# ACCOMMODATION OF GENOTYPE X ENVIRONMENT INTERACTIONS IN NATIONAL GENETIC EVALUATION PROGRAMS

## D.R. NOTTER

Animal Genetics and Breeding Unit[*]
University of New England
Armidale, NSW 2351.

## INTRODUCTION

Genotype x environment interactions (GEI) have been extensively studied and discussed but to date have not been explicitly included in any national genetic evaluation program for meat animals. Livestock breeders in geographically diverse areas such as Australia and North America often express concern over possible effects of GEI and important region by genotype interactions within North American Hereford cattle have been experimentally demonstrated (Burns et al. 1979; Parish et al. 1983, 1985). Yet documentation and quantification of GEI in large field data sets has been difficult, in part because of limitations in experimental design inherent in field data analyses and in part because of limitations in analytical procedures. This paper will discuss aspects of the definition, prediction and importance of GEI effects.

## DEFINITION OF ENVIRONMENTS AND OF GEI

At a practical level, GEI is usually acknowledged when candidates for selection can be shown to rank differently in true breeding value for some trait in different environments, implying that different animals will be selected in the different environments. At the gene level, we can consider N loci that may affect some trait over all possible environments. For the jth environment, $n_j$ loci actually have a measurable additive effect, $a_{jk}$, for locus k in environment j. At this level, then, GEI may alternatively be defined as a change in the number of loci affecting the trait or as a change in relative magnitude of the $a_{jk}$'s.

Accurate definition of the environment is a non-trivial problem in accommodating GEI. In general, an environment can be thought of as any factor that systematically affects expression of a trait. Environments can be defined in terms of climate, management, feeding level or a combination of these factors that may define environments unique to a region or even a single herd. Genotype x genotype interactions may also exist as a special case of GEI in which the breeding value (BV) of an animal may depend upon the genetic composition of the animal's mates. Thus beef bulls will not necessarily rank the same when bred to temperate beef, dairy or Zebu-cross cows.

Practically, incorporation of GEI into national genetic evaluations requires all animals to be unambiguously assigned to one of the environments. This is often more difficult than it appears. How much concentrate is required to differentiate feedlot from forage finishing, or is a 1/4-Zebu, 3/4-Hereford animal a 'temperate beef' or 'tropical crossbred' cow? Regional distinctions are unambiguous but may not be particularly useful, even if they are based on ecological rather than political boundaries. Seedstock herds are often managed to minimize environmental effects, which may mask genotype x regions effects acting in customers' commercial herds. Also, GEI can only be considered in selection when data from the various environments are regularly available. Often, the most important GEI may be in the commercial sector, where performance data are often not collected for use in genetic evaluation programs.

---

Several situations can produce apparent GEI but not affect the actual ranking of individuals for specific traits. Dickerson (1962) and Yamada (1962) noted that changes in mean and variance across environments may produce such apparent GEI. Methods for handling heterogeneous variances without resorting to GEI models have been discussed by Gianola (1986), Quaas et al. (1989) and Boldman and Freeman (1990). GEI for individual traits should also be clearly differentiated from GEI for the selection objective. Across environments, it is to be expected that the weightings assigned to individual traits in the selection objective or in some resulting selection index (Ponzoni and Newman, 1989) will differ. This situation necessarily results in GEI (re-ranking) at the level of the selection objective but does not in itself require consideration of GEI to predict BV for component traits. In particular, composite traits often proposed as 'natural' indices (e.g., lean growth rate; Simm et al. 1987) may be particularly liable to GEI.

## MODELLING GEI

In developing models for BV prediction in the presence of GEI, it is useful to discriminate between fixed and random environments. In practical terms, fixed environments may be thought of as regularly occurring, unambiguous environments with imperfect genetic correlation $r_{jj'}$ between pairs of environments. The objective of a GEI analysis with fixed environments would generally be estimation of BV for candidates for selection in each environment, with the potential to base selection decisions on BV specific to a given environment. In contrast, random environments are those that are less clearly repeatable (such as years) and that in general do not provide opportunity for environment-specific breeding programs. Instead, the goal is prediction of mean BV across environments, but after accounting for random GEI.

Fixed environments

In most cases, models to allow BV prediction in the presence of genotype x fixed environment interactions will utilize multiple-trait prediction methods (Henderson and Quaas 1976; Schaeffer 1984; Thompson and Meyer 1986) which consider performance in the different environments to be separate traits and which require pairwise estimates of genetic correlations among environments. The general multiple-trait mixed model prediction equations (Schaeffer 1984) for an individual animal model are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z+G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}Y \\ Z'R^{-1}Y \end{bmatrix}$$

where Y is a vector of data for all traits (environments); $\hat{\beta}$ and $\hat{u}$ are estimates of fixed and additive genetic animal effects, respectively; X and Z are incidence matrices relating observations to $\hat{\beta}$ and $\hat{u}$, respectively; R is a residual variance-covariance matrix of the observations; and G is the genetic additive variance-covariance matrix which includes both additive relationships among animals and genetic variances and covariances among traits (environments). Several characteristics of these general equations are unique to BV estimation in multiple environments. If animals have records in only one environment, $R^{-1}$ will usually be diagonal and, depending on heterogenity of variance, can perhaps be factored out of the equations. Fixed effects ($\beta$) will generally be environment-specific, such that X is block diagonal when observations are ordered within environment.

The main problem in solving these equations is that the number of equations usually increases in proportion to the number of traits (environments). However, this expansion assumes that BV are required for all animals in all environments. This assumption is generally made in multiple-trait prediction, in part because it allows $G^{-1}$, which can be very large, to be calculated efficiently as the direct product of the inverses of the numerator relationship matrix and the genetic variance-covariance matrix among traits (Henderson 1975, 1976). For

multiple environments, BV estimates in all environments may be required only for animals with descendants in multiple environments or who are otherwise candidates for selection across environments. BV for most dams and many progeny across environments would not be required. If environments are geographically distinct, the number of 'connecting' animals with descendants in multiple environments may be modest, allowing a fractional, rather than multiplicative, increase in the number of equations to be solved. This approach requires calculation of $G^{-1}$ for an incomplete G matrix. Fortunately, Elzo (1989), in a recent discussion of calculation of $G^{-1}$ in multi-breed designs, also gives rules for calculating $G^{-1}$ when BV are not estimated for all traits and all animals. Thus BV estimation with GEI may not require a multiplicative increase in the number of equations to be solved. Also, improved multiple-trait computing algorithms (Tier and Graser 1990; Tier 1991) appear to greatly facilitate BV estimation across environments.

Random environments

The GEI model with random environments may be written as:

$$Y = X\beta + Zu + Wi + e$$

where i is a vector of GEI constants, u is now a vector of mean BV and W is an incidence matrix relating Y and i. Mixed model equations may be written as:

$$
\begin{bmatrix}
X'X & X'Z & X'W \\
Z'X & Z'Z + A^{-1}\lambda_1 & Z'W \\
W'X & W'Z & W'W + A_w^{-1}\lambda_2
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\
\hat{u} \\
\hat{i}
\end{bmatrix}
=
\begin{bmatrix}
X'Y \\
Z'Y \\
W'Y
\end{bmatrix}
$$

where A is the overall numerator relationship matrix, $A_w$ is a block-diagonal relationship matrix with blocks corresponding to within-environment relationship matrices, $\lambda_1$ and $\lambda_2$ are the ratios of residual to additive genetic and genotype x environment interaction variances, respectively and $R^{-1}$ has been assumed factored from the equations. In most cases, i is considered a nuisance vector required for optimum prediction of u. This model has been discussed in the context of sire evaluation and with $A_w = I$ (i.e., no covariance among $i_j$'s for related sires) by Meyer (1987) who also pointed out that genotype (sire) x environment (herd) effects are usually considered indistinguishable from other, nongenetic sources of resemblance among half-sibs such as preferential treatment or unreported penning together of sire progenies. Discrimination between genetic and nongenetic covariances among half-sibs in the same environment might be achieved by comparing $\hat{i}_j$ values with $A_w = I$ for related sires with progeny in the same environment but could be biased if progeny of related sires received correlated levels of preferential treatment.

## IMPACT OF GEI

Fixed environments

The impact of GEI on sire evaluation was assessed for the case of two environments by comparing three alternative indices of sire BV in environment 1: I, the optimum index which weighted progeny records from the two environments in accord with their genetic correlation, r; I*, an index weighting all progeny records equally (i.e., assuming r=1.0); and $I_1$, an index using only data from environment 1. Results are expressed in terms of correlations of I* or $I_1$ with I and as such quantify the selection response expected from these non-

optimal indexes relative to that expected from selection on I. Results were calculated for heritability ($h^2$) of .1, .2 and .4, for r=.3, .5, .7 and .9, and for progeny numbers in environments 1 ($n_1$) and 2 ($n_2$) of 2, 8, 32, 128 and 512.

Table 1  Relative efficiencies of sire breeding value estimation using data from two environments when progeny from both environments are weighted equally ($I^*$) or when only progeny from environment 1 are used ($I_1$). The goal is breeding value estimation in environment 1 and efficiency is relative to an optimum index for a range of genetic correlations (r) and progeny numbers in environments 1 ($n_1$) and 2 ($n_2$) and for heritability = .2

| | | $n_2$ | | | | | | | | | |
| | | 2 | | 8 | | 32 | | 128 | | 512 | |
| $n_1$ | r | $I^*$ | $I_1$ | $I^*$ | $I_1$ | $I^*$ | $I_1$ | $I^*$ | $I_1$ | $I^*$ | $I_1$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | .3 | .87 | .97 | .79 | .90 | .74 | .82 | .73 | .77 | .73 | .76 |
| | .5 | .94 | .91 | .92 | .78 | .91 | .65 | .91 | .59 | .91 | .57 |
| | .7 | .98 | .84 | .98 | .66 | .98 | .52 | .98 | .46 | .97 | .44 |
| | .9 | 1.00 | .77 | 1.00 | .57 | 1.00 | .42 | 1.00 | .37 | 1.00 | .35 |
| 8 | .3 | .96 | .99 | .86 | .98 | .69 | .96 | .57 | .94 | .53 | .93 |
| | .5 | .98 | .98 | .93 | .94 | .85 | .89 | .79 | .85 | .77 | .84 |
| | .7 | .99 | .96 | .98 | .89 | .95 | .80 | .93 | .74 | .92 | .72 |
| | .9 | 1.00 | .94 | 1.00 | .84 | .99 | .71 | .99 | .63 | .99 | .61 |
| 32 | .3 | .99 | 1.00 | .97 | 1.00 | .84 | .99 | .58 | .99 | .44 | .99 |
| | .5 | 1.00 | 1.00 | .98 | .99 | .91 | .98 | .75 | .97 | .66 | .97 |
| | .7 | 1.00 | .99 | .99 | .98 | .96 | .96 | .89 | .94 | .84 | .93 |
| | .9 | 1.00 | .99 | 1.00 | .97 | .99 | .93 | .98 | .88 | .98 | .86 |
| 128 | .3 | 1.00 | 1.00 | 1.00 | 1.00 | .97 | 1.00 | .82 | 1.00 | .52 | 1.00 |
| | .5 | 1.00 | 1.00 | 1.00 | 1.00 | .98 | 1.00 | .88 | 1.00 | .69 | 1.00 |
| | .7 | 1.00 | 1.00 | 1.00 | 1.00 | .99 | 1.00 | .94 | .99 | .84 | .99 |
| | .9 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .99 | .99 | .98 | .96 | .98 |
| 512 | .3 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .98 | 1.00 | .81 | 1.00 |
| | .5 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .98 | 1.00 | .87 | 1.00 |
| | .7 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .94 | 1.00 | .94 | 1.00 |
| | .9 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .98 | 1.00 |

Results for $h^2$=.2 are shown in Table 1. In general, for $n_1 \geq 8 > n_2$, genetic progress exceeded 94% of maximum for both $I^*$ and $I_1$. Among remaining entries, two situations will be highlighted. The first is for sires that are accumulating progeny simultaneously in both environments (i.e., down the diagonal of Table 1). In this situation, $I^*$ is initially substantially better than $I_1$ unless r<.5. However, as progeny accumulate, this situation gradually reverses until at $n_1=n_2=128$, $I_1$ is superior to $I^*$ unless r=.9. At $n_1=n_2=512$, $I_1$ is always superior to $I^*$, with the advantage declining as r increases, but remaining at 6% for r=.7. This pattern indicates that as $n_1$ and $n_2$ increase simultaneously, the value of information from environment 2 declines, until, for large $n_1$, the failure to properly dilute information from environment 2 in accord with its imperfect r results in a substantial loss in accuracy for $I^*$.

The second situation is when a proven sire from environment 2 is introduced into environment 1 (i.e., when $n_1$ increases for $n_2=512$). Initially, one cannot afford to discard the large volume of data from environment 2; it is better to overweight it than to ignore it unless r is quite low. However, unless r approaches .9, by the time a few tens of progeny have been accumulated in environment 1, the marginal value of progeny from environment 2 has declined and $I_1$ approaches I.

Results were generally similar for $h^2=.1$ and .4. For all r and $h^2$, $I_1$ was at least 94% as efficient as I by the time $n_1$ reached 128. $I_1$ was at least 93% as efficient as I at $n_1=32$ unless r is quite high, $h^2$ is low and $n_2>n_1$. Thus by the time several tens of progeny have been recorded, there is little to be lost by considering only data from environment 1, suggesting that environment-specific genetic evaluations of proven sires may be reasonable, even if r is .7 to .8. In contrast, evaluation of young sires would suffer greatly if data from all environments are not used, even if r is only .3 to .5.

Random environments

The impact of GEI with random environments has been described by Meyer (1987) in terms of the reduction in effective progeny numbers associated with sire x herd environment interactions. Basically, as the sire x herd interaction variance increases, the value of additional progeny within a single environment declines relative to the value of additional progeny in other environments. Thus in the presence of sire x herd interaction the maximum effective progeny number asymptotically approaches $1/\gamma$ for any herd where $\gamma$ is the ratio of the sire x herd to residual (phenotypic minus sire minus sire x herd) variance.

CONCLUSIONS

Accommodation of genotype x environment interactions in national genetic evaluation programs will likely be most severely limited by inability to unambiguously define environmental factors, to estimate necessary genetic correlations across environments, and to collect data from commercial herds where GEI may be most important. Inclusion of GEI into BV estimation programs should not be undertaken lightly, since reporting of multiple BV for some or all animals will complicate reporting and merchandising and promote subdivision of the population. However, evaluation of GEI using designed experiments and improved analytical procedures should continue, since substantial improvements in rate of genetic changes can occur through proper consideration of documented cases of GEI.

ACKNOWLEDGEMENTS

REFERENCES

BOLDMAN, K.G. and FREEMAN, A.E. (1990) J. Dairy Sci. 73:503.
BURNS, W.C., KOGER, M., BUTTS, W.T., PANISH, O.F. and BLACKWELL, R.L. (1979) J. Anim. Sci. 49:403.
DICKERSON, G.E. (1962) Anim. Prod. 4:47
ELZO, M.A. (1989) J. Anim. Sci. 68:1215.
GIANOLA, D. (1986) Theor. Appl. Genet. 72:671.
HENDERSON, C.R. (1975) J. Dairy Sci. 58:1727.
HENDERSON, C.R. (1976) Biometrics 32:69.
HENDERSON C.R. and QUAAS, R.L. (1976) J. Anim. Sci. 43:1188.
MEYER, K. (1987) Livest. Prod. Sci. 17:95.

PANISH, O.F., KOGER, M., URICK, J.J., BURNS, W.C., BUTTS, W.T. and RICHARDSON, G.C. (1983) J. Anim. Sci. 56:1039.

PANISH, I.F., URICK, J.J., BURNS, W.C., BUTTS, W.T., KOGER, M. and BLACKWELL, R.L. (1985) J. Anim. Sci. 61:1146.

PONZONI, R.W. and NEWMAN, S.(1989) Anim. Prod. 49:35.

QUAAS, R.L., GARRICK, D.J. and MCELHENNEY, W.H. (1989) J. Anim. Sci. 67:2529.

SCHAEFFER, L.R. (1984) J. Dairy Sci. 67:1567.

SIMM, G., SMITH, C. and THOMPSON, R. (1987) Anim. Prod. 45:307.

THOMPSON, R. and MEYER, K. (1986) Livest. Prod. Sci. 15:299.

TIER, B. (1991) J. Anim. Breed. Genet. (In press).

TIER, B. and GRASER, H.-U. (1990) J. Anim. Breed. Genet. (In press).

YAMADA, Y. (1962) Japanese J. Genet. 37:498.