idea behind reduced-rank factorial regressions is straightforward: instead of working with the m measured environmental variables x_{kj}, we attempt to work with a new (and hopefully smaller) set of environmental factors y_{hj} that correspond to the linear combinations of the original variables,

$$y_{kj} = \sum_{h=1}^{m} f_{hj} x_{hj} \tag{39.27a}$$

The resulting GE term in the regression is modeled by

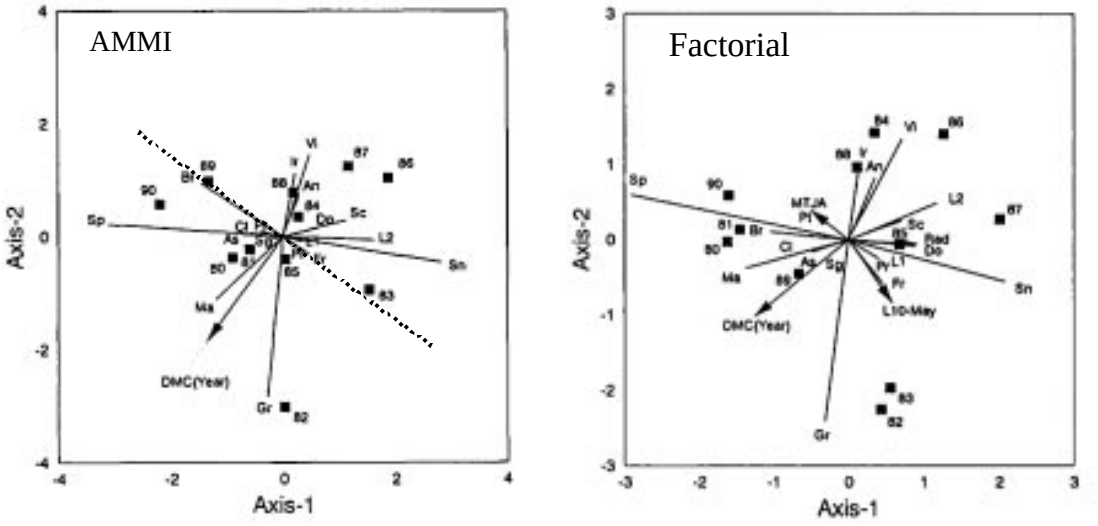
$$GE_{ij} = \sum_{k=1}^{p} \zeta_{ki}\, y_{kj} + \delta_{ij} = \sum_{k=1}^{p} \zeta_{ki} \left(\sum_{h=1}^{m} f_{hj} x_{hj} \right) + \delta_{ij} \tag{39.27b}$$

The motivation for this approach is that instead of (say) six environmental variables being required to account for the data, there might be just (say) two combinations of these original variables that account for most of the environmental variation underlying G x E. Finally, one can combine both AMMI and reduced rank factorial regressions by replacing the environmental axes η_{kj} in an AMMI analysis with the axes of environmental factors y_{kj}.

As Table 39.2 highlights, these various bilinear models that attempt to provide for a lower-dimensional representation of the **GE** interactions matrix appear quite similar. This arises because the motivation of all approaches is to estimate the interaction terms with a small sum of products. The difference between models is in the nature of the environmental variables used. Under an AMMI analysis, one uses values in the estimated **GE** matrix, with the environmental axis following from its singular value decomposition. With a factorial regression approach, we instead use direct measures of the environment (such as soil type and/or temperature) to look for those specific environmental factors that show significant genotypic sensitivity. Finally, in reduced-rank factorial regressions, a linear combination of

the environmental variables is used as the environmental terms in an AMMI model. If the majority of G x E interactions are functions of the measured environmental variables, them AMMI and reduced rank factorial regression will show very similar patterns. If this is not the case, these solutions will differ, suggesting that important environmental variable(s) impacting G x E have not been included in the analysis. As Example 39.9 shows, a biplot analysis is a useful way to compare the results of the two approaches.

Example 39.9. Returning to the Dutch maize variety trails examined in Example 39.7, van Eeuwiki et al. (1995) used reduced rank factorial regression to examine variety-by-year interactions. The biplot for the first two axes of this analysis is given below (right) along with the AMMI biplot (left) presented in Example 39.7. In addition to DMC (Example 39.7), four other measured environmental variables were considered, three of which have their vectors plotted on the biplot for the factorial regression analysis: L10-May (May days above 10^o C), Rad (total radiation over the growing season), and MTJA (mean temperate over July-August). The first axis separates L10-May and Rad (positive values) from MTJA and DMC (negative values).



Reduced rank factorial regression attempts to reconstruct the G x E interactions using just the measured environmental variables. Thus, while the variety vectors essentially have the same geometry in both plots, there are some noticeable differences in the location of the years between the two plots. For example, 86 and 87 are much more separated under the reduced-rank analysis (right), while 82 and 83 are much closer. The lack of concordance between these two plots indicates that not all relevant environmental variables have been considered by the factorial regression. This is especially important in that 82 was an exceptional year in terms of yield, yet there is little separation between 82 and 83 under the reduced-rank analysis. Likewise, 86 and 87 were roughly equal poor years, yet the reduced rank analysis (right plot) separates them by a larger amount than their actual performance dictates (left plot). Hence, just using the above environmental parameters to predict how good a year would not have been very successful for this data set.

The important message is that all of these tools provide breeders with additional information for selection decisions. Factorial regressions can be used to target those genotypes

with the smallest (for broad adaptation) or greatest (for locally-superior variety) sensitivity to important environmental factors in the target population of environments. AMMI analysis can be used to see how environments, and genotypes, cluster in the G x E space. For example, how similar are environments in different years from the same location? A biplot can suggest which genotype is likely to be best performing in a set of such environments.

Table 39.2. Summary of some of the bilinear models used to approximate the full space of G x E interactions in a lower dimension. Quantities subscripted with i refer to effects associated with genotype (or line) i, while those with j refer to effects associated with environment, or environmental factor, j. Here x denotes the observed value of a variable, while Greek letters denote quantities to be measured.

Model	Interpretation
Finlay-Wilkinson $GE_{ij} = \beta_i(E_j - \mu) + \delta_{ij}$	$\beta_i =$ sensitivity of genotype i to the average effect E_j of the environment.
AMMI $GE_{ij} = \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij}$	First m terms of the SVD of the $\mathbf{GE}$ matrix λ_k^2 is the amount of variation explained by axis k $\gamma_{ki} =$ sensitivity of genotype i to environmental axis k $\eta_{kj} =$ value of environment j on the kth environmental axis
Factorial Regression $GE_{ij} = \sum_{k=1}^{m} \zeta_{ki} \, x_{kj} + \delta_{ij}$	Modeling G x E using m measured environmental factors $x_{kj} =$ value of kth environmental factor in environment j $\zeta_{ki} =$ sensitivity of genotype i to kth environmental factor
Reduced rank Factorial Regression $GE_{ij} = \sum_{k=1}^{m} \zeta_{ki} \, (\sum_p c_{kp} \, x_{pj}) + \delta_{ij}$	Modeling G x E based on a reduced dimensional set of the observed environmental factors by constructing m combinations (axes) of these effects. $c_{kp} =$ loading of pth environmental factor on axis k. $\zeta_{ki} =$ sensitivity of genotype i to kth environmental combination (axis)
AMMI using Reduced rank Factorial Regression $GE_{ij} = \sum_{k}^{m} \lambda_k \gamma_{ki} (\sum_p c_{kp} x_{pj}) + \delta_{ij}$	The environmental axes η_{kj} under AMMI are replaced by the environmental axes generated by linear combinations of measured environmental factors generated by a reduced rank factorial regression, with $\eta_{kj} = \sum_p c_{kp} x_{pj}$.

MIXED-MODEL ANALYSIS OF G x E

Recall that a linear model, such as that given by Equation 39.13, can be analyzed as a fixed effects model (all effects except for ϵ are fixed), a random effects model (all effects except for μ are random), or a mixed model (some effects are fixed, others are random). The distinction between fixed and random effects can be very subtle, and the choice of which framework to use is influenced by both biological and statistical considerations. As we will see, even in settings where one can argue that a fixed-effects interpretation may be more "natural", there can be statistical reasons (such as less biased estimates and more efficient use of all the data) for treating certain effects as random.

Thus far in this chapter, we have been using a fixed-effects framework, assuming that G_i, E_j and GE_{ij} are fixed (but unknown) constants to estimate, while the residual error ϵ is a random variable (e.g., Equation 39.1). We further assumed that the variance of ϵ is

independent of both i and j (the genotype and the environment), so that

$$\sigma^2(\,\overline{\epsilon}_{ij}) = \frac{\sigma_e^2}{n_{ij}} \tag{39.28}$$

where n_{ij} is the number of replicates for this combination. In a fixed-effects framework, we have shown how this basic model can be expanded, for example by using AMMI or factorial regressions to replace the GE_{ij} terms. Since block and spatial effects create correlations among the residuals, the basic model often has additional terms attempting to account for these (e.g., Smith et al. 2001, 2002a,b; Qiao et al. 2004; Casanoves et al. 2005b).

By contrast, the models examined in Chapter 38 of the expected response to selection using the mean of a line or family over a sample of environments are based on random effects models, with all terms (except for the mean μ) being random. This is a reasonable model where the line/family members are randomly assigned to environments and we do not keep track of the environmental assignments. The rather simple forms for the selection response equations (38.10, 38.11) result from the very simplified covariance structure (compound symmetry) assumed for the vectors of random effects. Covariance structures are pivotal in the analysis of random and mixed models and much of our following discussion is on the development of more biologically realistic covariance models. Under a mixed model, we obtain BLUE estimates of the fixed effects and BLUP (best linear unbiased predictors) for the random effects. To compute the later, we need to know the variance components associated with the covariance structure. These are typically estimated using REML (restricted maximum likelihood). BLUP/REML machinery is covered in LW Chapters 26, 27 and Chapters 15, 16, and 35.

Our focus now becomes mixed-model analysis of G x E interactions. Besides proving a more natural framework in some cases, there are also very practical reasons for treating certain factors as random and moving from the strictly fixed-effects model to a mixed model. Under a fixed-effect model, missing values (combinations of genotypes and environments that are missing) cannot generally be accommodated. The result is that only a subset of the data can be analyzed, which can lead to significance estimation bias (Example 39.10). Under a mixed-effect framework, missing data is easily accommodated by exploiting the covariance structure of the random effects.

Example 39.10. Piepho and Möhring (2006) note that German field trails for new wheat cultivars typically use 15 locations in the first and second years, moving to 32 locations in year three. Usually the initial trails start with 100 cultivars, only half of which are passed onto year two, and only half of the year two lines are passed onto year three. The result is a highly unbalanced dataset. Piepho and Möhring used a simulation study to show that if only the completely-tested lines are used (those that survive all three years), the results are rather biased estimates of their true mean values. However, if all of the data (i.e., inclusion of the lines excluded in years one and two) are included under a mixed-model analysis, then BLUP/REML allows for unbiased estimates. This is in keeping with the theme in Chapters 15 and 35 that a REML analysis can yield unbiased variance component estimates under selection when all of the data are included. With these unbiased estimates in hand, empirical BLUP (BLUP using REML estimates of variances) can then to used to predict line values.

A second important reason for a mixed-model analysis is the notion of shrinkage that we have previously introduced. Estimating the true mean of the highest performing cultivar by its observed sample mean almost certainly results in an overestimation, and likewise using

the sample mean for the lowest-performing cultivar underestimates its true mean. Under a mixed-model framework, BLUP shrinks (or regresses) estimates back towards the mean, reducing this bias (Chapter 35). The importance of shrinkage is illustrated by Patterson and Silvey (1980), who found that the sample means of lines chosen by trials in the UK overestimate their true means by an average of around 27%. BLUPs (which are shrunken versions of BLUEs) shrink the estimates back towards their means, producing less biased estimators (Smith et al. 2001).

Several authors regard this shrinkage feature as significantly important enough to suggest that where the goal is predicting future performance, genotypes should always be treated as random effects (Hill and Rosenberger 1985; Panter and Allen 1995; Smith, Cullis and Gilmour 2001; Smith et al. 2005), even if the concern is only among a fixed collection of lines. The other suggestion that commonly appears in the literature is to treat genotypes as random during the early phases of selection and as fixed as a final few genotypes are chosen for evaluation. Our feelings is that there is considerable merit in treating genotypes as random, although we also present mixed models where they are treated as fixed.

We now examine several of the mixed models that start from Equation 39.13 by first declaring some of the effects to be random and then specify their particular covariance structure. The simple model of either E or G random with a compound symmetry covariance structure (all variances equal, all covariances equal) is examined first. We then build upon this by considering more realistic covariance structures, many of which are motivated by bilinear approximations of the interaction terms. Our goal here is to provide a general sketch of these methods while largely ignoring some of the minutia (the bookkeeping, while often tedious, can be quite important). More detailed reviews of mixed model applications to estimate G x E are given by Piepho (1997a,b, 1998b), Smith et al. (2001, 2005), Balzarini (2002), Crossa and Cornelius (2002), and Piepho et al. (2008).

BLUP Estimates Under Compound Symmetry Assumptions

The key element in moving from a fixed effects to a mixed model is the *specification of the covariance structure* of the random effects. Assume the genetic effects G in Equation 39.13 are random, and hence so are the G x E interactions, while E is treated as fixed. The simplest model of covariance structure of the random effects in this case is to assume

$$G_i \sim N(0, \sigma^2_G), \qquad GE_{ij} \sim N(0, \sigma^2_{GE}), \qquad \epsilon_{ijk} \sim N(0, \sigma^2_e) \tag{39.29a}$$

with G, GE, and ϵ all assumed to be independent of each other. Equation 39.29a assumes that all of the variances of homoscedastic, namely they do not vary over genotypes or environments (they are independent of any subscripts on the random variable). Clearly, we will relax this assumption shortly. Let's examine the implications of these assumptions for the covariance structure of this model. The expected genetic variance within a specific environment is just

$$\sigma(z_{ijk}, z_{ij\ell}) = \sigma(G_i + GE_{ij}, G_i + GE_{ij}) = \sigma(G_i, G_i) + \sigma(GE_{ij}, GE_{ij}) = \sigma^2_G + \sigma^2_{GE} \tag{39.29b}$$

where fixed effects (μ, E_j) do not enter. Likewise, the covariance of the same genotype (i) between two different environments (j and k) becomes

$$\sigma(z_{ij}, z_{ik}) = \sigma(G_i + GE_{ij}, G_i + GE_{ik}) = \sigma(G_i, G_i) = \sigma^2_G \tag{39.29c}$$

This is the **compound symmetry model** (Chapter 38), as the genetic variance within a given environment is constant (Equation 39.29b), as is the genetic covariance between any two environments (Equation 39.29b). Thus, the genetic correlation between any two environments

is also constant, with $\rho_G = \sigma_G^2/(\sigma_G^2 + \sigma_{GE}^2)$. The simple expressions in the previous chapter (38.10, 38.11) for the response based on the mean of a line or a family replicated over a number of environments follow by assuming this constant covariance structure. Specifically, genetic covariance matrix $\mathbf{G}$ in the multivariate breeder's equation in this case is given by

$$\mathbf{G} = \sigma_G^2\,\mathbf{J} + \sigma_{GE}^2\,\mathbf{I} \tag{39.29d}$$

where $\mathbf{J}$ is a matrix of ones.

Now let's contrast the estimate of cell means and interaction effects under this mixed model with those under a strictly fixed-effects model with a balanced design and uncorrelated homoscedastic residuals. For notational ease, we let BLUE$(X) = \widehat{X}$. As given by Equation 39.21a, the least-squares estimates for the fixed factors are

$$\widehat{\mu} = \overline{z}_{..}, \quad \widehat{E_j} = \overline{z}_{.j} - \overline{z}_{..}, \quad \widehat{G_i} = \overline{z}_{i.} - \overline{z}_{..}, \quad \widehat{GE}_{ij} = z_{ij} - \widehat{E_j} - \widehat{G_i} - \widehat{\mu} \tag{39.30a}$$

A little algebra allows us to express z_{ij} as

$$\begin{aligned} z_{ij} &= \overline{z}_{..} + (\overline{z}_{.j} - \overline{z}_{..}) + (\overline{z}_{i.} - \overline{z}_{..}) + (z_{ij} - \overline{z}_{.j} - \overline{z}_{i.} + \overline{z}_{..}) \\ &= \overline{z}_{..} + (\overline{z}_{.j} - \overline{z}_{..}) + (\overline{z}_{i.} - \overline{z}_{..}) + (z_{ij} - [\overline{z}_{.j} - \overline{z}_{..}] - [\overline{z}_{i.} - \overline{z}_{..}] - \overline{z}_{..}) \end{aligned} \tag{39.30b}$$

which nicely ties up the estimators offered by Equation 39.30a, giving the predicted mean as

$$\widehat{\mu}_{ij} = \widehat{\mu} + \widehat{E}_j + \widehat{G}_i + \widehat{GE}_{ij} = z_{ij} \tag{39.30c}$$

Thus, as we saw in Equation 39.21b, the LS estimate of the cell mean is the sum of the least squares estimates for each of the components, which in this case (balanced design with uncorrelated homoscedastic residuals) also equals the observed sample mean z_{ij}.

Now let's contrast this with the BLUP estimate. Piepho (1994) assumed E is fixed, while G and GE random with their covariance structure given by Equation 39.29a. Assuming equal number of replicates for each ij combination, the predicted yield μ_{ij} given an observed mean of z_{ij} is given by

$$\begin{aligned} \text{BLUP}(\mu_{ij}) &= \overline{z}_{.j} + h_G^2(\overline{z}_{i.} - \overline{z}_{..}) + h_{GE}^2(z_{ij} - \overline{z}_{.j} - \overline{z}_{i.} + \overline{z}_{..}) \\ &= \widehat{\mu} + \widehat{E}_j + h_G^2\,\widehat{G}_i + h_{GE}^2\,\widehat{GE}_{ij} \end{aligned} \tag{39.31a}$$

where for n_e environments, the repeatability of genetic effects and interactions, are respectively,

$$h_G^2 = \frac{\sigma_{GE}^2 + n_e\sigma_G^2}{\sigma_{GE}^2 + n_e\sigma_G^2 + \sigma_e^2}, \qquad h_{GE}^2 = \frac{\sigma_{GE}^2}{\sigma_{GE}^2 + \sigma_e^2} \tag{39.31b}$$

Contrasting the estimated cell mean under LS (given by the observed sample mean z_{ij}) with the predicted cell means under BLUP shows how BLUP shrinks the contributions from the two random effects (G_i, GE_{ij}). In particular, the BLUP contribution of the genotypic effect is $h_G^2\,\widehat{G}_i$, which is a shrinkage of the BLUE estimate $\widehat{G}_i$ back to its mean (zero). The same is true for the G x E effect. The amount of shrinkage is proportional to the lack of repeatability of these two contributions. If h^2 is near one, there is very little shrinkage, while if h^2 is near zero, its contribution is shrunk back towards nearly zero. An informal (but helpful) way of thinking about shrinkage is that the coefficient of shrinkage is the ratio of signal over signal plus noise, and is a measure of the "borrowing strength" from correlated observations. If there is little such information, there is much more noise than signal, and the resulting shrinkage is considerable, while if there is a strong signal, there is little shrinkage.

Example 39.11. The following hypothetical example of two genotypes and three environments shows how mixed-models easily handle unbalanced data. A single replicate of genotype (line) one is measured in each of three environments, while two replicates of genotype two are measured in environment one and one replicate in environment three. This design poses a problem for a strict fixed-effect analysis as the data are unbalanced and one combination is missing (genotype two in environment two). One would have to consider a reduced data set (single replicates in environments one and three), which ignores a significant fraction of the data. However, under a mixed model we can easily accommodate this data structure.

Assume environment E is a fixed effect, while G and GE are random. The resulting mixed model becomes

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{g} + \mathbf{Z}_2\mathbf{ge} + \mathbf{e}$$

The vectors of observations ($\mathbf{z}$), fixed effects ($\boldsymbol{\beta}$), and random effects (genotypic values $\mathbf{g}$, genotype x environment interactions $\mathbf{ge}$, and residuals errors $\mathbf{e}$) for this example are

$$\mathbf{z} = \begin{pmatrix} z_{111} \\ z_{121} \\ z_{131} \\ z_{211} \\ z_{212} \\ z_{231} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} E_1^* \\ E_2^* \\ E_3^* \end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}, \quad \mathbf{ge} = \begin{pmatrix} GE_{11} \\ GE_{12} \\ GE_{13} \\ GE_{21} \\ GE_{22} \\ GE_{23} \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} \epsilon_{111} \\ \epsilon_{121} \\ \epsilon_{131} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{231} \end{pmatrix}$$

Here $E_i^* = \mu + E_i$, with the E_i constrained to sum to zero. The resulting design matrices are

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The covariance matrix $\mathbf{V_z}$ for the vector of observations $\mathbf{z}$ becomes

$$\mathbf{V_z} = \mathbf{Z}_1\,\mathbf{V_g}\,\mathbf{Z}_1^T + \mathbf{Z}_2\,\mathbf{V_{ge}}\,\mathbf{Z}_2^T + \mathbf{V_e}$$

The final item to specify is the covariance structure for the three vectors of random effects ($\mathbf{g}, \mathbf{ge}, \mathbf{e}$), which we assumed are independent of each other. Under the assumptions of compound symmetry (Equation 39.29a),

$$\mathbf{V_g} = \sigma_G^2\,\mathbf{I}_{2\times 2}, \quad \mathbf{V_{ge}} = \sigma_{G\times E}^2\,\mathbf{I}_{6\times 6}, \quad \mathbf{V_e} = \sigma_e^2\,\mathbf{I}_{6\times 6}$$

REML estimation (Chapter 15, LW Chapter 26) can be used to estimate the variance components, and these can be used to provided empirical BLUP estimates by using the mixed model equations (Chapter 15, 36; LW Chapter 26).

Under compound symmetry, all the covariance matrices are a variance component times an identity matrix. More realistic models replace these simple matrices with more complex ones. We could allow residual variances to vary over lines [$\sigma^2(\epsilon_{ijk}) = \sigma_{e_i}^2$] or environments [$\sigma^2(\epsilon_{ijk}) = \sigma_{e_j}^2$], in which case $\mathbf{V_e}$ becomes a diagonal matrix with the diagonal the appropriate residual variance component (e.g., Cullis et al. 1996). For our hypothetical design, if residual variances are genotype-dependent,

$$\mathbf{V_e} = \text{diagonal}(\sigma_{e_1}^2, \sigma_{e_1}^2, \sigma_{e_1}^2, \sigma_{e_2}^2, \sigma_{e_2}^2, \sigma_{e_2}^2)$$

while if they are environment-dependent

$$\mathbf{V_e} = \text{diagonal}(\sigma^2_{e_1}, \sigma^2_{e_2}, \sigma^2_{e_3}, \sigma^2_{e_1}, \sigma^2_{e_1}, \sigma^2_{e_3})$$

Again, these can be estimated via REML. Another modification is when pedigree information exists on the genotypes, in which case $\mathbf{V_g}$ may have off-diagonal elements reflecting relationships among genotypes (Crossa et al. 2006, Oakey et al. 2007, Piepho et al. 2008). Finally, one can allow for differential correlations among genotype-environment interactions by suitably modifying $\mathbf{V_{ge}}$, a point we develop in detail shortly.

Piepho (1994) also examine the case of a fully random model where E is also assumed to be random. In this case, the covariance structure given by Equation 39.29a holds, with the addition that $E \sim N(0, \sigma^2_E)$. In this case, the BLUP prediction of cell means is given by

$$\begin{aligned}\text{BLUP}(\mu_{ij}) &= \overline{z}_{..} + h^2_E(\overline{z}_{.j} - \overline{z}_{..}) + h^2_G(\overline{z}_{i.} - \overline{z}_{..}) + h^2_{GE}(z_{ij} - \overline{z}_{.j} - \overline{z}_{i.} + \overline{z}_{..}) \\ &= \widehat{\mu} + h^2_E\widehat{E_j} + h^2_G\widehat{G_i} + h^2_{GE}\widehat{GE_{ij}} \end{aligned} \tag{39.32a}$$

where for n_g genotypes,

$$h^2_E = \frac{\sigma^2_{GE} + n_g\sigma^2_E}{\sigma^2_{GE} + n_g\sigma^2_E + \sigma^2_e} \tag{39.32b}$$

is the repeatability of environmental effects. In this case, the resulting BLUP estimate of the cell mean is the LS estimate of the overall mean plus the BLUP estimates for G_i, E_j, and GE_{ij}, the latter all being shrunken versions of their BLUE estimators under a fixed-effects analysis (Equation 39.30a). Again, the simple form for Equation 39.32a results from assuming a balanced design. For an unbalanced designs, Example 39.11 shows how using the appropriate design matrices easily handles such cases, which can then be solved using the standard mixed-model equations (Chapters 15, 35).

Moreno-González et al. (2003) contrasted the accuracy in predicting cell means using the fully random effects BLUP model (Equation 39.32) with shrinkage methods based on shrinking the factors in a fixed-effects AMMI model (by shrinking their singular values, Equation 39.22), finding that shrunken AMMIs have better predictive properties than this BLUP model. They attribute this to the fact that different shrinkage factors are applied to each separate AMMI contribution (as each singular value has its own shrinkage value) while the compound symmetry model uses only a single shrinkage correction factor (h^2_{GE}) for interaction terms. However, shrunken AMMI (being fixed-effects models) cannot account for missing cell values, while BLUP easily can. Also, the comparison is a bit unfair, as it uses the simplest covariance structure (compound symmetry) for a G x E BLUP model. Peipho (1998b) found that BLUPs with more rich covariance structures (detailed below) tended to out-perform shrunken AMMI predictors of cell means.

A simple extension of the compound symmetry model is what Peipho (1998a) calls **Shukla's model**, where Equation 39.29a is modfied to allow for different G x E variances over genotypes,

$$GE_{ij} \sim N(0, \sigma^2_{G_iE}) \tag{39.33a}$$

Note that $\sigma^2_{G_iE}$ corresponds to Shukla's stability variance (σ^2_i) for genotype i, and hence the suggested name for this model. The resulting covariance structure under Shukla's model for the observed data becomes

$$\sigma(z_{ijk}, z_{ij\ell}) = \sigma^2_G + \sigma^2_{G_iE}, \qquad \sigma(z_{ij}, z_{ik}) = \sigma^2_G \tag{39.33b}$$

Under this model, Equation 39.29d becomes

$$\mathbf{G} = \sigma^2_G \, \mathbf{J} + \text{diag}(\sigma^2_{G_1 E}, \cdots, \sigma^2_{G_{n_g} E}) \tag{39.33c}$$

The compound symmetry and Shukla models deal with assumptions about the covariance structure $\mathbf{V_{ge}}$ of the G x E interactions. As we saw in Example 39.11, we can also modify the basic mixed-model by modifying the covariance structure for the residuals $\mathbf{V_e}$.

Structured Covariance Models: Introduction

While compound symmetry and Shukla's extension are obvious starting points, both are clearly deficient. Ideally, we wish to account for two difference sources of variability: heterogeneity of residual errors (allowing different genotypes, or environments, to have a different residual variance) and heterogeneity in genotype-environment interactions (allowing different pairs of environments to have different genetic covariances and allowing genetic variances to differ over environments). Conversely, the other extreme of a completely **unstructured** covariance matrix (an arbitrary symmetric semipositive-define matrix) is usually unworkable given the difficulty in gathering enough data to estimate all covariance components with any precision. **Structured covariance matrices**, which are closely related to the multiplicative terms in biadditive models, offer a class of models that bridge these two extremes. These models start with a set of independent underlying random factors, with the covariance structure for the observations being generated by the particular statistical model chosen. As is shown below, multiplicative terms in a mixed model generate correlations, which provides the connection between covariance structures and biadditive models.

Structured Covariance Models: Finlay-Wilkinson Regressions

A good starting point to introduce these models is our old friend, the Finlay-Wilkinson regression,

$$z_{ijk} = \mu + G_i + (1 + \beta_i)E_j + \delta_{ij} + \epsilon_{ijk} \tag{39.34a}$$

Previously we analyzed this model assuming a fixed-effects framework. Digby (1979) showed that one could use an iterative least-squares approach to accommodate missing data (certain genotype-environment combinations are missing). This is reasonable, as one can borrow information from other observations in an attempt to predict the missing observation. Suppose that line five was not measured in environment three. Data from the other genotypes can be used to estimate E_3 while observations on genotype five from other environments can be used to estimate β_5, with GE_{53} being estimated by $(1 + \beta_5)E_3$. This shows how information can be borrowed from other observations under this model by using correlations between observations. In a mixed-model framework, such information borrowing occurs through the covariance matrix associated with the vector of random effects.

Treating the genetic effects (G_i and β_i) as fixed and the environmental effects (E_j and δ_{ij}) as random allows for a mixed-model analysis of the Finlay-Wilkinson regression (Oman 1991, Gogel et al. 1995, Piepho 1997a, Denis et al. 1997). Our treatment follows Piepho (1997a). Assume that the environmental effect, regression deviation, and residual error are all independent random effects and have constant variances,

$$E_j \sim N(0, \sigma^2_E), \qquad \delta_{ij} \sim N(0, \sigma^2_\delta), \qquad \epsilon_{ijk} \sim N(0, \sigma^2_e) \tag{39.34b}$$

Hence,

$$\sigma(E_j, E_\ell) = \begin{cases} 0 & j \neq \ell \\ \sigma^2_E & j = \ell \end{cases}, \qquad \sigma(\delta_{ij}, \delta_{k\ell}) = \begin{cases} 0 & ij \neq k\ell \\ \sigma^2_\delta & ij = k\ell \end{cases} \tag{39.35b}$$

The variance of the trait value from an individual from line i randomly drawn accross environments follows by using Equation 39.35b and taking the variance of Equation 39.34a, giving

$$\sigma^2(z_{ijk}) = (1+\beta_i)^2\sigma_E^2 + \sigma_\delta^2 + \sigma_e^2 \tag{39.35c}$$

Peipho (1997a) notes that the regression residual and normal residual variances (if both homoscedastic) cannot be separately estimated and hence can be combined into a single general residual variance. The covariance between two different genotypes (i and k) in the same environment (j) similarly becomes

$$\begin{aligned}\sigma^2(z_{ij}, z_{kj}) &= \sigma\left[\,(1+\beta_i)E_j + \delta_{ij}, (1+\beta_k)E_j + \delta_{kj}\,\right] \\ &= (1+\beta_i)(1+\beta_k)\sigma_E^2\end{aligned} \tag{39.35b}$$

Let $\mathbf{z}_j$ be a vector of observations of the line means within environment j (for simplification we assume a single observation, but multiple, and unequal, replication is easily accommodated by modification of $\mathbf{V_e}$). In matrix form, the covariance matrix for $\mathbf{z}_j$ is

$$\mathbf{V}_{\mathbf{z}_j} = \sigma_E^2\,\boldsymbol{\lambda}\boldsymbol{\lambda}^T + (\sigma_\delta^2 + \sigma_e^2)\mathbf{I}, \quad \text{where} \quad \boldsymbol{\lambda} = \begin{pmatrix} 1+\beta_1 \\ \vdots \\ 1+\beta_{n_g} \end{pmatrix} \tag{39.36a}$$

Observe that the assumed structure of the model (Equation 39.34a) translates underlying independent random effects (E, δ) into correlated effects across the vector $\mathbf{z}$ of observations. This is a simple example of a **factor-analytic covariance structure** (Jennrich and Schluchter 1986, also see Example 31.5) where the covariance structure is determined by a small number of interacting factors. Piepho (1997a) and Denis et al. (1997) showed how this general framework can be extended to cases (e.g., Shulka 1972) where the lines have different variances,

$$\mathbf{V}_{\mathbf{z}_j} = \sigma_E^2\,\boldsymbol{\lambda}\boldsymbol{\lambda}^T + \text{diag}(\sigma_{\delta_1}^2, \cdots, \sigma_{\delta_{n_g}}^2) + \sigma_e^2\,\mathbf{I} \tag{39.36b}$$

Likewise much more general covariance structures for the residuals can be incorporated,

$$\mathbf{V}_{\mathbf{z}_j} = \sigma_E^2\,\boldsymbol{\lambda}\boldsymbol{\lambda}^T + \text{diag}(\sigma_{\delta_1}^2, \cdots, \sigma_{\delta_{n_g}}^2) + \mathbf{V_e} \tag{39.36c}$$

The curious reader might have asked what would happen if we treated environmental effects as fixed and genetic effects as random in Equation 39.34a. In this case, E_j is now a fixed effect, while G_i, β_i and δ_{ij} are now random, and the analog of Equation 39.33b becomes

$$G_i \sim N(0, \sigma_G^2), \qquad \beta_i \sim N(0, \sigma_\beta^2), \qquad \delta_{ij} \sim N(0, \sigma_{\delta_j}^2), \qquad \epsilon_{ijk} \sim N(0, \sigma_e^2) \tag{39.37a}$$

The resulting variance of a single observation becomes

$$\sigma^2(z_{ijk}) = \sigma_G^2 + E_j^2\sigma_\beta^2 + \sigma_{\delta_j}^2 + \sigma_e^2 \tag{39.37b}$$

while the covariance of interest in now that between the same genotype (i) over different environments (j and k),

$$\sigma^2(z_{ij}, z_{ik}) = \sigma^2(G_i + \beta_i E_j + \delta_{ij}, G_i + \beta_i E_k + \delta_{ik}) = \sigma_G^2 + E_j E_k \sigma_\beta^2 \tag{39.37c}$$

giving the covariance matrix for the vector $\mathbf{z}_i$ of genotype i over all n_e environment as

$$\mathbf{V}_{\mathbf{z}_i} = \sigma_G^2\,\mathbf{J} + \sigma_\beta^2\boldsymbol{\lambda}\boldsymbol{\lambda}^T + \sigma_e^2\mathbf{I}, \quad \text{with} \quad \boldsymbol{\lambda} = \begin{pmatrix} E_1 \\ \vdots \\ E_{n_e} \end{pmatrix} \tag{39.37d}$$

Structured Covariance Models: AMMI Models

The same logic used to generated a mixed-model Finlay-Wilkinson regression easily extends to other biadditive models, such as AMMI (Piepho 1997a,1998a; Denis et al. 1997; Smith et al. 2005). Our treatment follows Piepho (1998a). Recalling Equation 39.19b, an AMMI$_m$ model is given by

$$\begin{aligned} z_{ij\ell} &= \mu + G_i + E_j + \sum_{k=1}^{m} \lambda_k \, \gamma_{ki} \, \eta_{kj} + \delta_{ij} + \epsilon_{ij\ell} \\ &= \mu + G_i + E_j + \sum_{k=1}^{m} w_{ki} \, v_{kj} + \epsilon^*_{ij\ell} \end{aligned} \tag{39.38a}$$

The second line simplifies the AMMI model (to allow for estimability) in two ways. First, the singular value λ_k is absorbed into the genotype sensitivity (w_{ki}) and environmental (v_{kj}) coefficients. Second, the error in predicting GE from the AMMI approximation (δ_{ij}) and the model residual ($\epsilon_{ij\ell}$) are combined into a single residual $\epsilon^*_{ij\ell}$. Assume genotypes are random and environments are fixed, so that E_j and v_{kj} are fixed, while G_i and w_{ki} are random (as is, of course, $\epsilon^*_{ij\ell}$). Assume these underlying components are independent and homoscedastic,

$$G_i \sim N(0, \sigma^2_G), \qquad w_{ki} \sim N(0, \sigma^2_k), \qquad \epsilon^*_{ijk} \sim N(0, \sigma^2_e + \sigma^2_\delta) \tag{39.38b}$$

The resulting variance for the trait value of a random individual drawn from environment j becomes

$$\sigma^2(z_{ijk}) = \sigma^2_G + \sum_{k=1}^{m} \sigma^2_k \, v^2_{kj} + \sigma^2_\delta + \sigma^2_e \tag{39.38c}$$

Hence, the resulting genetic variance in environment j is just

$$\sigma^2(z_{ij}) = \sigma^2_G + \sum_{k=1}^{m} \sigma^2_k \, v^2_{kj} + \sigma^2_\delta \tag{39.38d}$$

Likewise, the covariance between the same random genotype over different environments (j and ℓ) becomes

$$\begin{aligned} \sigma^2(z_{ij}, z_{i\ell}) &= \sigma \left(G_i + \sum_{k=1}^{m} w_{ki} \, v_{kj}, G_i + \sum_{k=1}^{m} w_{ki} \, v_{k\ell} \right) \\ &= \sigma^2_G + \sum_{k=1}^{m} \sigma(w_{ki} \, v_{kj}, w_{ki} \, v_{k\ell}) = \sigma^2_G + \sum_{k=1}^{m} v_{kj} \, v_{k\ell} \, \sigma(w_{ki}, w_{ki}) \\ &= \sigma^2_G + \sum_{k=1}^{m} \sigma^2_k \, v_{kj} \, v_{k\ell} \end{aligned} \tag{39.38d}$$

The resulting covariance matrix for the vector $\mathbf{z}_i$ of observations of genotypes over the n_e environments can be written in the form

$$\mathbf{V}_{\mathbf{z}_i} = \sigma^2_G \, \mathbf{J} + \boldsymbol{\lambda}\boldsymbol{\lambda}^T + \sigma^2_e \mathbf{I}$$

where $\mathbf{J}$ is a matrix of ones, and $\boldsymbol{\lambda}$ is the $n_e \times m$ matrix,

$$\boldsymbol{\lambda} = (\, \boldsymbol{\lambda}_1 \quad \cdots \quad \boldsymbol{\lambda}_m \,), \quad \text{where} \quad \boldsymbol{\lambda}_k = \sigma^2_k \begin{pmatrix} v_{k1} \\ \vdots \\ v_{kn_e} \end{pmatrix} = \begin{pmatrix} \lambda_{k1} \\ \vdots \\ \lambda_{kn_e} \end{pmatrix} \tag{39.38e}$$

Because the model is over-parameterized, we can set $\sigma_k^2 = 1$ for $1 \leq k \leq m$ (Piepho 1998a) so that the factor-loading terms $\lambda_{kj} = \sigma_k^2 v_{kj}$ simply equal the environment coefficients v_{kj}. Equation 39.38e is not yet in final form, as while $\boldsymbol{\lambda}$ is an $n_e \times k$ matrix, the number of free variance parameters is $n_e k - k(k-1)/2$, so that $k(k-1)/2$ elments of $\boldsymbol{\lambda}$ are constrained. Peipho (1998a) accommodates this by setting the first element in the vector $\boldsymbol{\lambda}_2$ to zero, the first two elements in $\boldsymbol{\lambda}_3$ to zero, and in general the first $k-1$ elements of $\boldsymbol{\lambda}_k$ to zero.

Note that this model is the random effects equivalent of an AMMI$_m$ model. As such, one can also generate the equivalent of biplots, plotting the first two scaled sensitivity values $(\lambda_{1j}, \lambda_{2j})$ to generate the biplot vector the environment j, while the BLUPs of (w_{1i}, w_{2i}) generate the biplot vector for genotype i. This is the basic idea, but their are a few subtleties. For example, Peipho (1997b) suggested that the sensitivity show be shifted (using $\lambda_{kj} - \lambda_{k\cdot}$) to remove $\lambda_{k\cdot} w_{jk}$ which is the contribution to the environment main effect. Also, given the constraint of setting $\lambda_{21} = 0$, the resulting biplots are generally rotated (see Smith et al. 2001 for detail), but this does not change the relative orientation of the environmental and genotypic vectors with respect to each other. See Casanoves et al. (2005a) for an example of using such a random AMMI model.

Table 39.3 gives the resulting covariance structures for the various mixed-models we have assumed. When G is random and E fixed, the variances $\sigma^2(z_{ij})$ are for a randomly-drawn genotype (i is the index of the random variable) in environment j and the covariances $\sigma(z_{ij}, z_{ik})$ are for a randomly-drawn genotype measured in environments j and k. When we take E as random, the roles in these covariances are reversed, see the table caption for details.

Table 39.3. Summary the covariance structures for various mixed-models for G x E discussed in this chapter, where either G or E is taken as random with the other fixed. When G is taken as random with E fixed, $\sigma^2(z_{ij})$ is genetic variance in environment j, while $\sigma(z_{ij}, z_{ik})$ is the covariance between a random genotype (i) measured in environments j and k. When E taken as random, $\sigma^2(z_{ij})$ corresponds to the variance for an individual from genotype i drawn from a random environment, while $\sigma(z_{ij}, z_{kj})$ is the covariance between genotypes i and k when measured in across a random environment (j). C corresponds to the Compound Symmetry model (Equation 39.29), S is Shukla's extension (Equation 39.33), FW is Finlay-Wilkinson (Equations 39.34, 37) where $\alpha_i = 1 + \beta_i$, FA(m) is factor-analytic model (i.e., a mixed AMMI-type model) with m factors (Equation 39.38), U is the completely Unstructured model.

	G random, E Fixed		E random, G Fixed	
	$\sigma^2(z_{ij})$	$\sigma(z_{ij}, z_{ik})$	$\sigma^2(z_{ij})$	$\sigma(z_{ij}, z_{kj})$
Index range	$1 \leq j \leq n_e$	$1 \leq k, j \leq n_e$	$1 \leq i \leq n_g$	$1 \leq i, k \leq n_g$
C	$\sigma_G^2 + \sigma_{GE}^2$	σ_G^2	$\sigma_E^2 + \sigma_{GE}^2$	σ_E^2
S	$\sigma_G^2 + \sigma_{GE_j}^2$	σ_G^2	$\sigma_E^2 + \sigma_{G_iE}^2$	σ_E^2
FW	$\sigma_G^2 + E_j^2\sigma_\beta^2 + \sigma_{\delta_j}^2$	$\sigma_G^2 + E_j E_k \sigma_\beta^2$	$\alpha_i^2\sigma_E^2 + \sigma_{\delta_i}^2$	$\alpha_i\alpha_k\sigma_E^2$
FA(m)	$\sigma_G^2 + \sum_\ell^m \lambda_{\ell j}^2$	$\sigma_G^2 + \sum_\ell^m \lambda_{\ell j}\lambda_{\ell k}$	$\sigma_E^2 + \sum_\ell^m \lambda_{\ell i}^2$	$\sigma_E^2 + \sum_\ell^m \lambda_{\ell i}\lambda_{\ell k}$
U	σ_j^2	σ_{jk}	σ_i^2	σ_{ik}

PUTTING IT ALL TOGETHER

We are now in a position to unify the three major topics of this chapter (measures of stability,

approximation of interaction terms using a bilinear models, and mixed models) and show how these are connected with selection response. The key is the covariance matrix of the z_{ij}, whose structure can be approximated by the elements in a bilinear mixed-model (Table 39.3). Elements of the covariance matrix correspond to various stability measures and the covariance matrix can be used to obtain the $\mathbf{G}$ matrix in order to apply the multivariate breeder's equation to selection response when G x E is present. Generally, issues of stability assume that G is fixed, E is random as we are interested in the stability of a particular genotype over random environments. Conversely, selection response usually assumes that E is fixed and G is random, as we are interested in how a random genotype covaries over a known (fixed) set of environments.

Structured Covariance Matrices and Stability Measures

The connection between stability measures and structured covariance matrices was stressed in an important (and undercited) paper by Piepho (1998a), who emphasized two key points: "the problem of choosing the appropriate stability measure may be view as a problem of choosing the appropriate variance covariance structure" in a mixed model, and the "usefulness of any measure of stability depends crucially on how well the underlying model approximates the real data".

Example 39.12. Let's further examine the connection between Type 1 and Type 2 stability. Throughout we take G as fixed (as we are interested in the stability of a particular genotype) and E as random. Recall that Type 1 (static) stability is the variance of a genotype over environments. The underlying statistical model is the simple ANOVA $z_{ij} = \mu + G_i + \epsilon_{ij}$ where ϵ_{ij} is the random error about this effect (due to environmental effects, G x E interaction, and residual error). We assume ϵ_{ij} are uncorrelated across genotypes, but not necessarily homoscedastic, so that $e_{ij} \sim (0, \sigma^2_{G_i})$. Type 1 stability is indicated by a small value of $\sigma^2_{G_i}$, namely a small variance of genotype i over environments. As noted by Piepho (1998a), the difference in Type 1 stability for genotypes i and k is $\sigma^2_{G_i} - \sigma^2_{G_k}$.

By constrast, Type 2 stability is indicated by a small G x E variance associated with a genotype. One underlying statistical model for this is

$$z_{ijk} = \mu + G_i + E_j + GE_{ij} + \epsilon_{ijk}$$

where E, GE, and ϵ are taken to be random and independent. Under Shukla's model of covariance structure, $E_j \sim (0, \sigma^2_E)$, $GE_{ij} \sim (0, \sigma^2_{G_iE})$, and $\epsilon_{ijk} \sim (0, \sigma^2_e)$, giving the variance of a single observation of genotype i from a random environment as

$$\sigma^2_{G_i} = \sigma^2_E + \sigma^2_{G_iE} + \sigma^2_e$$

Type 2 (dynamic) stability is defined by having a small value of $\sigma^2_{G_iE}$, the variance associated with the G x E interaction terms for genotype i. Under Shukla's model we have

$$\sigma^2_{G_i} - \sigma^2_{G_k} = \left(\sigma^2_E + \sigma^2_{G_iE} + \sigma^2_e\right) - \left(\sigma^2_E + \sigma^2_{G_kE} + \sigma^2_e\right) = \sigma^2_{G_iE} - \sigma^2_{G_kE}$$

Thus, for the covariance structure assumed in Shukla's model, the difference between two genotypes in Type 1 stability also equals their difference under Type 2 stability. The key assumption leading to this equivalence is that $\sigma(E_j, GE_{ij}) = 0$ (Weber et al 1996).

Now suppose that the correct biological model is best captured by the mixed model version of Finlay-Wilkinson regression, so that

$$GE_{ij} = \beta_i E_j + \delta_{ij}$$

where the residual $\delta_{ij} \sim (0, \sigma^2_{\delta_i})$ and is uncorrelated with E. Under this model, GE and E are no longer uncorrelated as

$$\sigma(GE_{ij}, E_j) = \sigma(\beta_i E_j + \delta_{ij}, E_j) = \beta_i \sigma^2_E$$

The variance for genotype i over random environments under this model becomes

$$\sigma^2_{G_i} = (1 + \beta_i)^2 \sigma^2_E + \sigma^2_{\delta_i} + \sigma^2_e$$

while the genotype-specific interaction variance can be expressed as

$$\sigma^2_{G_i E} = \sigma^2(GE_{ij}) = \sigma^2(\beta_i E_j + \delta_{ij}) = \beta_i \sigma^2_E + \sigma^2_{\delta_i}$$

A little algebra shows that

$$\left(\sigma^2_{G_i} - \sigma^2_{G_k}\right) - \left(\sigma^2_{G_i E} - \sigma^2_{G_k E}\right) = 2(\beta_k - \beta_i)\sigma^2_E$$

with two genotypes (i and k) showing the same difference in Type 1 and 2 stability only when $\beta_i = \beta_k$. As Piepho (1998a) notes, "*what is often seen as a difference of* [stability] *concepts may be viewed as a difference of models, i.e., a difference of variance-covariance structures*".

Besides important theoretical connections to stability measures, mixed-models have two very practical features. One is there enormous flexibility in handling unbalanced designs. The second is the shrinkage effects of BLUP estimates resulting in less bias relative to least-squares estimates. For example, simulations by Peipho (1993) shows that ML-based estimates of Shukla's stability variance tend (as the number of environments increases) to provide a better ranking of genotypes relative to stability estimates based on Equation 39.6c. When the data are unbalanced, these simple least-squares expressions no long apply, but REML estimated under a mixed-model framework can easily be applied. More generally, mixed-model analysis can be expanded into a bayesian framework (Edwards and Jannink 2006, Cotes et al. 2006). A bayesian analysis explicitly accounts for the uncertainty in using REML estimates of variances (as opposed to the true variances) in the BLUP equations (Chapter 15, Appendix 2).

Structured Covariance Matrices and Selection Response

We started our introduction of selection under G x E in Chapter 38 by placing it within the framework of selection on correlated traits. Under this framework, the expected vector of responses over a series of environments given selection within each (or some subset) is given by $\mathbf{GP}^{-1}\mathbf{S}$. As mentioned in Chapter 38, the problem with widely applying the multivariate breeder's equation is the difficulty of estimating $\mathbf{G}$ with any precision for more than two environments. The concern is that $\mathbf{G}$ matrix is completely unstructured, requiring a large number of covariances to be estimated with some precision. However, when $\mathbf{G}$ has some structure, we now have a potential solution, namely using factor-analytic estimates of $\mathbf{G}$ as obtained from mixed-models (Table 39.3). In particular, treating environments as fixed effects (as the concern is response over a specified set of environments, for example locations), then AMMI-type mixed models (taking G random) offer estimates of the covariance matrix $\mathbf{G}$.

If the concern is predicting selection response over a set of environments (such as particular locations), then if $\mathbf{z}$ is a vector of line means over this set of environments and $\boldsymbol{\mu}$ is the vector of means over this environment, the vector $\mathbf{S}$ of selection differentials is given by

$$\mathbf{S} = E_s(\mathbf{z} - \boldsymbol{\mu}) \tag{39.39a}$$

where the expectation is taken over those lines selected. Thus, the jth entry in $\mathbf{S}$ is just

$$s_j = \mu_j^* - \mu_j \tag{39.39b}$$

where μ_j^* is the mean of the selected lines in that environment. The resulting selection response follows from the multivariate breeder's equation,

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \mathbf{G}\left(\mathbf{G} + \mathbf{V_e}\right)^{-1}\mathbf{S} \tag{39.39c}$$

where $\mathbf{G}$ is estimated using one of the models in Table 39.3 and $\mathbf{V_e}$ is the covariance matrix of the residuals. Standard results index selection theory also apply. Suppose the goal is the expected response over n environments, where the desired gains are weighted by

$$I = a_1 g_1 + \cdots + a_n g_n = \sum_{i=1}^{n} a_i g_i = \mathbf{a}^T\mathbf{g}$$

The response in I is simply

$$\mathbf{a}^T\mathbf{R} = \mathbf{a}^T\mathbf{G}\left(\mathbf{G} + \mathbf{V_e}\right)^{-1}\mathbf{S} \tag{39.40a}$$

Likewise, a factor-analytic approximation of a structured $\mathbf{G}$ matrix could also be used in a Smith-Hazel index to obtain the weights $\mathbf{b}$ to maximinze the response in I (Chapter 33), namely

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a} = \left(\mathbf{G} + \mathbf{V_e}\right)^{-1}\mathbf{G}\mathbf{a} \tag{39.40b}$$

Likewise, one could base selection on BLUP estimates of genotype means over environments. When the number of replicates of lines over environments very similiar, then BLUP and the multivariate breeder's equation are also very similar. The strength of BLUP is that it allows for more complex structures of residuals (such as differences in the number of replicates over genotypes) to be easily handled. Piepho and Möhring (2005, 2007) examine BLUP applications to selection when G x E is present.

Literature Cited

Abou-El-Fittouh, H. A., J. O. Rawlings, and P. A. Miller. 1969. Genotype by environment interactions in cotton — their nature and related environmental variables. *Crop Sci.* 9: 377–381. [39]

Annicchiarico, P. 1997a. Additive main effects and multiplicative interaction (AMMI) analysis of genotype-location interaction in variety trails repeated over years. *Theor. Appl. Genet.* 94: 1072–1077. [39]

Annicchiarico, P. 1997b. Joint regression vs AMMI analysis of genotype-environment interactions for cereals in Italy. *Euphytica* 94: 53–62. [39]

Annicchiarico, P., and M. Perenzin. 1994. Adaptation patterns and definition of macro-environments for selection and recommendation of common-wheat genotypes in Italy. *Plant Breeding* 113: 197–205. [39]

Barah, B. C., H. P. Binswanger, B. S. Rana, and N. G. P. Rao. 1981. The use of risk aversion in plant breeding; concept and application. *Euphytica* 30: 451–458. [39]

Baril, C. P. 1992. Factor regression for interpreting genotype-environment interaction in bread-wheat trails. *Theor. Appl. Genet.* 83: 1022–1026. [39]

Baril, C. P., J.-B. Denis, R. Wustman, and F. A. van Eeuwijk. 1995. Analysing genotype by enviornment interactions in Dutch potato variey trails using factorial regression. *Euphytica* 82: 149–155. [39]

Bawa, V. S. 1975. Optimal Rules for Ordering Uncertain Prospects. *J. Financial Econ.* 2: 95–121. [39]

Becker, H. C. 1981. Correlations among some statistical measures of phenotypic stability. *Euphytica* 30: 835–840. [39]

Becker, H. C., and J. Léon. 1988. Stability analysis in plant breeding. *Plant Breeding* 101: 1–23. [39]

Brancourt-Hulmel, M., and C. Lecomte. 2003. Effect of environmental variates on genotype x environment interaction of winter wheat: a comparison of biadditive factorial regression to AMMI. *Crop Sci.* 43: 608–617. [39]

Breese, E. L. 1969. The measurement and significance of genotype-environment interactions in grasses. *Heredity* 24: 27—44. [39]

Calderini, D. F., and G. A. Slafer. 1999. Has yield stability changed with genetic improvement of wheat yield? *Euphytica* 107: 51–59. [39]

Calinski, T. 1960. On a certain statistical method of investigating interaction in serial experiments with plant varieties. *Bull. L'Academie Polonasise des Sciences Cl. II.* Vol VIII, no. 12: 565–568. [39]

Campos, M. Cooper, J. E. Habben, G. O. Edmeades, and J. R. Schussler. 2004. Improving drought tolerance in maize: a view from industry. *Field Crops Res.* 90: 19–34. [39]

Casanoves, F., J. Baldessari, and M. Balzarini. 2005a. Evaluation of multienvironment trials of peanut cultivars. *Crop Sci.* 45: 18–26. [39]

Casanoves, F., R. Macchiavelli, and M. Balzarini. 2005b. Error variation in multienvironment peanut trails: within-trail spatial correlation and between-trail heterogenity. *Crop Sci.* 45: 1927–1933. [39]

Cooper, M. and I. H. DeLacy. 1994. Relationships among analytical methods used to study environmental variation and genotype-by-environment interaction in plant breeding multi-environmental experiments. *Theor. Appl. Genet.* 88: 561–572. [39]

Cornelius, P. L. 1993. Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trails. *Crop Sci.* 33: 1186–1193. [39]

Cornelius, P. L., and J. Crossa. 1999. Prediction assessment of shrinkage estimators of multiplicative models for multi-environmental cultivar trails. *Crop Sci.* 39: 998–1009. [39]

Cornelius, P. L., J. Crossa, and M. S. Seyedsadr. 1996. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. *In* M.S. Kang and H.G. Gauch Jr. (eds) *Genotype-by-environment interaction: New perspectives*, pp. 199–234. CRC Press, Boca Raton, Florida. [39]

Cotes, J. M., J. Crossa, A. Sanches, and P. L. Cornelius. 2006. A bayesian approach for assessing the stability of genotypes. *Crop Sci.* 46: 2654–2665. [39]

Crossa, J. 1990. Statistical analysis of multilocation trails. *Adv. Agron.* 44: 55–85. [39]

Crossa, J., J. Burgueño, P. L. Cornelius, G. McLaren, R. Trethowan, and A. Krishnamachari. 2006. Modeling genotype x environment interaction using additive genetic covariances of relatives for predicting breeding vaues of wheat genotypes. *Crop Sci.* 46: 1722–1733. [39]

Crossa, J., P. L. Cornelius, 1997. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances *Crop Sci.* 37: 406–415. [39]

Crossa, J., and P. L. Cornelius. 2002. Linear-bilinear models for the analysis of genotype-environment interaction. *In* M. S. Kang (ed) *Quantitative genetics, genomics, and plant breeding*, pp. 305–322. [39] CABI, Wallingford.

Crossa, J., P. L. Cornelius, and W. Yan. 2002. Biplots of linear-bilinear models for studying genotype x enviroment interaction. *Crop Sci.* 42: 619–633. [39]

Cullis, B. R., R. M. Thomson, J. A. Fisher, A. R. Gilmour, and R. Thompson. 1996. The analysis of the NSW wheat variety database. I. Modelling trail error variance. *Theor. Appl. Genet.* 92: 21–27. [39]

Dashiell, K. E., O. J. Ariyo, L. Bello, and K. Ojo. 1994. Genotype x environment ineraction and simultaneous selection for high yield and stability in soybeans (*Glycine max* (L.) Merr.). *Ann. Appl. Biol.* 124: 133–139. [39]

Denis, J.-B., and J. C. Gower. 1996. Asymptotic confidence regions for biadditive models: interpreting genotype-environment interactions. *Appl. Stat.* 45: 479–493. [39]

Denis, J.-B., H.-P. Piepho, and F. A. van Eeuijk. 1997. Modelling expectation and variance for genotype by environment data. *Heredity* 79: 162–171. [39]

Digby, P. G. N. 1979. Modified joint regression analysis for incomplete variety x environment data. *J. Agric. Sci.* 93: 81–86. [39]

dos S. Dias, C. T., and W. J. Krzanowski. 2003. Model selection and cross validation in additive main effect and multiplicative interaction models. *Crop Sci.* 43: 865–873. [39]

Eberhart, S. A., and W. A. Russell. 1966. Stability parameters for comparing varieties. }si Crop Sci. 6: 36–40. [39]

Eckart, C., and G. Young. 1936. The approximation of one matrix by another of lower rank. *Pyschometrika* 1: 211–218. [39]

Edmeades, G. O., G. S. McMaster, J. W. White, and H. Campos. 2004. Genomics and the physiologist: bridging the gap between genes and crop response. *Field Crops research* 90: 5–18. [39]

Edwards, J. W., and J.-L. Jannink. 2006. Bayesian modeling of heterogeneous eroor and genotype x environment interaction variances. *Crop Sci.* 46: 820–833. [39]

Epinat-Le Signor, C., S. Dousse, J. Lorgeou, J.-B. Denis, R. Bonhomme, P. Carolo, and A. Charcosset. 2001. Interpretation of genotype x environment interactions for early maize hybrids over 12 years. *Crop Sci.* 41: 663–669. [39]

Eskridge, K. M. 1990. Selection of stable cultivars using a safety-first rule. *Crop Sci.* 30: 369–374. [39]

Eskridge, K. M. 1991. Screening cultivars for yield stability to limit the probability of disaster. *Maydica* 36: 275–282. [39]

Eskridge, K. M. and R. F. Mumm. 1992. Choosing plant cultivars based on the probability of outperforming a check. *Theor. Appl. Genet.* 84: 494–500. [39]

Fan, X.-M., M. S. Kang, H. Chen, Y. Zhang, J. Tan, and C. Xu. 2007 Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. *Agron. J.* 99: 220–228. [39]

Fataunla, T., and K. J. Frey. 1976. Repeatability of regression stability indexes for grain yield in oats. *Avena sative* L.). *Euphytica* 25: 21–28. [39]

Fernandez, G. C. J. 1991. Analysis of genotype x environment interaction by stability estimates. *HortSci.* 26: 947–950. [39]

Finlay, K. W., and G. N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programe. *Aust. J. Agr. Res.* 14: 742–754. [39]

Fisher, R. A., and W. A. MacKenzie. 1923. Studies in crop variation: II. The maurial response of different potato varieties. *J. Agric. Sci.* 13: 311–320. [39]

Flores, F., M. T. Moreno, and J. I. Cubero. 1998. A comparison of univaraite and multivariate methods to analyze G x E interaction. }si Field Crops Res. 56: 271–286. [39]

Francis, T., and L. W. Kannenberg. 1978. Yield stability studies in short-season maize. 1. A descriptive method for grouping genotypes. *Can. J. Plant Sci.* 58: 1029–1034. [39]

Freeman, G. H. 1973. Statistical methods for the analysis of genotype-environment interactions. *Heredity* 31: 339–354. [39]

Freeman, G. H., and J. M. Perkins. 1971. Environmental and genotype-environmental components of variability. VIII. Relations between genotypes grown in different environment and measures of these environments. *Heredity* 27: 15–23. [39]

Frey, K. J. 1964. Adaptation reaction of oat strains under stress and non-stress environments. *Crop Sci.* 4: 55–58. [39]

Fripp, Y. J., and C. E. Caten. 1971. Genotype-environmental interactions in*Schizophyllum commune.* I. Analysis and character. *Heredity* 27: 393–407. [39]

Gabriel, K. R. 1971. The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58: 453–467. [39]

Gabriel, K. R. 1978. Least squares approximation of matrices by additive and multiplicative models. *J. Roy. Stat. Soc. b* 40: 186–196. [39]

Gauch, H. G. 1988. Model selection and validation for yield traits with interaction. *Biometrics* 44: 705–715. [39]

Guach, H. G., Jr. 2006. Statistical analysis of yield trails by AMMI and GGE. *Crop Sci.* 46: 1488–1500. [39]

Guach, H. G., Jr., H.-P. Piepho, and P. Annicchiarico. 2008. Statistical analysis of yield trails by AMMI and GGE: Further considerations. *Crop Sci.* 48: 866–889. [39]

Gauch, H. G. and R. W. Zobel. 1988. Predicitive and postdictive success of statistical analysis of yield trails. *Theor. Appl. Genet.* 76: 1-10. [39]

Gauch, H. G. and R. W. Zobel. 1997. Identifying mega-environments and targeting genotypes. *Crop Sci.* 37: 311–326. [39]

Gogel, B. J., B. R. Cullis, and A. P. Verbyla. 1995. REML estiamtion of multiplicative effects of multi-environment variety trails. *Biometrics* 51: 744–749. [39]

Gollob, H. F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika* 33: 73–115. [39]

Good, I. J. 1969. Some applications of the singular decomposition of a matrix. *Technomet.* 11: 823–831. [39]

Gower, J. C., and D. J. Hand. 1996. *Biplots.* Chapman and Hall, UK. [39]

Hardwick, R., and J. T. Wood. 1972. Regression methods for studying genotype-environment interactions *Heredity* 28: 209–222. [39]

Haufe, W., and H. Geidel. 1978. Zur beurteilung der ertagssicherheit von sorten und zuchstammen. I. Definitionen. Stabilitätsparameter und deren interpretationsmöglichkeiten. *Z. Pflanzenüchtg* 80: 24–37. [39]

Hawkins, D. M., L. Liu, and S. S. Young. 2001. Robust singular value decomposition. National Institute of Statistical Sciences. Technical Report 122. [39]

Hill, J. 1975. Genotype-environment interactiuons — a challenge for plant breeding. *J. Agric. Sci.* 85: 477–493. [39]

Hill, R. R. Jr., and J. L. Rosenberger. 1985. Methods for combining data from germplasm evaluation trails. *Crop Sci.* 25: 467–470. [39]

Horner, T. W., and K. J. Frey. 1957. Methods for determining natural areas for oat varietal recommendations. *Agron. J.* 49: 313–315. [39]

Huehn, M. 1990a. Nonparametric measures of phenotypic stability. Part 1: Theory. *Euphytica* 47: 189–194. [39]

Huehn, M. 1990b. Nonparametric measures of phenotypic stability. Part 2: Applications. *Euphytica* 47: 195–201. [39]

Hühn, M. 1987. Quantitative erfassung der ertragssicherheit — aktueller stand, offene fragen und mögliche entwicklungen. *Vortr. Pflanzenzüchtg.* 12: 37–61. [39]

Hühn, M. 1979. Beiträge zur Erfassung der phänotypischen stabilität. I. Vorschlag einiger auf ranginformationen beruhenden stabilitätsparameter. *EDV in Medizin und Biologie* 10: 112–117. [39]

Hühn, M., and J. Léon. 1985. Genotype-environment interactions and phenotypic stability of *Brassica napus* L. *Z. Pflanzenzüchtg.* 95: 135–146. [39]

Jennrich, R. L., and M. D. Schluchter. 1986. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* 42: 805–820. [39]

Jensen, N. F. 1976. Floating checks for plant breeding nurseries. *Ceral Res. Comm.* 4: 285–295. [39]

Kang, M. S. 1998. Using genotype-by-environment interaction for crop cultivar development. *Adv. Argon.* 62: 199–252. [39]

Kang, M. S., and H. G. Gauch, Jr. (eds). 1996. *Genotype-by-environment interaction* CRC Press, Boca Raton, Florida. [39]

Kang, M. S., J. D. Miller, and L. L. Darrah. 1987. A note on the relationship between stability variance and ecovalence. *J. Heredity* 78: 107. [39]

Kang, M. S., and H. N. Pham. 1991. Simultaneous selection for high yielding and stable crop genotypes. *Agron. J.* 83: 161–165. [39]

Kempton, R. A. 1984. The use of biplots in interpreting variety by environment interactions. *J. Agric. Sci. Camb.* 103: 123–135. [39]

Kroonenberg, P. M. 1995. Introduction to biplots for G x E tables. Research Report No. 51, Centre for Statistics, University of Queensland, Brisbane. [39]

Laffont, J.-L., M. Hanafi, and K. Wright. 2007. Numerical and graphical measures to facilitate the interpretation of GGE biplot. *Crop Sci.* 47: 990–996. [39]

Léon, J. 1986. Methods of simultaneous estimation ofyield and yield stability. *In: Biometrics in plant breeding*. Proc. 6th Meeting ECUARPIA Section biometrics in plant breeding, Birimingham, UK, pp. 299–308. [39]

Léon, J., and H. C. Becker. 1988. Repeatability of some statistical measures of phenotypic stability — correlations between single year results and multi year results. *Plant Breeding* 100: 137–142. [39]

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: 425–430. [39]

Lin, C. S., and M. R. Binns. 1991. Genetic properties of four types of stability parameter. *Theor. Appl. Genet.* 82: 505–509. [39]

Lin, C. S., and M. R. Binns. 1994. Concepts and methods for the analysis of regional trail data for cultivar and location selection. *Plant Breed. Rev.* 12: 271–297. [39]

Lin, C. S., M. R. Binns, and L. P. Lefkovitch. 1986. Stability analysis: where do we stand? *Crop Science* 26: 894–900. [83]

Mandel, J. 1961. Non-additivity in two-way analysis of variance. *J. Am. Stat. Assoc.* 65: 878–888. [3]

Mandel, J. 1969. The partititioning of interaction in analysis of variance. *J. Res. Int. Bur. Stand. B* 73: 309–328. [39]

Mead, R., J. Riley, K. Dear, and S. P. Singh. 1986. Stability comparisons of intercropping and monocropping systems. *Biometrics* 42: 253–266. [39]

Miezan, K., G. A. Milliken, and G. H. Liang. 1979. Using regression coefficients as a stability parameter in plant breeding programs. *Theor. Appl. Genet.* 54: 7–9. [39]

Mooers, C. A. 1921. The agronomic placement of varieties. *J. Amer. Soc. Agron.* 13: 337–352. [39]

Moreno-González, J. Crossa, and P. L. Cornelius. 2003. Additive main effects and multiplicative interaction model: I. Theory on variance components for predicting cell means. *Crop sci.* 43: 1967–1975. [39]

Moreno-González, J. Crossa, and P. L. Cornelius. 2003. Additive main effects and multiplicative interaction model: II. Theory on shrinkage factors for predicting cell means. *Crop sci.* 43: 1976–1982. [39]

Moreno-González, J. Crossa, and P. L. Cornelius. 2004. Genotype x environment interaction in multi-environment trails using shrinkage factors of ammi models. *Euphytica* 137: 119-127. [39]

Nassar, R., and M. Hühn. 1987. Studies on estimation of phenotypic stability: tests of significance for nonparametric measures of phenotypic stability. *Biometrics* 43: 45–53. [39]

Oakey, H., A. P. Verbyla, B. R. Cullis, X. Wei, and W. S. Pitchford. 2007. Joint modeling of additive and non-additive (genetic line) effects in multi-environment trails. *Theor. Appl. Genet.* 114: 1319–1332. [39]

Oman, S. D. 1991. Multiplicative effects in mixed model analysis of variance. *Biometrika* 78: 729–739. [39]

Ortiz, R., W. W. Wagoire, J. Hill, S. Chandra, S. Madsen, and O. Stølen. 2001. Heritability of and correlations among genotype-by-environment stability statistics for grain yield in bread wheat. *Theor. Appl. Genet.* 103: 469–474. [39]

Panter, D. M., and F. L. Allen. 1995. Using best linear unbiased predictions to enhance breeding for yield in soybean: I. Choosing parents. *Crop Sci.* 35: 397–405. [39]

Patterson, H. D., and V. Silvey. 1980. Statutory and recommended list trails of crop varieties in the United Kingdom. *J. Royal Stat Sco, A* 143: 219–252. [39]

Perkins, J. M. 1972. The principal component analysis of genotype-environmental interactions and physcial measures of the environment. *Heredity* 29: 51–70. [39]

Perkins, J. M., and J. L. Jinks. 1968. Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity* 23: 339–356. [39]

Pham, H. N. and M. S. Kang. 1988. Interrelationships among and repeatability of several stability statistics estimated from international maize trials. *Crop Sci.* 28: 925–928. [39]

Peipho, H.-P. 1993. Use of the maximum likelihood method in the analysis of phenotypic stability. *Biom. J.* 35: 815–822. [39]

Peipho, H.-P. 1994. Best linear unbaised prediction (BLUP) for regional yield trails: a comparison to additive main effects and multiplicative interaction (AMMI) analysis. *Theor. Appl. Genet.* 89: 647–654. [39]

Piepho, H. P. 1995a. Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trails. *Theor. Appl. Genet.* 90: 438–443. [39]

Piepho, H. P. 1995b. Implications of correlations among some common stability statistics – a Monte Carlo simulations. *Theor. Appl. Genet.* 90: 457–461. [39]

Piepho, H. P. 1996. Analysis of genotype-by-environment interactions and phenotypic stability. *In* M.S. Kang and H.G. Gauch Jr. (eds) *Genotype-by-environment interaction: New perspectives*, pp. 151–174. CRC Press, Boca Raton, Florida. [39]

Piepho, H.-P. 1997a. Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* 53: 761–766. [39]

Piepho, H.-P. 1997b. Mixed models with multiplicative terms for cultivar trail data. *In* P. Krajewski and Z. Kaczmarek (eds), *Proedd. 10th meeting of EUCARPIA section in Biometrics*, pp./ 215–220. [39]

Piepho, H.-P. 1998a. Methods for comparing the yield stability of cropping systems. *J. Agron. Crop Sci.* 180: 193–213. [39]

Piepho, H.-P. 1998b. Empirical best linear unbiased prediction in cultivar trails using factor-analytic variance-covariance structures. *Theor. Appl. Genet.* 97: 195–201. [39]

Peipho, H.-P. and J. Möhring. 2005. Best linear unbaised predicion of cultivar effects for subdivided target regions. *Crop Sci.* 45: 1151–1159. [39]

Peipho, H.-P. and J. Möhring. 2006. Selection in cultivar trails — is it ignorable? *Crop Sci.* 46: 192–201. [39]

Piepho, H.-P., and J. Möhring. 2007. Computing heritability and selection response form unbalanced plant breeding trails. *Genetics* 177: 1881–1888. [39]

Piepho, H.-P., J. Möhring, A. E. Melchinger, and A. Büchse. 2008. BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161: 209–228. [39]

Pinthus, M. J. 1973. Estimate of genotypic value: a proposed method. *Euphytica* 22: 121–123. [39]

Plaisted, P. L., and L. C. Peterson. 1959. A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *Amer. Potato J.* 36: 381–385. [39]

Pritts, M., and J. Luby. 1990. Stability indices for horticultural crops. *HortScience* 25: 740–745. [39]

Qiao, C. G., K. E. Basford, I. H. DeLacy, and M. Cooper. 2004. Advantage of single-trail models for response to selection in what breeding multi-environment trails. *Theor. Appl. Genet.* 108: 1256–1264. [39]

Römer, T. 1917. Sind die ertragreichen Sorten ertragssicherer? *Mitteilungen der DLG* 32: 87–89. [39]

Russell, W. A., and S. A. Eberhart. 1968. Testcrosses of one- and two-ear types of corn belt maize inbreds. II. stability of performance in difference environments. *Crop Sci.* 8: 248–251. [39]

Sabaghnia, N., H. Dehghani, and S. H. Sabaghpour. 2006 Nonparametric methods for interpreting genotype x environment interaction of lentil genotypes. *Crop Sci.* 46: 1100–1106. [39]

Salmon, S. C. 1951. Analysis of variance and longtime variety test of wheat. *Agron. J.* 43: 562–570. [39]

Schnell, F. W. 1967. Die methoden zur erfassung der phänotypsichen stabilität von zuchtsorten. *Arbeitsgemeinschafter Biometrie DLG-Pflanzenzuchtabteilung, Rundschreiben* 1/1967: 18-21. [39]

Shukla, G. K. 1972. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29: 237–245. [39]

Simmonds, N. W. 1991. Selection for local adaptation in plant breeding programme. *Theor. Appl. Genet.* 82: 363–367. [39]

Skrøppa, T. 1984. A critical evaluation of methods available to estimate genotype $\times$ environment interaction. *Studia Forestalia Suecica* 166: 3–14. [39]

Smith, A., B. Cullis, and A. Gilmour. 2001. The analysis of crop variety evaluation data in Australia. *Aust. N. Z. J. Stat.* 43: 129–145. [39]

Smith, A., B. Cullis, D Luckett, G. Hollamby, and R. Thompson. 2002b. Exploring variety-environment data using random effects AMMI models with adjustments for spatial field trend: Part 2: applications. *In* M. S. Kang (ed) *Quantitative genetics, genomics, and plant breeding*, pp. 337–351. CABI, Wallingford. [39]

Smith, A., B. Cullis, and R. Thompson. 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57: 1138–1147. [39]

Smith, A., B. Cullis, and R. Thompson. 2002a. Exploring variety-environment data using random effects AMMI models with adjustments for spatial field trend: Part 1: theory. *In* M. S. Kang (ed) *Quantitative genetics, genomics, and plant breeding*, pp. 323–335. CABI, Wallingford. [39]

Smith, A., B. 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. [39]

Snoad, B., and A. E. Arthur. 1976. The use of regression techniques for predicting the response of peas to environment. *Theor. Appl. Genet.* 47: 9–19. [39]

Stringfeild, G. H., and R. M. Salter. 1934. Differential response of corn varieties to fertility levels and to seasons. *J. Agric. Res.* 49: 991–1000. [39]

Tai, G. C. C. 1971. Genotypic stability analysis and its application to potato regional trails. *Crop Sci.* 11: 184–190. [39]

Utz, H. F. 1972. Die zerlegung der genotype x umwelt - interaktionen. *EDV Medizin Biologie* 3: 52–59. [39]

van Eeuwijk, F. A. 1992. Interpreting genotype-by-environment interaction using redundancy analysis. *Theor. Appl. Genet.* 85: 89–100. [39]

van Eeuwijk, F. A. 1995. Linear and bilinear models for the analysis of multi-environment trails: I. An inventory of models. *Euphytica* 84: 1–7. [39]

van Eeuwijk, F. A., M. Cooper, I. H. DeLacy, S. Ceccarelli, and S. Grando. 2001. Some vocabulary and grammar for the analysis of multi-environemntal trials, as applied to the analysis of FPB and PPB trails. *Euphytica* 122: 477–490. [39]

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 Jr. (eds) *Genotype-by-environment interaction: New perspectives*, pp. 15-49. CRC Press, Boca Raton, Florida. [39]

van Eeuwijk, F. A., and A. Elgeersma. 1993. Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. *Heredity* 70: 447–457. [39]

van Eeuwijk, F. A., L. C. P. Keizer, and J. J. Bakker. 1995. Linear and bilinear models for the analysis of multi-environment trails: II. An application to data from the Dutch maize variety trails. *Euphytica* 84: 9–22. [39]

Verma, M. M., and G. S. Chahal. 1978. Limitations of conventional regression analysis: a proposed modification. *Theor. Appl. Genet.* 53: 89–91. [39]

Weber, W. E., and T. Westermann. 1994. Prediction of yield for specific locations in German winter-wheat trails. *Plant Breeding* 113: 99–105. [39]

Weber, W. E., and G. Wricke. 1987. Wie zuverlässig sind schätzungen von parametern der phänotypischen stabilität? *Vortr. Pflanzenzüchtg* 12: 120–133. [39]

Weber, W. E., G. Wricke, and T. Westermann. 1997. Selection of genotypes and prediction of performance by analyzing genotype-by-environmnet interactions. *In* M.S. Kang and H.G. Gauch Jr. (eds) *Genotype-by-environment interaction: New perspectives*, pp. 353–371. CRC Press, Boca Raton, Florida. [39]

Westcott, B. 1986. Some methods of analysing genotype-environment interaction. *Heredity* 56: 243–253. [39]

Williams, E. J. 1952. The interpretation of interactions in factorial experiments. *Biometrika* 39: 65–81. [39]

Wood, J. T. 1976. The use of environmental variables in the interpretation of genotype-environment interaction. *Heredity* 37: 1–7. [39]

Wricke, G. 1962. Über eine methode zur erfassung der ökologischen streubreite in feldversuchen. *Z. Pflanzenzüchtg* 47: 92–96. [39]

Wricke, G. 1964. Zur berechnung der ökovalenz bein sommerweizen und hafer. *Z. Pflanzenzüchtg* 52: 127–138. [39]

Wricke, G., and W. E. Weber. 1980. Erweiterte analyse von wechselwirkungen in versuchsserien. *In* W. Köpcke and K. Überla (eds) *Biometrie — heute und morgen*, pp. 87–95. Springer. Berlin. [39]

Wright, A. J. 1976. Bias in the estimation of regression coefficients in the analysis of genotype-environment interaction. *Heredity* 37: 299–303. [39]

Yan, W., P. L. Cornelius, J. Crossa, and L. A. Hunt. 2001. Two types of GGE biplots for analyzing multi-environment trail data. *Crop Sci.* 41: 656–663. [39]

Yan, W., L. A. Hunt, O. Sheng, and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40: 597–605. [39]

Yan, W. and M. S. Kang. 2003. *GGE biplot analysis: A graphical tool for breeders, geneticists, and aggronomists.* CRC Press. [39]

Yan, W., M. S. Kang, B. Ma, S. Woods, and P. L. Cornelius. 2007. GGE biplots vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* 643–655. [39]

Yan, W., and I. Rajcan. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Sci.* 42: 11–20. [39]

Yan, W., and N. A. Tinker. 2006. Biplot analysis of multi-environment trail data: principles and applications. *Can. J. Plant Sci.* 86: 623–645. [39]

Yang, R.-C., J. Crossa, P. Cornelius, and J. Burgueno. 2009. Biplot analysis of genotype x environment interaction: proceed with caution. *Crop Sci.* **IN PRESS**. [39]

Yates, F., and W. G. Cochran. 1938. The analysis of groups of experiments. *J. Agricult. Sci.* 28: 556–580. [39]

Yin, X., P. C. Struik, and M. J. Kropff. 2004. Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Sci.* 9: 462–432. [39]