



Managing Genotype \times Environment Interaction in Plant Breeding Programs: A Selection Theory Approach

G.N. Atlin^{1*}, K. Kleinknecht², K.P. Singh³ and H.P. Piepho²

¹*CIMMYT, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico*

²*Bioinformatics Unit, Institute of Crop Science, University of Hohenheim,
Fruwirthstrasse 23, 70599 Stuttgart, Germany*

³*Directorate of Maize Research, Pusa Campus, New Delhi-110012, India*

Received 05 February 2011; Revised 11 April 2011; Accepted 18 April 2011

SUMMARY

Two forms of genotype \times environment interaction (GEI) are of concern to plant breeders. One consists of fixed GEI associated with predictable environmental, geographical, or management factors that can be used to delineate a target population of environments (TPE) for cultivar development and testing. The other consists of random and unexplained rank changes among trials within the TPE which are not associated with any known factor. These two types of GEI must be managed differently by plant breeding programs; fixed GEI is managed by developing or identifying cultivars with adaptation to the specific fixed factor causing the interaction, while random GEI is a noise stratum that is managed through wide-scale testing that adequately samples environmental variation in the TPE, and through the use of best linear unbiased prediction (BLUP). There is substantial evidence that fixed GEI is of limited importance within well-designed TPE. Management of GEI in cultivar development programs, and the estimation of means from multi-environment trials with appropriate measures of precision (METs) has been hampered by the widespread use of inappropriate models that designate trials or trial locations as fixed effects in the combined analysis of cultivar testing data, resulting in unnecessary division of TPEs, identification of putative patterns of adaptation that are not repeated in subsequent testing, and over-estimation of the precision of entry means in multi-environment trials. Mixed model approaches to testing the relative importance of fixed and random GEI in METs are presented.

Keywords : Genotype \times environment interaction, BLUP, Mixed models, Cultivar development genetic correlation, Adaptation.

INTRODUCTION

Genotype \times environment interaction (GEI) is considered to be a critical problem in the development of crop varieties, and much effort has been devoted to the development of methods for the description and measurement of GEI in variety trials. It is often argued that selection for local adaptation, often with farmer participation, is required for the development of cultivars optimized to particular locales and production systems (e.g. Ceccarelli 1989). However, modern crop cultivars developed by seed companies, international

crop research centers, and large national breeding programs often exhibit very wide geographical adaptation, as well as broad adaptation to different management practices and the range of climate variation that affects individual sites over years. For example, the popular rice varieties *Swarna* and IR64 have each been grown for many years on many millions of hectares in several Asian countries, and the CIMMYT maize inbred line CML312 has contributed to hybrids throughout the Latin American and African subtropics (Braun *et al.* 2010). In the case of wheat, several varieties have consistently occupied large areas

* *Corresponding author* : G.N. Atlin
E-mail address : g.atlin@cgiar.org

of production on several continents. Green Revolution wheat cultivars like Siete Cerros (also named Mexipak and Kalyansona) were grown on millions of hectares from North Africa to South Asia, and selections from the CIMMYT cross Veery were released in more than 40 countries. Broad adaptation and high levels of stress tolerance are consistently achieved by large modern plant breeding programs that test germplasm widely. This seems to be evidence that GEI, in the form of narrow cultivar adaptation to a specific set of environmental conditions, is less pervasive and important than is commonly assumed. In this review, it is argued that “fixed” GEI, in the sense of narrow adaptation of elite cultivars to a particular set of environmental conditions, is of limited importance if target environments are well-defined, and that commonly-used GEI analyses consistently overemphasize and overestimate the importance of local adaptation because of the use of inappropriate fixed-effect models that confound individual trials in a multi-environment trial (MET) set with locations, while at the same time underestimating the importance of random trial-to-trial noise in the estimation of genotype means. Appropriate and predictive mixed and random models for assessing genotypic effects and dealing with genotype \times environment interaction in a practical cultivar development context are presented, based on theoretical approaches, developed by Falconer (1952) and applied to plant breeding by Atlin and Frey (1990), that treat the expression of traits such as yield in different environment as correlated traits. The main purpose of this article is to provide breeders with guidance on managing fixed and random GEI in their own breeding and testing programs, with a view to maximizing genetic gains.

THE TARGET POPULATION OF ENVIRONMENTS (TPE) AND FIXED VERSUS RANDOM GEI

The breeding target, or *target population of environments* (TPE) (Comstock 1997) is a key concept in the organization of breeding programs, and is very useful in thinking about and managing GEI. The TPE is the set of fields and future seasons in which the varieties produced by a breeding program will be grown. The TPE is not a single environment, but a variable set of future production environments that must be sampled by the breeding program that serves it. TPE are delineated by some combination of fixed and

repeatable environmental factors that influence crop adaptation, and that are common to all the fields and future seasons in the TPE, or that occur within the TPE at a predictable frequency.

These factors can include geographical region, soil type, daylength, average temperature, and average rainfall, among many others. These factors, which may or may not contribute to changes in relative cultivar performance in different trials, are fixed attributes of a location in which trials are conducted. Some, like temperature or rainfall, vary from year to year, but their long-term averages can be used to characterize environments. Such factors are fixed in the statistical sense, i.e., we wish to make inferences about them and their interaction with genotype; they are not included in the design or analysis of METS as an error sampling stratum. Fixed environmental or geographic factors can be used to group trial sites into relatively uniform TPE. However, even if two trial sites differ in some fixed environmental factor, it is not necessarily the case that a set of genotypes will have different ranks at the two sites, or that they need to be treated as separate breeding targets. Fixed environmental factors, at most, provide hypotheses that must be tested when breeders delineate their TPE. For example, if, within the mandated region of responsibility of a particular breeding program, there exist two distinct soil types, it is not necessarily true that each soil type requires a different, specifically-adapted cultivar, although it is possible that they do. This is a hypothesis that must be tested by evaluating a fairly large set of genotypes at several locations within each soil type, and testing whether the interaction between genotypes and soil types is greater than the random variation among trials within soil types.

In general, a TPE is delineated such that consistent and repeatable genotype rank changes associated with locations or management systems do not occur within it in association with some fixed environmental factor. If they do, these locations do not belong in the same TPE. Another way to put this is that there should not be consistent and repeatable **fixed GEI** (GEI associated with a fixed environmental factor such as location or soil type) within a TPE. This does not mean that there will not be differences in ranks among entries in breeding trials conducted at different locations within the TPE, but rather that these differences are primarily random and ephemeral, and are not consistently associated with particular locations, changing from trial

to trial and year to year. These unpredictable and non-recurring rank changes among entries over trials within the TPE constitute **random GEI**. Random GEI is more properly considered a form of sampling variance, akin to plot-to-plot variation in a field trial, but occurring at a different level. Of course, improved biological understanding may result in an explanation of some factor causing variation in ranking among trials. When this occurs, the factor may be introduced as a fixed effect into the modeling of cultivar response across environments, and may be used in a refined delineation of the TPE.

There is great confusion among both plant breeders and biometricians with respect to fixed versus random GEI, resulting in the use of inappropriate fixed effects models in situations where mixed or random models are more appropriate. If fixed GEI is large within a TPE, then it may be appropriate to subdivide it into two or more smaller TPE. However, if no fixed factor can explain the GEI observed in a MET series, then the observed random GEI needs to be treated as a sampling error. Random GEI is managed similarly to the management of plot error in field trials: with appropriate sampling and adequate replication. Random environmental fluctuations from trial to trial need to be treated as experimental noise that can reduce the accuracy of selection. Trial-to-trial variation within a TPE is always present. Even in an extremely uniform region, two trials of the same cultivar set planted a few kilometers apart, or in the same spot in different years, will normally exhibit some differences in the rankings of cultivars due to annual fluctuations in rainfall totals and distribution, disease pressure, planting date, etc. Random GEI is an error stratum that can contribute as much or more “noise” to estimates of cultivar means as the plot error or residual stratum. Mistaking random for fixed GEI can lead to serious errors in breeding program design, notably the inappropriate splitting of regions into smaller TPE when there is really no repeatable difference in cultivar ranking between them.

It should be understood that, although the TPE normally has defined geographical limits, it is not necessary that one geographic region be treated as a single TPE. For example, within a given region, there may be great differences in the cultivars needed by farmers with access to irrigation, as opposed to farmers whose fields are purely rainfed. In such a situation, it is reasonable to delineate two TPE in the same region,

and perhaps operate two breeding and testing programs (resources permitting), one for rainfed farmers and one for irrigated farmers.

THE MANAGEMENT OF FIXED GEI IN BREEDING PROGRAMS

The principal approach to the management by breeders of fixed and repeatable GEI associated with environmental or management factors is to “exploit” it by developing cultivars that are specifically adapted to the particular environmental condition causing the interaction. In other words, one can subdivide the TPE into more homogeneous breeding targets. This is equivalent to “shifting” GE effects to genotypic effects, and may increase genetic variance and broad-sense heritability (H). To take a simple example, short-duration cultivars have an advantage in environments where growth is terminated at the end of the crop cycle by drought, heat, or cold, but long-duration cultivars are higher-yielding in environments where the length of the growing season does not prematurely terminate the crop cycle. In such cases, if the growing season length is reasonably consistent at individual locations, it makes sense to target long-duration cultivars at sites with a longer growing season, and short-duration cultivars at sites with a short season, i.e., to subdivide the TPE. However, great care should be taken in deciding whether or not to divide a TPE to ensure that subdivision is truly warranted, because it usually results in a reduction of testing resources within the new, smaller TPE. The trade-offs involved are discussed in detail by Atlin *et al.* (2000a,b) and Curnow (1988), and elaborated below.

Some environmental factors may not appear to fall neatly into the fixed versus random schema, but still should be considered fixed factors for the delineation of TPE. For example, drought events occur unpredictably; it is not possible to forecast droughts with enough certainty to determine whether a drought-tolerant variety will give an advantage at a particular location in a given season. However, within any given TPE the breeder usually has an idea about how frequently severe droughts occur. For example, if severe droughts occur in one season out of four, the breeder can make an informed decision about whether it is worthwhile to screen materials under development for drought tolerance, and how much weight to give to the results of such screening in selection decisions. In this

sense, drought-proneness is more like a fixed than a random effect. Sporadically-occurring stress factors such as drought may be best dealt with by selecting for them in simulated managed-stress screening systems, wherein an analogue to the naturally-occurring stress is applied artificially, for example (in the case of drought) by withholding irrigation during a rain-free period. Managed stress screening for drought tolerance has been incorporated in the cultivar development pipelines of many breeding programs, and has proven effective for improving drought tolerance of rice (Verulkar *et al.* 2010) and maize (Bänziger *et al.* 2006). The weight that should be given to data from such screens in selection decisions, relative to weighting of trials that randomly sample the TPE, is a question that has been inadequately researched and requires urgent attention. Managed stress screening data are often characterized by low repeatability, and care needs to be taken not to weight them excessively. Too much weight placed on screens with low broad-sense heritability or repeatability (H) relative to high-H data from wide-scale testing under normal conditions can reduce gains in both stressed and non-stressed environments.

THE MANAGEMENT OF RANDOM GEI IN BREEDING PROGRAMS

Random GEI resulting from site-to-site variations in cultivar rankings that change from year to year cannot be managed by subdivision of the TPE, or by any form of site-specific prediction, because it does not result from consistent site effects. Random GEI is a true error stratum, like plot error, and must be handled by breeders in the same way they handle other sources of noise that obscure the value of genotypes under test. The principal tool for reducing the effect of random GEI is replication over sites and years, such that the unexplained environmental variation within the TPE is adequately sampled, and precise and repeatable estimates of genotype means are obtained, with realistic standard errors. Increasing or decreasing the number of sites and years of testing will affect H, standard errors, and selection response. Variance component estimates may be used to model these effects to determine the optimal allocation of testing resources (for an example, see Atlin and McRae (1994)).

Another key tool for managing random GEI (and random plot error) in breeding programs is the use of *best linear unbiased prediction (BLUP)* to account for

selection bias. The purpose of multi-location variety evaluation is almost always to *select* the best cultivars for further testing or for recommendation to farmers, based on mean yields estimated in a set of METs that sample the TPE. The process of ranking and selection of the highest-yielding entries, either for further testing or for recommendation to farmers, introduces *selection bias*, a problem that has been widely recognized in animal breeding, and for which appropriate analytical approaches have been developed (Henderson 1976), but which is little recognized or understood in plant breeding. Selection bias arises because the highest-yielding entries in a cultivar trial are usually high-yielding both because of large genotypic deviations or effects, which can recur in subsequent testing and in farmers' fields, and because they also fortuitously include larger-than-average random environmental deviations, which were specific to the trials in which they were measured and unlikely to be repeated in future testing. BLUP is an estimation method that treats genotypic effects as random (Smith *et al.* 2005), and shrinks them towards the mean by a factor which is closely related to H; in essence it uses population variance component estimates to generate a repeatability value specific to each entry (i.e. adjusted for the number of sites, years, and replicates in which the entry was tested) and shrinks the effect estimate back towards the population mean in proportion to this repeatability. Means with a higher information content (i.e. means tested over more sites and years) are subject to less shrinkage than means with a lower information content (i.e. means from testing over fewer sites and years). BLUP is an important tool for dealing with the imbalance inherent in most cultivar testing systems, wherein new entries are added and poor-performing entries dropped each year, and only a fraction of entries are tested over two years or more. BLUP also permits information from other regions or cropping systems, if available, to be used in improving predictions of performance in the TPE. Piepho and Möhring (2005) pointed out that genotype performance is often correlated in different TPE, and showed that, when appropriately weighted, information from different but similar regions can usually improve predictions of performance in the TPE. There are many reports of high correlations among genotypic effects estimated in widely-separated TPE (Atlin *et al.* 2000b, Braun *et al.* 1992, Cooper *et al.* 1993a, Peterson and Pfeiffer 1989). A practical illustration of use of information

generated outside the TPE is provided by Cooper *et al.* (1993b).

PARTITIONING FIXED VERSUS RANDOM GEI: THE GE MODEL

The importance of GEI caused by a fixed environmental or geographic effect, relative to the importance of within-TPE noise caused by random and unexplainable variation among genotypic effects across locations is easily assessed by incorporating a fixed factor into the model for the combined analysis of a MET series. Consider the simplest GEI model (referred to herein as the “GE model” for the analysis of a series of trials:

$$Y_{ijkl} = M + E_i + R(E)_{j(i)} + G_k + GE_{ik} + e_{ijkl} \quad (1)$$

where

Y_{ijkl} = the measurement on plot l in environment i , block j , containing genotype k ;

M = the overall mean of all plots in all environments;

E_i = the effect of environment (trial) i ;

$R(E)_{j(i)}$ = the effect of replicate j within environment i ;

G_k = the effect of genotype k ;

GE_{ik} = the interaction of genotype i with environment k ;

e_{ijkl} = the plot residual.

This model assumes that trials are laid out in randomized complete blocks and that error variance is constant across trials. This assumption simplifies the subsequent arguments, but it can be modified, for example, to account for the use of resolvable incomplete block designs and heterogeneity of error variance.

In this model, the variance of the estimate of a genotype or entry mean is

$$\sigma_{\bar{Y}}^2 = \frac{\sigma_{GE}^2}{e} + \frac{\sigma_e^2}{er} \quad (2)$$

where $\sigma_{\bar{Y}}^2$ is the variance of a cultivar mean, e is the number of trials, and r is the number of replicates per trial.

The repeatability or broad-sense heritability (H) of estimates of genotype or entry means is

$$H = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GE}^2}{e} + \frac{\sigma_e^2}{re}} \quad (3)$$

with effects associated with variances defined as in Eq. 1, and e and r as in Eq. 2. If the environments can be classified into two subgroups, such as soil types or regions, then the environmental effect E_i can be subdivided into an effect of subgroups, S_m , and trials within subgroups, $E_i(S_m)$ (Atlin *et al.* 2000a). The GE effect is similarly partitioned into an effect GS_{mk} , the interaction of genotypes with environmental subgroup, and an effect $GE(S)_{ki(m)}$, or the interaction of genotypes with trials within environmental subgroups

$$Y_{ijkl} = M + S_m + E_i(S_m) + R(E(S))_{j(i(m))} + G_k + GS_{mk} + GE(S)_{ki(m)} + e_{ijklm} \quad (4)$$

Environmental subgroups are considered to be fixed effects, because specific inferences are to be made about the subgroups, but trials within subgroups are simply samples of the environmental variation within the subgroups, and therefore must be considered to be random effects. Genotypes may be considered fixed or random, but should generally be considered random if they are being selected for recommendation or further testing (Smith *et al.* 2005).

To illustrate the utility of this model, we will consider the CIMMYT-managed Southern African early and intermediate maize hybrid trials conducted in 2009. These trials consisted of 63 hybrids evaluated at 24 sites in three-replicate trials. The trials were located in 6 countries. To determine if each country should be treated as a separate TPE, countries were included as a factor in the model, by partitioning site effects and their interactions with genotypes into fixed effects of countries and random effects of locations within countries. Results are presented in Table 1. The genotype x site variance was larger than the genotypic variance, but over 80% of this GEI was due to genotype x site interaction within countries, and only 20% among countries. Atlin *et al.* (2000a) showed that the genetic correlation between means estimated in an undivided TPE and a subdivision (r'_G) can be expressed as

$$r'_G = \sqrt{\frac{\sigma_G^2}{\sigma_G^2 + \sigma_{GS}^2}} \quad (5)$$

This is based on a specific assumption on the random effects, i.e. all have constant variance. This is

Table 1. Genotype, genotype \times country, genotype \times location within country, and error variances for the 2009 Southern African CIMMYT early and intermediate hybrid trials.

Source	Variance component
Hybrid	0.100
Hybrid \times location	0.125
Hybrid \times country	0.030
Hybrid \times location within country	0.125
Error	0.518

a very simplistic model, which could be modified to account for heterogeneity. For example, the variance of GS could be allowed to vary between TPE.

By this method, the average genetic correlation between yields in any one Southern African country and the mean of all countries is 0.88. Given this high genetic correlation, and the fact that many more test sites are available in the region as a whole than in any one country, it is likely that genetic gains would be greater if the Southern African region were treated as the TPE than if each country made separate testing decisions. The effect of subdividing a target region on response to selection can be expressed as

$$CR/DR = r'_G \sqrt{\frac{H_W}{H_S}} \quad (6)$$

where CR is the predicted correlated response in the subregion to selection in the undivided TPE, DR is direct response to selection in the subregion, and H_W and H_S are the repeatabilities in the whole TPE and the subregions, respectively (Atlin *et al.* 2000b). H_W is estimated as in Eq. (3), while H_S is estimated as

$$H = \frac{\sigma_G^2 + \sigma_{GS}^2}{\sigma_G^2 + \sigma_{GS}^2 + \frac{\sigma_{GE(S)}^2}{e} + \frac{\sigma_E^2}{re}} \quad (7)$$

Substituting into Eq. (6) and assuming that four trials would be conducted in each country whereas 24 would be conducted in the region as a whole, it is found that gains in each country would be, on average, 4% higher if selections were made based on performance across the whole region, rather than based on data from each country alone.

PARTITIONING FIXED VERSUS RANDOM GEI : THE GLY MODEL

Eq. (1) is widely used for analyzing METs, but it should be noted that it is a flawed and misleading model, because it confounds genotype \times year, genotype \times location, and genotype \times year \times location effects in an undifferentiated “genotype \times environment” effect, resulting in a great loss of information and frequent over-estimation of the importance of location effects. In general, this model should not be used for the analysis of METs unless it is for the analysis of a series of trials conducted in a single year for the purpose of estimating a more realistic standard error than the one derived from a single-site analysis.

A more realistic and complete model for the analysis of cultivar trials than Eq. (1) recognizes years and sites as random factors used in sampling the TPE. The resulting model is

$$Y_{ijklm} = M + Y_i + L_j + YL_{ij} + R(YL)_{k(ij)} + G_l + GL_{li} + GY_{lj} + GLY_{lij} + e_{ijklm} \quad (8)$$

This model (referred to herein as the “GLY” model) differs from the GE model in that the E term has been partitioned into location (L) and year (Y) effects and their interaction (LY). Similarly, the GE term has been partitioned into GL, GY, and GLY components. H for this model is

$$H = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GL}^2}{l} + \frac{\sigma_{GY}^2}{y} + \frac{\sigma_{GLY}^2}{ly} + \frac{\sigma_E^2}{rly}} \quad (9)$$

where l , y , and r are the number of locations, years, and replicates, respectively. In this model, location effects may be taken as fixed, if they refer to a specific subregion within the TPE rather than to specific trials, with the same locations used in the trials over years. However, if locations are included simply to randomly sample fields within the TPE, they are best considered random. Year effects, and their interactions with locations and genotypes, are clearly random.

It is important for breeders and biometricians to clearly understand the relationship between the GE and the GLY model, and the implication of this relationship for estimates of genotype and genotype \times environment means and variances. In the GE model, if the trials are

Table 2. Variance component estimates from variety testing programs in several species and target populations of environments (TPE)

Crop	TPE	σ_G^2	σ_{GL}^2	σ_{GY}^2	σ_{GLY}^2	σ_E^2	Source
Maize hybrids	Southern Africa: rainfed	0.17	0	0.03	0.51	0.29	CIMMYT, unpublished data
Rainfed rice	Thailand	0.07	0.04	0.06	0.32	0.54	Cooper <i>et al.</i> (1999)
Barley	Eastern Canada	0.17	0.08	0.05	0.17	0.53	Atlin and McRae (1994)
Spring wheat	Eastern Canada	0.48	0.01	0	0.27	0.25	Atlin and McRae (1994)
Winter wheat	Eastern Canada	0.36	0.03	0.02	0.29	0.30	Atlin and McRae (1994)
Spring wheat	Western Canada	0.29	0.11	0.02	0.27	0.31	Baker (1969)
Barley	UK	0.10	0.06	0.12	0.27	0.45	Talbot (1984)
Spring wheat	UK	0.13	0.12	0.14	0.28	0.33	Talbot (1984)
Winter wheat	UK	0.27	0.02	0.06	0.31	0.35	Talbot (1984)
Winter wheat	Italy	0.18	0.07	0.04	0.29	0.41	Annicchiarico and Perezin (1994)
Spring wheat	Australia	0.05	0.13	0.11	0.12	0.58	M. Cooper, personal communication
Mean		0.20	0.06	0.06	0.28	0.39	

conducted over more than one year, estimates of genotypic effects are not confounded by GEI effects, and genotypic variance estimates are unbiased by GEI variances. However, because GL, GY, and GLY effects and variances are not estimated separately by the GE model, it is impossible to determine if GEI is caused by random noise (GY and GLY effects) or fixed adaptation to specific locations or subregions (GL effects). In general, for TPE that are delineated by reasonable criteria, the random GLY component is usually the largest GEI component, and little repeatable adaptation to specific locations is observed. Evidence for this statement is found in the variance components derived from 12 MET series conducted over a range of crop species in both tropical and temperate environments and presented in Table 2, standardized to a phenotypic variance of 1.0. On average, the GLY variance resulting from unexplained changes in genotype rank from site to site and year to year was nearly five times as large as the genotype \times location variance.

The very large size of the GLY effects, which are truly random, relative to the GL effects that may in some cases be considered to be fixed attributes of particular genotypes in particular environments, is a potential source of misinterpretation in the analysis of METs using the GE model. This is particularly true when analyses are conducted based on fixed-effect

models such as AMMI (Gauch 2006) or GGE (Yan *et al.* 2007). Yang *et al.* (2009) showed that treating random GEI effects as fixed causes lack of repeatability and over-interpretation of “which-won-where” patterns observed in GGE and AMMI biplots, and Yang (2007) showed that it resulted in the declaration of excess crossover interactions in COI models. If the MET series is from a single year of testing, the danger of misinterpretation is especially great. In single-year analyses, estimates of genotypic effects are confounded with genotype \times year effects, and estimates of genotypic variances are biased upward by genotype \times year variance. GE variances estimated from single-year analyses contain both the GL and GLY variance. Because of the large size of the GLY variance relative to the GL variance in most TPE, analyses of single-year MET series are prone to misinterpretations wherein large GE effects are assumed to be caused by local adaptation to specific locations, whereas they really result from random year-to-year and site-to-site variability causing non-repeatability of rankings from different trials. These errors and misinterpretations can be largely avoided by estimating variance components from the GLY model and determining if there is evidence of large genotype by site effects before proceeding to analyses designed to group environments based on interaction patterns. Also, using BLUP based on a mixed model version of AMMI or GGE, known as factor-analytic models (Piepho 1997, 1998), will

result in reduced selection bias in the discovery of interaction patterns.

Why has the GE model remained in wide use, while the GLY model is rarely used? In the past, breeders and biometricians often used the GE model in place of the GLY model because it was less demanding of computer memory, but this reason is no longer valid; most breeding programs have access to adequate computing power to analyze METs with appropriate models, and residual maximum likelihood (REML) algorithms capable of analyzing unbalanced datasets with mixed models are available in statistical packages such as SAS, GENSTAT, and R. Another reason may be the widespread availability of AMMI and GGE biplot software that does not allow the partitioning of location and year effects, and that many breeders use as a “black-box” approach to the analysis of METs (Yang *et al.* 2009). Neither of these reasons justifies using an inappropriate model which leads breeders to make wrong decisions.

As in the GE model, the site effects in the GLY model, and their interactions with genotypes, may be partitioned according to some fixed environmental factor. The resulting model is

$$Y_{ijklm} = M + Y_i + S_j + L(S)_{jk} + YS_{ij} + YSL_{ijk} + R(YSL)_{l(ijk)} + G_m + GY_{mi} + GS_{mj} + GSL_{mjk} + GYS_{mij} + GYSL_{mijk} + e_{ijklmn} \quad (10)$$

where S is the effect of the subdivision of locations, $L(S)$ the effect of locations within subregions, with interactions with genotypes similarly partitioned. The explanatory power of the fixed factor in accounting for GEI can then be quantified and tested. As an example, we consider the power of the current geographical recommendation zones of the All-India Coordinated Maize Improvement Program (AICMIP) maize hybrid trials to explain GEI in sets of new hybrids tested India; these trials are used to provide information for the release of new hybrids. The maize-growing regions of the country are divided into five large geographical zones, with a total of 25-30 test locations distributed among them annually. New sets of hybrids, segregated by maturity group (extra-early, early, medium, and late) are composed annually in these trials and evaluated over all five zones in their first year of testing; hybrids that perform well in the first year are then retested in one or more of the five zones for up to two more years. A common set of check hybrids is tested in each zone.

In a combined analysis conducted for each set occurring in the period 1994-2006, the location effects and their interactions with hybrids were partitioned into zones and sites within zones. Variance component estimates pooled over the combined analyses are presented in Table 3 for the late and medium hybrid maturity groups. These results indicate that, although in aggregate GEI variances were large relative to the genetic variance the cultivar \times zone interaction explained almost none of the GEI within each maturity group, and that the cultivar \times location \times year interaction within zones was the largest variance component. The limited proportion of the GEI explained by the large geographical recommendation zones indicates that environmental variability within each zone is nearly as great as the variation among zones, and that it may be possible to select hybrids within each maturity group that are broadly adapted across zones. The limited hybrid \times

Table 3. The magnitude of cultivar and cultivar \times zone variances relative to other components of genotype \times environment interaction estimated from the All-India Coordinated Maize Improvement Program medium and late maturity hybrid trials conducted in India 1994-2006.

LATE MATURITY		
Effect	Variance component	Standard error
Cultivar	229862	36873
Cultivar \times Zone	0	.
Cultivar \times Location within Zone	263903	44706
Cultivar \times Year within Zone	0	.
Cultivar \times Year \times Location within Zone	735292	44554
MEDIUM MATURITY		
Effect	Variance component	Standard error
Cultivar	313310	28960
Cultivar \times Zone	365	9482
Cultivar \times Location within Zone	78019	18254
Cultivar \times Year within Zone	42527	14687
Cultivar \times Year \times Location within Zone	656587	22777

zone interaction also indicates that information from neighbouring zones may be used to increase the precision of hybrid means estimated within zones as outlined by Piepho and Möhring (2005).

CONSEQUENCES OF INAPPROPRIATELY DESIGNATING SITE, YEAR, OR TRIAL EFFECTS AS FIXED ON ESTIMATES OF PRECISION

An important consequence of wrongly designating location, year, or trial effects as fixed rather than random in combined analyses conducted for the purposes of selection within the TPE is that the standard error of the difference (SED) between two cultivar means, and related statistics such as the least significant difference (LSD), are greatly underestimated, leading to a false idea of the precision of estimation of entry means. To illustrate this, the standard error of the difference for the GSY model with sites and years considered fixed (Eq. 11) and random (Eq. 12) are presented below.

$$SED_{fixed} = \sqrt{2 \left(\frac{\sigma_e^2}{lyr} \right)} \quad (11)$$

$$SED_{random} = \sqrt{2 \left(\frac{\sigma_{GL}^2}{l} + \frac{\sigma_{GY}^2}{y} + \frac{\sigma_{GLY}^2}{ly} + \frac{\sigma_e^2}{lyr} \right)} \quad (12)$$

For the model that incorrectly designates these effects as fixed, the GL, GY, and GLY variances do not enter in the standard error of the difference. Since the purpose of METs, however, is to predict the mean future performance of cultivars in the TPE, these variances affect the accuracy of predictions. Excluding

them from the SED gives an extremely overoptimistic impression of the precision of the means from cultivar trials, with the extent of underestimation of SED dependent on the size of the GEI variances relative to the plot residual variance. This is illustrated in Table 4, using variance component estimates from Thai rainfed lowland rice trials reported by Cooper *et al.* (1999). For this MET series, designation of sites and years as fixed effects resulted in underestimation of SEDs by 44 to 52%, depending on the extent of testing.

IMPLICATIONS OF LIMITED FIXED GEI FOR BREEDING STRATEGIES

Selection for broad adaptation and stress tolerance has been achieved via different strategies in different breeding programs, but the most important tool has been the extensive field testing of experimental breeding lines and hybrids in many environments during the selection process. Increases in tolerance to a range of stresses such as drought, low fertility, and cold in US Corn Belt maize, for example, have contributed greatly to yield gains (Castleberry *et al.* 1984; Duvick and Cassman 1997; Tollenaar and Lee 2002) but have not resulted from direct selection for any of these stresses. Rather, they have resulted mainly from the broad-scale multi-location hybrid testing programs of commercial maize breeding companies that effectively sample conditions occurring in farmers' fields. The small size of the genotype \times location interaction variance relative to other components of GEI reported in many studies (Table 2) is strong evidence that this approach is sound, and that selection based on wide-scale testing is likely to outperform selection for local adaptation unless fixed genotype \times location

Table 4. Predicted standard error of the difference (SED) ($t \text{ ha}^{-1}$) estimates from three-replicate Thai rainfed lowland rice variety trials (Cooper *et al.* (1999)) assuming locations and years fixed or random

No. of years	No. of locations	SED assuming locations and years fixed	SED assuming locations and years random	% underestimation of SED by fixed-effects model
1	1	0.77	1.38	44
1	5	0.34	0.67	49
1	10	0.24	0.50	52
2	1	0.54	0.96	43
2	5	0.24	0.45	47
2	10	0.17	0.33	49

interaction is very large (Atlin *et al.* 2001). This is consistent with the breeding and commercialization strategies of seed companies and large public-sector breeding programs that deliver most cultivars of staple food crops to farmers around the world; to be cost-effective, these breeding programs must serve very large TPE. The best approach to the management of GEI in breeding programs is to (a) design the TPE such that fixed GEI within it is limited or non-existent (b) sample the TPE adequately to estimate genotypic effects with high repeatability, and (c) if available, use information from relevant trials outside the TPE to increase the precision with which means are estimated in the TPE.

CONCLUSIONS

METs, which are expensive to conduct, are important tools in the selection of new cultivars, and should be analyzed so as to extract the maximum information possible from the investment. Random GEI is an important but under-recognized error stratum in the analysis and interpretation of METs. However most METs are analyzed using inappropriate fixed-effect models that are highly susceptible to selection bias, grossly underestimate standard errors of means and differences, and overestimate heritability and the importance of repeatable genotype \times environment interaction. Breeders and statisticians usually fail to use the optimal BLUP algorithm, which corrects for selection bias and permits the analysis of unbalanced data sets over years and locations, permitting new entries with only one or a few years of evaluation to be compared with long-term checks without discarding data on the long-term entries. It is important that both local breeding programs and national cultivar testing systems keep pace with developments in statistical software and algorithms, and adopt modern approaches to data analysis. Cultivar trial series should be analyzed using the GLY model rather than the GE model even when unbalanced; the REML algorithm available in widely-used statistical packages makes this possible for all plant breeding programs. In general, the importance of highly localized adaptation is overestimated by plant breeders, largely as a result of the use of inappropriate fixed effects models that deliver results that are contradicted by the practical experience of the world's largest and most effective public and private sector breeding programs. Selection in METs conducted

across enough sites and years to adequately sample environmental variation in the TPE remains the most reliable way to develop high-yielding and broadly-adapted crop cultivars.

REFERENCES

- Annicchiarico, P. and Perezin, M. (1994). Adaptation patterns and definitions of macro-environments for selection and recommendation of common-wheat genotypes in Italy. *Plant Breeding*, **113**, 197-205.
- Atlin, G.N. and Frey, K.J. (1990). Selecting oat lines for yield in low-productivity environments. *Crop Sci.*, **30**, 556-561.
- Atlin, G.N. and McRae, K.B. (1994). Resource allocation in Maritime cereal cultivar trials. *Can. J. Plant Sci.*, **74**, 501-505.
- Atlin, G.N., Baker, R.J., Lu, X. and McRae, K.B. (2000a). Selection response in subdivided target regions. *Crop Sci.*, **40**, 7-13.
- Atlin, G.N., Lu, X. and McRae, K.B. (2000b). Genotype \times region interaction for yield in two-row barley in Canada. *Crop Sci.*, **40**, 1-6.
- Atlin, G.N., Cooper, M. and Bjornstad, A. (2001). A comparison of formal and participatory breeding approaches using selection theory. *Euphytica*, **122**, 463-464.
- Baker, R.J. (1969). Genotype-environment interactions in yield of wheat. *Can. J. Plant Sci.*, **49**, 743-751.
- Bänziger, M., Setimela, P.S., Hodson, D. and Vivek, B. (2006). Breeding for improved abiotic stress tolerance in maize adapted to southern Africa. *Agric. Water Manage.*, **80**, 212-224.
- Braun, H.J., Atlin, G.N. and Payne, T. (2010). Multi-location testing as a tool to identify plant response to global climate change. pp. 115-138. In: Reynolds, MP, ed. *Climate Change and Crop Production*. CABI Climate Change Series I. CAB International, Wallingford, UK.
- Braun, H.J., Pfeiffer, W.H. and Pollmer, W.G. (1992). Environments for selecting widely adapted spring wheat. *Crop Sci.*, **32**, 1420-1427.
- Castleberry, R.M., Crum, C.W. and Krull, F. (1984). Genetic yield improvement of U.S. maize cultivars under varying fertility and climatic environments. *Crop Sci.*, **24**, 33-36.
- Ceccarelli, S. (1989). Wide adaptation: How wide? *Euphytica*, **40**, 197-205.

- Comstock, R.E. (1977). Quantitative genetics and the design of breeding programs. pp. 705-718. In : Pollack E., Kempthorne O., Bailey Jr T.B. (eds), *Proceedings of the International Conference on Quantitative Genetics*, August 16-21, 1976, Iowa State University Press, Ames.
- Cooper, M., Byth, D.E., DeLacy, I.H. and Woodruff, D.R. (1993a). Predicting grain yield in Australian environments using data from CIMMYT international wheat performance trials. 1. Potential for exploiting correlated response to selection. *Field Crops Res.*, **32**, 305-322.
- Cooper, M., Delacy, I.H., Byth, D.E. and Woodruff, D.R. (1993b). Predicting grain yield in Australian environments using data from CIMMYT international wheat performance trials. 2. The application of classification to identify environmental relationships which exploit correlated response to selection. *Field Crops Res.*, **32**, 323-342.
- Cooper, M., Rajatasereekul, S., Immark, S., Fukai, S. and Basnayake, J. (1999). Rainfed lowland rice breeding strategies for northeast Thailand. I. Genotypic variation and genotype \times environment interactions for grain yield. *Field Crops Res.*, **64**, 131-151.
- Curnow, R.N. (1988). The use of correlated information on treatment effects when selecting the best treatment. *Biometrika*, **75**, 287-293.
- Duvick, D.N. and Cassman, K.G. (1999). Post-green revolution trends of temperate maize in the North-Central United States. *Crop Sci.*, **39**, 1622-1630.
- Falconer, D.S. (1952). The problem of environment and selection. *Am. Nat.*, **86**, 293-298.
- Henderson, C.R. (1976). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, **31**, 423-447.
- Gauch, H.G. (2006). Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.*, **46**, 1488-1500.
- Gauch, H.G. and Zobel, R.W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.*, **37**, 311-326.
- Peterson, C.J. and Pfeiffer, W.H. (1989). Breeding widely adapted, popular maize hybrids evaluation : Relationships among test sites based on cultivar performance. *Crop Sci.*, **29**, 276-282.
- Piepho, H.P. (1997). Analyzing genotype-environment data by mixed models with multiplicative effects. *Biometrics*, **53**, 761-766.
- Piepho, H.P. (1998). Empirical best linear unbiased prediction in cultivar trials using factor analytic variance-covariance structures. *Theo. Appl. Genetics*, **97**, 195-201.
- Piepho, H.P. and Möhring, J. (2005). Best linear unbiased prediction of cultivar effects for subdivided target regions. *Crop Sci.*, **45**, 1151-1159.
- Smith, A.B., Cullis, B.R. and Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. *J. Agric. Sci.*, **143**, 449-462.
- Talbot M. (1984). Yield variability of crop varieties in the UK. *J. Agric. Sci.*, (Camb.) **102**, 315-321.
- Tollenaar, M. and Lee, E.A. (2002). Yield potential, yield stability and stress tolerance in maize. *Field Crops Res.*, **75**, 161-169.
- Verulkar, S.B., Mandal, N.P., Dwivedi, J.L., Singh, B.N., Sinha, P.K., Mahato, R.N., Dongre, P., Singh, O.N., Bose, L.K., Swain, P., Robin, S., Chandrababu, R., Senthil, S., Jain, A., Shashidhar, H.E., Hittalmani, S., Vera Cruz, C., Paris, T., Raman, A., Haefele, S., Serraj, R., Atlin, G. and Kumar, A. (2010). Breeding resilient and productive genotypes adapted to drought-prone rainfed ecosystem of India. *Field Crops Res.*, **117**, 197-208.
- Yan, W., Kang, M.S., Ma, B., Woods, S. and Cornelius, P.L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.*, **47**, 643-655.
- Yang, R.C. (2007). Mixed-model analysis of crossover genotype-environment Interactions. *Crop Sci.*, **47**, 1051-1062.
- Yang, R.C., Crossa, J., Cornelius, P.L. and Burgueño, J. (2009). Biplot analysis of genotype \times environment interaction: Proceed with caution. *Crop Sci.*, **49**, 1564-1576.