Theor Appl Genet (1994) 89:647–654

H.-P. Piepho

Best Linear Unbiased Prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interaction (AMMI) analysis

Received: 15 July 1993 / Accepted: 18 March 1994

Abstract Multilocation trials are often used to analyse the adaptability of genotypes in different environments and to find for each environment the genotype that is best adapted; i.e. that is highest yielding in that environment. For this purpose, it is of interest to obtain a reliable estimate of the mean yield of a cultivar in a given environment. This article compares two different statistical estimation procedures for this task: the Additive Main Effects and Multiplicative Interaction (AMMI) analysis and Best Linear Unbiased Prediction (BLUP). A modification of a cross validation procedure commonly used with AMMI is suggested for trials that are laid out as a randomized complete block design. The use of these procedure is exemplified using five faba bean datasets from German registration trails. BLUP was found to outperform AMMI in four of five faba bean datasets.

Key words Two-way classification · Genotype × environment interaction · Additive main effects multiplicative interaction · Best linear unbiased prediction · Predictive accuracy · Cross validation

Introduction

The evaluation of new cultivars is usually done by testing a set of cultivars in different environments. Such multilocation trials are essential for obtaining well-adapted cultivars. Basically, one is interested in identifying the genotypes best adapted for each particular environment. To extract a maximum of information from the data we seek a "best" estimate of yield. The most common estimate is the arithmetic mean of a genotype across replicates in an environment. This mean is often called the cell mean. It has the virtue of simplicity but the shortcoming of not fully exploiting all information contained in the complete genotype × environment dataset.

Gauch (1988, 1992) has advocated the use of what he terms AMMI (Additive Main Effects Multiplicative Interaction) analysis of multilocation trial data. In many cases, this procedure has been shown to increase estimation accuracy by separating the pattern from the noise in the residuals of the additive model (Gauch 1988, 1992).

In the AMMI model both genotypes and environments are regarded as fixed. When the number of genotypes is large, however, modelling genotypic effects as random may be preferable despite the fact that it would be classified as fixed using traditional definitions (Stroup and Mulitze 1991). In the case of random genotypic effects, the assessment of the mean yield of a genotype in a certain environment may be viewed as a problem of prediction rather than one of estimation (see Searle et al. 1992, p 18). Random genotypes also imply random genotype-environmental interaction, so the prediction of yield involves prediction of a genotype's random interaction with a specific environment. The prediction of the outcome of random variables is commonly done by Best Linear Unbiased Prediction (BLUP), as originally suggested by Henderson (1975). BLUP of additive main effects for yield trials has been investigated by Hill and Rosenberger (1985) and Stroup and Mulitze (1991).

The purpose of this paper is to propose a BLUP procedure that includes a prediction of genotype-environmental interaction. The accuracy of BLUP and AMMI are compared based on the root mean square prediction difference (RMSPD) suggested by Gauch and Zobel (1988). Gauch and Zobel's procedure for computing the RMSPD is appropriate for a completely randomized (CR) design. In this paper the procedure is slightly modified for application to a randomized complete block (RCB) design.

Communicated by A. R. Hallauer

H.-P. Piepho
Faculty of Agriculture, University of Kassel, Steinstrasse 19, 37213 Witzenhausen, Germany

Methods

We assume the following linear model for a two-way classification with interaction:

$$y_{ij} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + e_{ij} \quad (i = 1,\ldots,K; \quad j = 1,\ldots,N) \tag{1}$$

y_{ij} = mean of i-th genotype in j-th environment
μ = overall mean
α_i = effect of i-th genotype
$(\alpha\tau)_{ij}$ = interaction of i-th genotype with j-th environment
e_{ij} = error of mean of i-th genotype in j-th environment

It should be noted that if we employ a RCB design, the environmental effect τ_j may be written as $\tau_j = \tau'_j + b_j$, where τ'_j is a "pure" environmental effect and b_j is the mean of the effects of all blocks laid out in the j-th environment. Throughout this paper, we will assume that $b_j = 0$, whence $\tau_j = \tau'_j$, which implies that block effects in a given environment are fixed and sum up to zero. The assumption of fixed blocks is quite common in agricultural research (Kempthorne 1952, p 152; Steel and Torrie 1980, p 218f). In order to make a decision as to whether blocks are fixed or random, it is helpful to answer the following question: may blocks be regarded as a random sample of a larger population of blocks, and if so, what are the geographical boundaries of this population of blocks? In most cases, if blocks are regarded as a sample from a larger population, it is seldom appropriate to assume that the sample is a random one. We will therefore consider blocks as fixed, which means that the population of blocks, and hence the environment under consideration, is confined to those blocks that were actually included in the experiment.

For the two procedures to be described, it is assumed that all effects (except error) are fixed. The mean y_{ij} of genotype i in environment j is the arithmetic mean of R replicates:

$$y_{ij} = \textstyle\sum_r y_{ijr}/R, \quad (r = 1,\ldots,R) \tag{2}$$

The simplest approach is to take the arithmetic mean y_{ij}; i.e. the cell mean as an estimate of the yield of genotype i in environment j. The arithmetic mean is a Best Linear Unbiased Estimator (BLUE) of $\mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$ (Searle 1987, p 81). Gauch (1988) has produced evidence that an improved estimate can be obtained by AMMI. The principle of AMMI is to first fit additive main effects for genotypes and environments by an ordinary ANOVA procedure and then to apply principle component analysis (PCA) to the matrix of residuals that remain after the fitting of main effects (for details see Gauch 1992, p 85f.). The interaction plus mean error $(\alpha\tau)_{ij} + e_{ij}$ can be decomposed into S PCA axes:

$$y_{ij} = \mu + \alpha_i + \tau_j + \sum_{s=1}^{S} l_s a_{is} t_{js} + \theta_{ij} \tag{3}$$

where l_s is the singular value for PCA axis s, a_{is} the genotype eigenvector for axis s and t_{js} is the environment eigenvector. Eigenvectors are scaled as unit vectors. A residual θ_{ij} remains if not all axes are used. There are at most $\min(K-1, N-1)$ axes. Depending on the number of PCA axes retained, the models are denoted as AMMIO, AMMI1,..., AMMIF. With AMMIO, no PCA axis is fitted, while with AMMIF the full model i.e. the cell means model is used. The latter is the same as the ordinary arithmetic mean in Eq. 2.

AMMI may be viewed as a procedure to separate pattern [i.e. the interaction effect $(\alpha\tau)_{ij}$] from noise (i.e. the mean error e_{ij}). This is achieved by PCA, were the first axes (i.e. the axes with the largest eigenvalues), recover most of the pattern, while most of the noise ends up in later axes. In what follows we will present a different approach to achieve the same goal. The basic idea is to estimate the effects in the linear model and then to weight some or all of the effects by an estimate of the pattern-to-noise ratio associated with the respective effect. It must be emphasized that this approach is limited to linear models, while AMMI includes multiplicative terms.

Up to now we have assumed that all effects, including genotypes, are fixed. However, one may also take the view that genotypes are random (Stroup and Mulitze, 1991). If genotypes are random μ and e_j are fixed, while α_i, $(\alpha\tau)_{ij}$ and e_{ij} are randomly distributed with zero mean and with variances σ^2_α, $\sigma^2_{\alpha\tau}$ and σ^2, respectively. Clearly, we then have a mixed model. It should be emphasized that by regarding genotypes as random rather than fixed, we do not necessarily eliminate our interest in the yields of specific genotypes. Rather, we now look for estimators of the realized value of random variables instead of estimating fixed effects. In the common statistical terminology (see, e.g. Searle et al. 1992, p 13), estimation of random effects is referred to as prediction; the corresponding procedure for mixed linear models is BLUP.

Whether an effect is most appropriately deemed fixed or random does not so much depend on whether or not we are interested in a particular set of treatment effects, but rather on whether or not the levels of the effect being considered may reasonably be assumed to come from a probability distribution (Searle et al. 1992, p 18). More often than not, it is quite appropriate to assume that genotypes constitute a random sample from a certain population. Stroup and Multize (1991) have emphasized that in variety trials, where the number of treatments (varieties) may be between 20 and 100, BLUP is typically more efficient than BLUE, provided that the distribution of treatment effects is reasonably symmetric. The same authors suggest that modelling an effect as random may be preferable even though it would be regarded as fixed using traditional definitions. Stroup and Mulitze (1991) hold the view that the traditional distinction between fixed and random is not useful and that this distinction may in fact lead the data analyst to choose the less efficient alternative.

We are interested in predicting the "true" yields $w_{ij} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$, given the observed yields y_{ij}. To this end we will apply BLUP as suggested by Henderson (1975) and described in Searle et al. (1992, p 269f.). A derivation of BLUP is given in the Appendix, which shows that the BLUP of w_{ij} can be expressed as

$$\begin{aligned}\mathrm{BLUP}(w_{ij}) = \bar{y}_{\cdot j} &+ \frac{\sigma^2_{\alpha\tau} + N\sigma^2_\alpha}{\sigma^2_{\alpha\tau} + \sigma^2 + N\sigma^2_\alpha}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \\ &+ \frac{\sigma^2_{\alpha\tau}}{\sigma^2_{\alpha\tau} + \sigma^2}(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot} + \bar{y}_{\cdot\cdot})\end{aligned} \tag{4}$$

This expression demonstrates the shrinkage effect associated with BLUP. To see this, consider the following identity:

$$y_{ij} = \bar{y}_{\cdot j} + (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) + (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}) \tag{5}$$

= estimated mean of j-th environment
+ estimated effect of i-th genotype
+ estimated interaction of i-th genotype with j-th environment

The effect of BLUP is to weight the estimated random effects (genotypic and interaction effects) in Eq. 5 by a factor that could be termed "repeatability". From Eq. 4 the BLUP of w_{ij} can be written as

$$\mathrm{BLUP}(w_{ij}) = \bar{y}_{\cdot j} + h^2_g(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) + h^2_{ge}(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}) \tag{6}$$

with

$$h^2_g = \frac{\sigma^2_{\alpha\tau} + N\sigma^2_\alpha}{\sigma^2_{\alpha\tau} + \sigma^2 + N\sigma^2_\alpha} = \text{"repeatability of estimated genetic effect"}$$

$$h^2_{ge} = \frac{\sigma^2_{\alpha\tau}}{\sigma^2_{\alpha\tau} + \sigma^2} = \text{"repeatability of estimated interaction"}$$

It is seen that repeatabilities increase with decreasing error variance σ^2. When σ^2 equals zero the BLUP is identical to the cell mean. Since $1 \geq h^2_g \geq h^2_{ge} \geq 0$, the effect of BLUP is to shrink the estimated genotypic and interaction effects towards their zero mean whenever $\sigma^2 > 0$, which will be the rule. The smaller the repeatabilities, the larger is the shrinkage effect. This shrinkage effect is considered a desirable property of BLUP. As Hill and Rosenberger (1985) pointed out, "intution tells most breeders to suspect a new entry with an exceptionally large or low mean, and the shrinkage property makes an adjustment consistent with the need for caution."

BLUP assumes that all variances (σ_α^2, $\sigma_{\alpha\tau}^2$, and σ^2) are known. This requirement is rarely met in practice, so that we will have to use variance component estimates and accept a loss in efficiency. A recent account of variance component estimation procedures is given in Searle et al. (1992). We will use the ANOVA method to estimate σ_α^2 and $\sigma_{\alpha\tau}^2$. σ^2 may be estimated as

$$\hat{\sigma}^2 = \sum_j s_j^2 / RN,$$

where s_j^2 is the error mean square in the j-th environment. Computation of s_j^2 involves separate analyses of the statistical design in each environment.

In this paper, focus will be on the case that the variance of a mean is the same in each environment. Often, however, the design and the number of replicates are not the same in each environment, implying that the variance of a mean (σ^2) is not necessarily constant across environments (Cochran and Cox 1957, p 553). In this case the BLUP procedure is still applicable. With heterogeneous mean variances, the diagonal elements of V associated with the j-th environment are $\sigma_\alpha^2 + \sigma_{\alpha\tau}^2 + \sigma_j^2$, where σ_j^2 denotes the variance of a mean in the j-th environment. For BLUP, σ_j^2 in V may be appropriately estimated by s_j^2/R_j, where R_j is the number of replicates in the j-th environment. If this is done, one has to use the general formula in Eq. A.2 of the Appendix in place of Eq. 4, which is valid only for balanced data and homogeneous error variances. For estimation of variance components from unbalanced data, see Searle et al. (1992).

When some genotype-environment combinations are missing, BLUPs (using the general formula in Eq. A.2 of the Appendix) may still be obtained for those combinations that were observed. To compute BLUPs from Eq. A.2, we need to estimate the genotypic, interaction and error variances, and this is possible even if some genotype-environment combinations are missing, e.g. by Maximum Likelihood or REstricted Maximum Likelihood methods (see Searle et al. 1992, p 232f).

In the discussion so far we have assumed that genetic and interaction effects are random, while environments are fixed. Following Stroup and Mulitze (1991) we might as well regard all effects as random. It may be shown that for the completely random model the BLUP (w_{ij}) is given by

$$\text{BLUP}(w_{ij}) = \bar{y}.. + h_g^2(\bar{y}_{i\cdot} - \bar{y}..) + h_e^2(\bar{y}_{\cdot j} - \bar{y}..) + h_{ge}^2(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}..)$$

where h_{ge}^2 and h_g^2 are as defined above and

$$h_e^2 = \frac{\sigma_{\alpha\tau}^2 + K\sigma_\tau^2}{\sigma_{\alpha\tau}^2 + \sigma^2 + K\sigma_\tau^2} = \text{"repeatability of estimated environmental effect"}$$

In a similar way we may obtain the BLUP for the model with fixed fixed genotypes and random environments. The following notation will be used for the different BLUPs: BLUP_e = BLUP for genotypes fixed/environments random; BLUP_g = BLUP for genotypes random/environments fixed; and BLUP_{ge} = BLUP for all effects random. It is conjectured that BLUP for the different models do not differ appreciably. The difference between any of the above BLUPs and the cell mean will mainly accrue from the interaction effect. The important feature common to all three BLUPs is that interactions are random. Note that in the cell mean each estimated effect receives a weight of 1, while with BLUP the estimates of random effects receive a weight between 0 and 1, the weights being equal to the repeatabilities as defined above. It is easily seen that h_e^2 and h_g^2 will be close to unity if σ_α^2 and σ_τ^2 are not much smaller than $\sigma_{\alpha\tau}^2$ and σ^2. At the limit h_g^2 appoaches a value of 1 as N tends to infinity. Similarly, h_e^2 tends to unity with increasing K. On the contrary, h_{ge}^2 is independent of K and N. Thus, it will generally not make a big difference whether genotypes or environments, or both are assumed random. What is important is that one of the two is considered random, such that interactions are random. It should be pointed out that ranking of genotypes within environments is always identical for BLUP_g and BLUP_{ge}.

Example

We will now compare the predictive accuracy of AMMI (AMMIO, AMMI1, ..., AMMIF) to that of BLUP, employing, five faba bean datasets from German registration trials (1985–1989). Each of these data sets has four replicates. The number of genotypes and environments is displayed in Table 1. The trials in each environment were laid out as randomized complete block (RCB) design.

Table 2 shows the ANOVA for the five faba bean datasets. In all of the datasets the genotype-environment interaction was significant at the 0.001 level, so the BLUP procedure proposed in this article and AMMI were deemed appropriate. Table 3 shows the estimated repeatabilities of different effects. h_g^2 and h_e^2 are very close to 1 in all datasets, so in the examples considered here, BLUP_e, BLUP_g and BLUP_{ge} yield very similar results.

Predictive accuracy was assessed by randomly splitting the original data into modelling data (three replicates = blocks of each environment) and validation data (the remaining replicate of each environment). The models AMMIO–AMMIF and BLUP were fitted to the modelling data. Then the residual mean square prediction difference (RMSPD) was calculated as the square root of the mean squared difference between model predictions and validation observations (Gauch and Zobel 1988). The procedure was repeated 1000 times

Table 1 Number of genotypes in sugar faba beans data from German registration trials

Year	Number of Genotypes	Number of Environments
1985	14	9
1986	31	9
1987	32	9
1988	35	10
1989	35	10

Table 2 ANOVA mean squares, (MS) for five faba bean datasets

[a] The environment MS and the genotype MS were tested against the interaction MS. In each data set, all effects were significant at the 0.001 level

	1985	1986	1987	1988	1989
Genotype[a]	628.23	678.82	771.64	631.75	522.21
Environment[a]	59 42.14	28 437.23	9374.99	26 530.82	12 104.18
Block	55.33	341.06	127.80	141.16	198.74
G × E	208.32	94.71	83.41	84.78	74.13
Error	15.19	29.61	19.56	15.63	13.38

Table 3 Repeatabilities of different effects in faba bean datasets

Year	h^2_{ge}	h^2_g	h^2_e
1985	0.972	0.976	0.998
1986	0.687	0.956	0.999
1987	0.766	0.975	0.998
1988	0.816	0.975	0.999
1989	0.819	0.974	0.999

using different randomizations, and the result was averaged. Computations were done using the SAS matrix procedure IML (SAS Institute 1990). It is stressed that – for reasons to be discussed in the next section – the data splitting procedure used here is slightly different from that originally proposed by Gauch and Zobel (1988). Moreover, as will also be discussed in the next section, in a strict sense the RMSPD of $BLUP_e$ and $BLUP_{ge}$ are not comparable to the RMSPD of $BLUP_g$ and AMMI. But, since h^2_e is almost equal to unity, the comparison is approximately valid (see next section).

Predictive accuracy as assessed by RMSPD is shown in Table 4. In four of the five datasets BLUP was more accurate than the best AMMI model, while in the 1985 dataset, AMMI5 was identified as the best AMMI model. $BLUP_g$ and $BLUP_{ge}$ yielded the same RMSPD, while $BLUP_e$ had only very slightly larger RMSPD. Thus, the three BLUP methods were essentially equivalent. For all datasets the best AMMI model had many axes. In the 1987 and 1988 datasets, the full model was the best of all AMMI models. This is in contrast to many other results, where AMMI0, AMMI1 or AMMI2 were identified as the most accurate models (see e.g. Gauch 1988, 1992, p 139; Crossa et al. 1990). Obviously, the interaction pattern in the faba bean examples is too complex to be captured by only a few axes. This is also reflected by the sums of squares associated with the different PCA axes (see example of the 1986 faba bean dataset in Table 5). Another reason is the difference in the data splitting procedure compared to that suggested by Gauch and Zobel (1988).

Table 5 Sums of squares, degrees of freedom and mean square for interaction, PCA axes and error of 1986 faba bean dataset

Source	SS	df[a]	MS
Interaction	28411.51	240	94.71***
PCA1	7472.63	37	201.96***
PCA2	6634.81	35	189.57***
PCA3	4463.16	33	135.25***
PCA4	3364.96	31	108.55***
PCA5	2942.72	29	101.47***
PCA6	1756.24	27	65.05**
PCA7	1369.64	25	54.79**
PCA8	407.35	23	17.71[ns]
Error	24783.57	837	29.61

* Significant at $P \leq 0.5$; ** significant at $P \leq 0.01$; *** significant at $P \leq 0.001$; ns = non-significant

[a] Degrees of freedom associated with a PCA are computed as $df = K + N - 1 - 2s$ df for PCA axis s (see Gollob 1968) Note that the Gollob-test is liberal for the first PCAs. For alternatives see Cornelius (1993)

Cross validation for the RCB design

The cross validation procedure by Gauch and Zobel (1988) has been applied to data from RCB designs by many authors (e.g. Crossa et al. 1990, 1991; Nachit et al. 1992). Crossa et al. (1990) used data from trials that were laid out in a RCB design with four blocks (= replicates) in each environment (this is the same design as in the faba bean example). For each treatment (i.e. genotype and environment combination) two replicates were selected at random to be modelled by AMMI, and the other two were reserved as validation observations. Thus, observations from a block were randomly split into modelling and validation data, thus destroying the original block structure. In contrast, the results in Table 4 are based on randomly selecting (three) complete blocks of each environment. In other words, while by the method of Gauch and Zobel (1988) single observations are randomized within treatments, we randomized complete blocks within environments. The implications of this difference will be discussed later in this article. First

Table 4 Predictive accuracy of AMMI and BLUP for five faba bean datasets (1985–1989) based on random splitting complete blocks in each environment

Model	1985	1986	1987	1988	1989
	RMSPD	RMSPD	RMSPD	RMSPD	RMSPD
AMMI0	7.725	7.591	6.195	5.993	5.820
AMMI1	6.469	7.548	6.043	5.628	5.411
AMMI2	5.647	7.348	5.745	5.386	5.208
AMMI3	5.182	7.306	5.659	5.189	5.087
AMMI4	5.012	7.301	5.504	5.080	4.921
AMMI5	4.874[a,b]	7.224	5.554	5.100	4.897
AMMI6	4.915	7.221	5.554	5.121	4.893[b]
AMMI7	4.895	7.194[b]	5.516	5.107	4.893[b]
AMMI8	–	–	–	5.074	4.929
AMMIF	4.915	7.220	5.492[b]	5.041[b]	4.947
$BLUP_g$	4.879	7.019[a]	5.359[a]	4.945[a]	4.864[a]
$BLUP_e$	4.878	7.022	5.360	4.946	4.865
$BLUP_{ge}$	4.879	7.019[a]	5.359[a]	4.945[a]	4.864[a]

[a] Most accurate model
[b] Most accurate AMMI model

consider the result of cross validation for the faba bean data based on randomizing single observations instead of complete blocks (see Table 6). The results indicate that a model selection based on randomizing single observations tends to favour simpler AMMI models than does a selection based on randomizing complete blocks. Only in 1985 was the same AMMI model selected with both randomization procedures.

The tendency to select simpler models is a result of an inflated noise created by not keeping complete blocks together. The increased level of noise contributes to interaction and error in the modelling data. This is best seen by writing down the model for a single observation, which must include b_{jr} effect of the r-th block in the j-th environment:

$$y_{ijr} = \mu + \alpha_i + \tau'_j + b_{jr} + (\alpha\tau)_{ij} + e_{ijr}, \quad (7)$$

where e_{ijr} is the error associated with the r-th replicate of the i-th genotype in the j-th environment. Now assume that there are R blocks in each environment and that we randomly selected M blocks for modelling. The means that the model for the average across these M replicates can be writted as

$$\begin{aligned} y_{ij} &= \mu + \alpha_i + \tau'_j + b_j + (\alpha\tau)_{ij} + e_{ij} \\ &= \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + e_{ij} \end{aligned} \quad (8)$$

with $y_{ij} = \Sigma y_{ijr}/M$, $e_{ij} = \Sigma e_{ijr}/M$, $\tau_j = \tau'_j + b_j$, and $b_j = \Sigma b_{jr}/M$, where summation (Σ) is across the M selected blocks. It is noted that the block mean effect b_j can simply be added to the "pure" environmental effect τ'_j, yielding a modified environmental effect τ_j. Therefore, a randomization of blocks adds noise to an additive effect. On the contrary, a randomization of single observations adds noise to the nonadditive effects, i.e. the interaction plus error term. This is demonstrated by deriving the appropriate means from Eq. 7. To take account of the fact that for each genotype by environment combination the mean is not necessarily taken across the same replicates, we denote the block mean effect for the i-th genotype and the j-th environment as $b_j(i)$ and write the means model as

$$y_{ij} = \mu + \alpha_i + \tau'_j + b_j(i) + (\alpha\tau)_{ij} + e_{ij}. \quad (9)$$

Now, the nonadditivity to be fitted by AMMI is given by the term $b_j(i) + (\alpha\tau)_{ij} + e_{ij}$. The is not the same as when AMMI is fitted to the entire dataset, for which the nonadditity is $(\alpha\tau)_{ij} + e_{ij}$. Thus, by randomizing individual observations we have added noise to the nonadditivity. This is likely to have a bearing on the AMMI model selected by cross validation. It is to be expected that we select fewer axes than would be appropriate for the complete data. This is so because all axes will be contaminated with additional noise from the data-splitting procedure and if, as a result of this contamination, an axis becomes mostly noise, it stands a high chance of being discarded in cross validation, even though it may capture a significant portion of the nonadditive pattern in the entire dataset.

With an RCB design, randomization of complete blocks is more appropriate, for then the nonadditivity in the modelling data is not artificially inflated by an added $b_j(i)$. Splitting of single observations is perfectly adequate with a CR design, for which $b_j(i) = 0$. But as Gauch (1992, p 21) points out, probably over 90% of yield trials employ the RCB design.

Having selected the best model based on the RMSPD, one would like to know how much was gained by using AMMI or BLUP in place of the cell mean. The gain may be expressed in terms of the number of replicates needed for the different models in order to achieve the same accuracy. Gauch and Zobel (1988) suggested a gain factor (GF) to be computed from the RMSPD. We will mainly use the notation given by Crossa et al. (1990), where computation of GF is discussed at some detail. They consider the square of RMSPD:

$$\begin{aligned} (\text{RMSPD})^2 &= \text{MSE}(\text{model} - \text{validation}) \\ &= \text{MSE}(\text{model}) + \text{Var}(\text{validation}) \end{aligned} \quad (10)$$

where $\text{MSE}(\text{model}) = \text{Var}(\text{model}) + (\text{Bias})^2$.

Let us, for the moment, assume a CR design. We are interested in a comparison of MSE(model) and the MSE of the cell means model [MSE(cell means)]. Denote the residual error variance of a single replicate as σ^2_O. If there are R replicates, the expected value of MSE(cell means) is σ^2_O/R. The approximate number of replicates needed for the cell means model to equal the performance of the

Table 6 Predictive accuracy of AMMI for five faba bean datasets (1985–1989) based on randomly splitting single observations in each environment

	1985	1986	1987	1988	1989
Model	RMSPD	RMSPD	RMSPD	RMSPD	RMSPD
AMMI0	7.637	7.294	6.086	5.845	5.615
AMMI1	6.377	7.318	6.047	5.486	5.202
AMMI2	5.582	7.196[a]	5.748	5.268	5.031
AMMI3	5.089	7.244	5.664	5.087	4.955
AMMI4	4.944	7.263	5.512[a]	5.010[a]	4.819[a]
AMMI5	4.814[a]	7.220	5.567	5.038	4.867
AMMI6	4.876	7.245	5.570	5.083	4.889
AMMI7	4.867	7.236	5.539	5.091	4.903
AMMI8	–	–	–	5.067	4.957
AMMIF	4.894	7.277	5.514	5.045	4.988

[a] Most accurate AMMI model

best AMMI model is obtained by equation MSE(cell means) to MSE(model) and solving for R:

$$R^* = \sigma_O^2/\text{MSE(model)}. \quad (11)$$

If, as in cross validation, the model is based on $R-1$ replicates, the gain factor is

$$\text{GF} = R^*/(R-1). \quad (12)$$

In order to estimate GF, an estimate of MSE(model) is needed. From Eq. 11, MSE(model) = (RMSPD)^2 − Var(validation) is the mean square error of a single observation, which equals σ_O^2. It may be estimated by $\sum_j s_j^2/N$. Then, an estimate of MSE(model) is given by

$$\hat{\text{MSE}}(\text{model}) = (\text{RMSPD})^2 - \textstyle\sum_j s_j^2/N. \quad (13)$$

So far we have assumed a CR design with no block effects. If we have a RCB design, the block effects will inflate the RMSPD. Since in the complete data the mean of block effects in an environment equals zero, block effects are absent in the mean y_{ij}. On the contrary, block effects are present in the means of the modelling data. Therefore, a correction for block effects is needed in Eq. 13.

From Eq. 7 we know that in the modelling data the mean of block effects b_j is contained in a modified environmental effect τ_j. In the expression for BLUP (w_{ij}) in Eq. 4, only the environmental mean $\bar{y}_{.j}$ is a function of b_j. Clearly, $\bar{y}_{.j}$ is the sum of b_j plus a linear combination of other model effects. Note that b_j (as well as τ_j) cancels out in the estimated genotypic and interaction effects. The estimate for the additive part of AMMI may be written as $\bar{y}_{.j} + (\bar{y}_{i.} - \bar{y}_{..})$. Again, in this expression it is only $\bar{y}_{.j}$ that is affected by b_j. The multiplicative part of AMMI is not influenced by b_j. The RMSPDs for AMMI and BLUP (w_{ij}) are comparable because the added noise due to randomizing complete blocks is modelled in the same way by both procedures, namely by $\bar{y}_{.j}$. On the contrary, BLUP with random environmental effects models the added noise by $h_e^2(\bar{y}_{.j} - \bar{y}_{..})$. Its RMSPD is therefore not comparable to that of AMMI and BLUP (w_{ij}), except when $h_e^2 = 1$.

Since b_j is not part of the model for the original data, a correction of MSE (model) is needed. If $R-1$ blocks are selected for modelling, b_j is given by

$$b_j = \sum_{r \neq r'} b_{rj}/(R-1)$$

where the r'-th replicate is selected for validation. Considering the restriction $\sum_r b_{jr} = 0$, we find that this term is equal to $-b_{jr'}/(R-1)$. In computing RMSPD, validation data are subtracted from prediction from a model fitted to the modelling data. Thus, for each genotype and environment combination, the difference of "model − validation" contains the quantity $-b_{jr'}/(R-1) - b_{jr'} = -Rb_{jr'}/(R-1)$. The expected mean square of this quantity is

$$\frac{R^2}{(R-1)^2}\frac{\sum_{jr} b_{jr}^2}{RN} = \frac{R}{(R-1)}\Phi(b),$$

where $\Phi(b) = \sum_{jr} b_{jr}^2/N(R-1)$. The expected value of crossproducts of $b_{jr'}$ with other effects (x, say) is zero, since r' may take on the value from 1 to R with equal probability, such that $E(b_{jr'}x) = R^{-1}(b_{j1} + \ldots + b_{jR})x = 0$. Therefore, $R\Phi(b)/(R-1)$ is the desired correction factor, and we estimate MSE (model) by

$$\hat{\text{MSE}}(\text{model}) = (\text{RMSPD})^2 - \textstyle\sum_j s_j^2/N - R\hat{\Phi}(b)/(R-1),$$

where $\hat{\Phi}(b)$ is an estimate of $\Phi(b)$. This may be obtained from the ANOVA mean square of blocks (MS_{Blocks}) in Table 2 as

$$\hat{\Phi}(b) = (MS_{\text{Blocks}} - \textstyle\sum_j s_j^2/N)/K.$$

With the estimate of MSE (model), the gain factor can be estimated using Eqs. 11 and 12.

The gain factors for the best AMMI model and for BLUP$_g$ are displayed in Table 7 (within rounding errors, the gain factor of BLUP$_e$ and BLUP$_{ge}$ are the same as for BLUP$_g$). The gain factor for BLUP$_g$ was higher than that for the best AMMI model in all cases except the 1985 dataset. In 1986, BLUP$_g$ had a gain factor of 1.457 (compared to only 1.043 for AMMI). Thus, for the arithmetic mean we would have needed almost six replicates in order to attain the same accuracy as with BLUP$_g$ based on only four replicates. The gain of BLUP$_g$ relative to that of the full model was small in 1985, and BLUP$_g$ was less efficient than AMMI in that year. This may be explained by the comparatively small number of genotypes, which results in poor estimates of the variance components needed for BLUP.

Concluding remarks

This article has suggested an application of BLUP for the purpose of obtaining reliable yield estimates in genotype × environment datasets. From the faba bean example it is concluded that BLUP may be a worthwhile alternative to AMMI, particularly if the dataset is large

Table 7 Gain factor of BLUP and the best AMMI model for five faba bean datasets (1985–1989)

	1985	1986	1987	1988	1989
Best AMMI	1.236	1.043	1	1	1.152
BLUP$_g$	1.074	1.457	1.310	1.237	1.253

enough so as to allow good estimates of variance components. It is stressed, however, that it should be decided on a case-by-case basis just which procedure is preferable. It is suggested that AMMI and BLUP be applied together routinely. An appropriate cross validation procedure will then show which model is better in a given situation. If a data-splitting procedure as that suggested by Gauch and Zobel (1988) is applied to RCB designs, complete blocks rather than single observations should be randomized, and the gain factor should be computed accordingly.

In this paper, focus has been on the RCB design because of its prevalence in many agricultural trials. When the number of genotypes is very large, however, complete blocks may become an inefficient means of error control due to systematic trends within the blocks. One approach to remedy this problem is a reduction in the block size by the formation of incomplete blocks (e.g. balanced incomplete block designs, lattice designs); another approach is to use nearest neighbour adjustment (NNA) methods. BLUP may be used in combination with both approaches. A description of NNA BLUP may be found in Stroup and Mulitze (1991). With incomplete blocks, the analysis must be based on adjusted genotypic means rather than on cell means in the respective environments. In yield trials, the most common design with incomplete blocks is the lattice design. If intra-block information is not recovered, the lattice design is appropriately analysed as an ordinary RCB design. If intra-block information is recovered, however, any combined analysis of lattice designs must be approximative (Cochran and Cox 1957, p 560). As an approximate analysis it is suggested that effects and variance components be estimated in much the same way as for the RCB design, with the only difference being that adjusted cell means are used in place of raw cell means. s_j^2 may be equated to the effective error MS of the corresponding incomplete block design (Cochran and Cox 1957, p 560).

An important advantage of AMMI is that it may be used for modelling and understanding interaction (Gauch 1992, chapter 6), a facility that is not offered by BLUP. AMMI has the further advantage that it allows the imputation of missing values (Gauch and Zobel 1990), while BLUP yields predictors only for those genotype-environment combinations that were actually observed. One might consider imputing missing values by AMMI prior to applying BLUP to the data.

AMMI and BLUP may be seen as two approaches to achieve the same goal, namely to separate pattern from noise. AMMI employs PCA, in which the first axes (i.e. those with the largest singular values) recover most of the pattern, while the bulk of noise concentrates in the later axes. On the contrary, BLUP first estimates effects of the ANOVA model and then weights them by an estimate of the corresponding pattern-to-noise ratios. In future work, it may be promising to combine the principles of AMMI and BLUP (H. G. Gauch, personal communication). Statistical theory for this type of approach appears to be as yet lacking. One would have to estimate the pattern-to-noise ratio of "repeatability" of each PCA axis in the AMMI model. Gauch (1992, p 123) proposed a simulation-based procedure for estimating the pattern-to-noise ratio that may be used in such analyses.

It would be interesting to compare BLUP to the Shifted Multiplicative Model (SHMM) recently suggested by Seyedsadr and Cornelius (1992) and Cornelius et al. (1992). This was not done here since the cross validation procedure, with which we were mainly concerned, is felt to be inappropriate for SHMM if data are from an RCB design. In future work, we will compare the accuracy of AMMI, BLUP and SHMM using Monte Carlo simulation techniques.

Appendix A

If genotypes and hence also the genotype-environment interactions are deemed random, BLUP for our particular problem may be derived by rewriting the model in Eq. 1 as

$$y_{ij} = \beta_i + u_{ij} + e_{ij}, \tag{A.1}$$

where

$\beta_j = \mu + \tau_j$ is the fixed effects part of the model and $u_{ij} = \alpha_i + (\alpha\tau)_{ij}$ is the random effects part. It will be assumed that all random effects are uncorrelated. Equation A.1 is conveniently expressed in general matrix notation:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{u} + \mathbf{e},$$

where $\mathbf{y}, \boldsymbol{\beta}, \mathbf{u}, \mathbf{e}$ are vectors of y_{ij}, β_j, u_{ij}, and e_{ij}, respectively, and $\mathbf{X}$ is an appropriate design matrix. We wish to predict the vector of "true" yields given by $\mathbf{w} = \mathbf{X\beta} + \mathbf{u}$. It is seen that $E(\mathbf{w}) = E(\mathbf{y}) = \mathbf{X\beta}$. We then have that $(\mathbf{w}, \mathbf{y})'$ is distributed with means vector and variance-covariance matrix as given by

$$\begin{bmatrix} \mathbf{w} \\ \mathbf{y} \end{bmatrix} \sim \left\{ \begin{bmatrix} \mathbf{X\beta} \\ \mathbf{X\beta} \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathbf{D} \\ \mathbf{D}' & \mathbf{V} \end{bmatrix} \right\},$$

where $\mathbf{V}$ and $\mathbf{D}$ are the covariance matrices of $\mathbf{y}$ and $\mathbf{w}$, respectively. Assume without loss of generality that the first N values in $\mathbf{y}$ (and likewise in $\mathbf{u}$ and $\mathbf{e}$) correspond to the first genotype, the next N values to the second genotype, and so on, i.e., $\mathbf{y}' = (\mathbf{y}_{11}, \ldots, \mathbf{y}_{1N}, \ldots, \mathbf{y}_{i1}, \ldots \mathbf{y}_{K1}, \ldots, \mathbf{y}_{KN})$. Then, $\mathbf{V}$ and $\mathbf{D}$ are block diagonal with $N \times N$ submatrices $\mathbf{V}^*$ and $\mathbf{D}^*$:

$$V = \begin{bmatrix} \mathbf{V}^* & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^* & \cdots & \mathbf{0} \\ & \cdot & \cdots & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}^* \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \mathbf{D}^* & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* & \cdots & \mathbf{0} \\ & \cdot & \cdots & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}^* \end{bmatrix}$$

$\mathbf{V}^*$ and $\mathbf{D}^*$ have diagonal elements $v_{jj} = \sigma_\alpha^2 + \sigma_{\alpha\tau}^2 + \sigma^2$ and $d_{jj} = \sigma_\alpha^2 + \sigma_{\alpha\tau}^2$, respectively. All offdiagonal elements of $\mathbf{V}^*$ and $\mathbf{D}^*$ are $v_{jj'} = d_{jj'} = \sigma_\alpha^2$. We have (see Searle et al. 1992, p 2709):

$$\text{BLUP}(\mathbf{w}) = \mathbf{X\beta}^\circ + \mathbf{DV}^{-1}(\mathbf{y} - \mathbf{X\beta}^\circ) \text{ and}$$

$$\text{BLUE}(\mathbf{X\beta}) = \mathbf{X\beta}^\circ = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \tag{A.2}$$

where $\mathbf{A}^-$ is a generalized inverse of $\mathbf{A}$. For the problem considered here, BLUE($\mathbf{X\beta}$) amounts to computation of the environmental mean across K genotype and R replicates (Compare Searle 1987, p 490). For calculating BLUP ($\mathbf{w}$) we need the inverse of $\mathbf{V}$, which is obtained

from the inverse of $\mathbf{V}^*$ as:

$$\mathbf{V}^{-1} = \begin{bmatrix} \mathbf{V}^{*-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{*-1} & \cdots & \mathbf{0} \\ & \cdot & \cdots & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}^{*-1} \end{bmatrix},$$

from which

$$\mathbf{DV}^{-1} = \begin{bmatrix} \mathbf{D}^*\mathbf{V}^{*-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^*\mathbf{V}^{*-1} & \cdots & \mathbf{0} \\ & \cdot & \cdots & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}^*\mathbf{V}^{*-1} \end{bmatrix}.$$

It can be shown (see Appendix B) that

$$\mathbf{D}^*\mathbf{V}^{*-1} = \frac{\sigma^2_{\alpha\tau}}{\sigma^2_{\alpha\tau}+\sigma^2}\mathbf{I}_N + \frac{\sigma^2\sigma^2_{\alpha}}{(\sigma^2_{\alpha\tau}+\sigma^2)(\sigma^2_{\alpha\tau}+\sigma^2+N\sigma^2_{\alpha})}\mathbf{J}_N, \tag{A.3}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\mathbf{J}_N$ is an $N \times N$ matrix with all elements equal to unity. With this result, the BLUP of w_{ij} can be expressed as

$$\begin{aligned} \mathrm{BLUP}(w_{ij}) = \bar{y}_{\cdot j} &+ \frac{\sigma^2_{\alpha\tau}+N\sigma^2_{\alpha}}{\sigma^2_{\alpha\tau}+\sigma^2+N\sigma^2_{\alpha}}(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}) \\ &+ \frac{\sigma^2_{\alpha\tau}}{\sigma^2_{\alpha\tau}+\sigma^2}(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot}), \end{aligned} \tag{A.4}$$

which is the formula given in Eq. 4.

Appendix B

In what follows it will be shown that $\mathbf{D}^*\mathbf{V}^{*-1}$ can be written as in Eq. A.3. We will use the following two results for matrices of the form $(a\mathbf{I}_n + b\mathbf{J}_n)$, where $\mathbf{I}_n$ is an $n \times n$ identity matrix and $\mathbf{J}_n$ is an $n \times n$ matrix with all elements equal to unity (Searle et al. 1992, p 443):

(1) $(a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = \frac{1}{a}(\mathbf{I}_n - \frac{b}{a+nb}\mathbf{J}_n)$

(2) $(a\mathbf{I}_n + b\mathbf{J}_n)(c\mathbf{I}_n + d\mathbf{J}_n) = ac\mathbf{I}_n + (ad+bc+bdn)\mathbf{J}_n$

Upon noting that

$\mathbf{V}^* = (\sigma^2_{\alpha\tau}+\sigma^2)\mathbf{I}_N + \sigma^2_{\alpha}\mathbf{J}_N$ and

$\mathbf{D}^* = \sigma^2_{\alpha\tau}\mathbf{I}_N + \sigma^2_{\alpha}\mathbf{J}_N$,

we find

$$\mathbf{V}^{*-1} = \frac{1}{\sigma^2_{\alpha\tau}+\sigma^2} - \left(\mathbf{I}_n - \frac{\sigma^2_{\alpha}}{\sigma^2_{\alpha\tau}+\sigma^2+\mathrm{N}\sigma^2_{\alpha}}\mathbf{J}_n\right),$$

and with (2) we obtain the formula for $\mathbf{D}^*\mathbf{V}^{*-1}$ given in Eq. A.3.

Acknowledgements I wish to thank Prof. H.G. Gauch and Prof. P.L. Cornelius for their numerous helpful comments and suggestions. Thanks are also extended to Dr. F. Laidig from The Bundessortenamt Hannover, Germany, for kindly providing the Faba bean data.

References

Cochran WG, Cox GM (1957) Experimental designs, 2nd edn. Wiley, New York

Cornelius PL (1993) Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. Crop Sci 33:1186–1192

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

Crossa J, Gauch HG, Zobel RW (1990) Additive main effects and muliplicative interaction analysis of two international maize cultivar trials. Crop Sci 30:493–500

Crossa J, Fox PN, Pfeiffer WH, Rajaràm S, Gauch HG (1991) AMMI adjustment for statistical analysis of an international wheat trial. Theor Appl Genet 81:27–37

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

Gauch HG (1992) Statistical analysis of regional yield trials. Elsevier, Amsterdam

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

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

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

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423–447

Hill RR Jr, Rosenberger JL (1985) Methods for combining data from germplasm evaluation trials. Crop Sci 25:467–470

Kempthorne O (1952) The design and analysis of experiments. Wiley, New York

Nachit MM, Nachit G, Ketata H, Gauch HG, Zobel RW (1992) Use of AMMI and linear regression to analyze genotype-environment interaction in durum wheat. Theor Appl Genet 83:597–601

SAS Institute (1990) SAS/IML user's guide, version 6, 1st edn. SAS Institute, Cary, N.C.

Searle SR (1987) Linear models for unbalanced data. Wiley, New York

Searle SR, Casella G, McCulloch CE (1992) Variance components. Wiley, New York

Seyedsadr M, Cornelius PL (1992) Shifted multiplicative models for nonadditive two-way tables. Commun Stat B Simul Comp 21:807–832

Steel RGD, Torrie JH (1980) Principles and procedures of statistics. A biometrical approach, 2nd edn. McGraw Hill, New York

Stroup WW, Mulitze DK (1991) Nearest neighbor adjusted best linear unbiased prediction. Am Stat 45:194–200