Methods for analysing multi-site plant variety trials
I. Estimating genotypic means at each site

G. Ye, D.L. McNeil and G.D. Hill

Plant Sciences Group, PO Box 84, Lincoln University, Canterbury, New Zealand

Abstract

This paper is the first of two that discuss statistical methods for analysing multi-site plant variety testing data sets. In this paper, methods for estimating genotypic means at each site are introduced. These methods include those that use data from each site (including the classical and spatial methods), as well as methods using data across all sites. The latter include the additive main effects and multiplicative interaction effect model (AMMI) and the multi-site best linear unbiased prediction (BLUP). The implementation of these methods using SAS programs is outlined.

Additional key words: spatial analysis, AMMI model, BLUP

Introduction

Multi-site testing is commonly used in plant breeding programs before new cultivars are recommended for release. The main purposes of multi-site testing are to investigate genotypic performance over representative environments of the target production area, and to determine the areas where the tested genotypes are adapted (i.e., quantify genotype stability).

Genotypic means at each site and across sites provide the basic information for achieving these purposes. Surprisingly, practical breeders have paid little attention to this important aspect of their testing. Consequently arithmetic means are generally used regardless of the situations. Conversely, statisticians have long recognised that simple averages provide correct estimation of means only if some very stringent assumptions are satisfied (Cochran and Cox, 1957; Snedecor and Cochran, 1980; Pearce *et al.*, 1988). Many approaches have been developed to obtain more accurate estimates and/or comparison when the assumptions for using simple procedures cannot be reasonably made, but practical breeders frequently do not appreciate these approaches. One of the reasons may be that the proposed methods are usually published in statistical journals that may not be in the reference list of practical breeders. In this paper several procedures for estimating genotypic means at each site are given. A companion paper (Ye *et al.*, 2001) summarises the methods for estimating genotypic means across sites and comparisons among these means, and methods for simultaneous selection of performance and stability. Together these two papers form a discussion of some methods that practical breeders can use relatively easily to more effectively use data provided by multiple site testing.

Using observations at each site

Classical methods

A common practice in analysing multi-site test data is to analyse the trial at each site separately. At each site, the trial is a typical one-factor experiment. Depending on the design used, standardised methods are available (Cochran and Cox, 1957; Pearce *et al.*, 1988). In the following sections, we assume the design at each site is a randomised complete block design because it is the most commonly used design in multi-site testing.

The arithmetic mean of a genotype across replicates is the easiest way to estimate a genotype's mean performance. This estimate is unbiased if observations are available for every replicate and the variance within replicates is the same (i.e., it is homogeneous). If there are missing values, the number of observations for each genotype is different and should be taken into consideration. The least-squares procedure has the property of providing estimates of means as if all the genotypes were included in all the replicates.

Spatial analysis

An important assumption for the above classical methods is that the residuals within each block are independently distributed with a constant variance. These methods are very inefficient if there is substantial heterogeneity within a block. This heterogeneity commonly arises if the number of genotypes tested is large. A group of statistical methods, collectively called spatial analysis, have been proposed to reduce the within-block heterogeneity. The following sections introduce several of these methods.

Papadakis's method and its modified versions

Papadakis (1937) first suggested that the performance of genotypes in a yield experiment should be adjusted for the local trend by an analysis of covariance with respect to the treatment (genotypic)-corrected yields of the adjacent plots. The working model for this method is

$$ y_{ij} = \mu + g_i + \beta x_{ij} + \varepsilon_{ij} \, , $$

where x_{ij} is the covariate calculated by taking the average of the genotype-corrected performance of its neighbouring plots, and β is the regression coefficient associated with the covariate.

Note that the block effect is removed from this model because it is accounted for by the covariate. However, it may be better keep the block effect in the model to ensure that the method does not perform significantly worse than the randomised blocks model.

Bartlett (1978) suggested that an iterated version of this method should be used. The first iteration is the same as the above. The analysis is then repeated iteratively based on the adjusted genotypic means from the previous iteration until the difference between the adjusted genotypic means in successive iterations is negligible.

Wilkinson *et al.* (1983) suggested that the original nearest neighbour means instead of the treatment–corrected neighbour means should be used to adjust the target observation.

The advantage of this class of methods is its simplicity. However, by using different neighbour plots, different results may be obtained. In practice various ways to form the covariate(s) can be tried and the best one used for final analysis. Usually the nearest four neighbours are used. Sometimes, two covariates are used. One is formed using the two longitudinal neigh-

bours and the other is formed using the two latitudinal neighbours.

The effectiveness of this method has been verified using real data sets (e.g., Mak *et al.*, 1978; Kempton and Howes, 1981; Bhatti *et al.*, 1991; Brownie *et al.*, 1993; Stroup *et al.*, 1994).

The necessary computations can be done as follows. Firstly, the unadjusted genotypic effect and the corresponding residuals are obtained by standard analysis of variance. Secondly, new variable(s) are formed using the average of the residuals of the neighbouring plots, and finally an analysis of covariance is done using the new variable(s) as covariate(s).

Schwarbach' weighted nearest-neighbour analysis

The weighted nearest-neighbour analysis proposed by Schwarbach (1985) is based on the following working model:

$$ y_{ij} = \mu + g_i + \frac{3}{4} NND_a + \frac{3}{4} NND_b + \varepsilon_{ij} $$

$$ = \tau_i + \frac{3}{4} NND_a + \frac{3}{4} NND_b + \varepsilon_{ij} \, , $$

where NND_a is the first nearest neighbour difference and is calculated as follows: for each plot, the observation is adjusted by subtracting the mean of the two nearest neighbours [$NND_1 = y_m - \frac{1}{2}(y_{m-1} + y_{m+1})$], then the mean of NND_1 for each genotype is computed and denoted as NND_a; NND_b is the second nearest neighbour difference and is calculated as: the mean of the genotype at the m-th plot adjusted by subtracting the average of the two genotypes at the two nearest neighbour plots [$NND_2 = \bar{y}_m - \frac{1}{2}(\bar{y}_{m-1} + \bar{y}_{m+1})$], and then the mean of the adjusted values for each genotype is computed and denoted as NND_b.

By simulation, Schwarbach (1985) showed that this procedure is better than the Wilkinson *et al.* (1983) method. The effectiveness of this method was also proven by Stroup *et al.* (1994).

Trend analysis based on specific function form

The local trend may be described by a function of plot position. The general model is

$$y_{ijk} = \mu + g_i + f(R,C) + \varepsilon_{ijk} = \tau_i + f(R_j, C_k) + \varepsilon_{ijk}$$

where y_{ijk} is the observed performance of the i-th genotype at the jk-th plot (j-th row and k-th column), and $f(Rj,Ck)$ is a function of the position given by the row (Rj) and column (Ck) values.

The difficulty with this type of method is to determine the right function form. Polynomial regression based on the row and column positions has been used successfully (Kirk *et al.*, 1980; Tamura *et al.*, 1988; Brownie, 1993). However, it is difficult to determine the optimal degree for the polynomial regression. Tamura *et al.* (1988) developed a program based on SAS to determine the most suitable polynomial function. Once the most suitable function form is defined the GLM, including all the variables, can be used to estimate the genotypic means.

Methods based on smoothing the trend

The spatial trend can also be modelled by using methods that do not rely on specific function forms. The least-square smoothing method proposed by Green *et al.* (1985) assumed that the second differences of the trend (i.e., $t_i - (t_{i-1} + t_{i+1})/2$) are independent and discovered the underlying trend pattern by choosing a weighted coefficient that gives a smooth trend and minimal residual variance. This approach can be understood as calculating the local trend as a weighted average of plot performance adjusted for differences among genotypes. Weights are inversely proportional to the distance from the adjusted plot and depend upon the observed 'fertility trend' (Clarke *et al.*, 1994). Similarly, Hackett *et al.* (1995) introduced the generalised additive model to represent the spatial trend. In their method, the residuals are smoothed against the explanatory variable 'X' (normally the position of the plot) by the use of the locally weighted running line smoother. For each value of X, say x, and some given value k, the k nearest neighbours are identified and assigned a weight according to their distance from x. The weight decreases from one for the neighbour with the same value of X to zero for the most distant of the k neighbours. These k neighbours of the plot are used to fit a linear relationship between the observation and the explanatory variables by weighted least squares and the fitted value at point x estimates the smooth function at that point.

Clarke *et al.* (1994) applied the least-squares smoothing to 12 experiments with hexaploid wheat and found that it was more efficient than the classical methods and the Papadakis method. They also developed a PASCAL program based on a procedure suggested by Green *et al.* (1985) to find the weight coefficient and the estimate of genotypic effect. For the generalised additive model, Hackett *et al.* (1995) showed how to implement it by use of the S-Plus software. SAS function PROC TPSPLINE can be used as well.

Correlated error model

The existence of local trends in a field trial usually implies that the neighbouring plots are more alike than those farther apart. By using a less restrictive assumption about the residuals, better estimates of the genotypic means can be obtained. To accommodate the spatial correlation structure of the residuals, the semivariogram concept used in geostatistics to model spatial correlation structure is used. Semivariogram is defined as one half the variance of the difference between two observations a given distance apart, and measures the spatial variability as a function of the distance between observations.

Zimmerman and Harville (1991) and Stroup *et al.* (1994) showed that the correlated error model approach was efficient, whereas Brownie *et al.* (1993) found that accounting for a trend with a correlated errors structure only was not effective.

SAS procedure PROC MIXED provides several covariance functions to model the correlated error structure and may be sufficient for analysing most field trials. However, PROC MIXED does not compute the semivariogram *per se*. External estimates of the parameters for the covariance function are required. SAS procedure VARIOGRAM can be used to determine the theoretical semivariogram model by computing the sample empirical semivariogram from the observed data set. The two companion functions 'Fvariogram' and 'Mvariogram' of Genstat were designed for finding suitable covariance functions from the observed data. Once a semivariogram model is selected, the unknown parameters describing variance and the spatial correlation can be estimated using the REML procedure and the genotypic means are estimated using generalised least squares.

Besag and Kempton (1986) first difference model

The first difference method of Besag and Kempton (1986) assumes that the first differences between adjacent plots (i.e., $z_i = y_i+1 - y_i$) are uncorrelated random variables with identical variance (σ_t^2) and other sources of variation are negligible. In other words, the systematic trend is assumed to be removed completely from the observations by first difference operation.

The advantages of this method are that the computation is simple and that the linear component of the trend can be removed completely, and consequently the estimates of genotypic effects are better than the classical methods. The disadvantage is that the resulting estimates may be less accurate if the linear component of the trend only takes a small proportion of its overall variation. Baird and Mead (1991) applied this method to analyse data sets generated from a range of yield models and concluded that this method was more efficient than a randomised block analysis and an incomplete block analysis when the yields were from models with trend components.

A generalised linear model using the differenced observations as the raw data provides the estimates of the genotypic means (see Appendix 1).

Linear variance model

An extension of the above first difference model is the so-called linear variance model described by Williams (1986), and Besag and Kempton (1986). The linear variance model superimposes a white noise term with variance (σ_ε^2) on the plot observations of the first difference model. This model removes the assumption that the first difference operation eliminates the local trend completely. It can be understood as a two-step detrending process; the first step is to remove the linear component of the trend and the second step is to model the remaining trend as a random effect with mean zero and variance σ_t^2. Therefore, it increases the power of detrending as confirmed by Besage and Kempton (1986), Baird and Mead (1991) and Wu and Dutilleul (1999).

The procedure LVARMODEL of Genstat is specially developed for this method.

Autoregressive integrated moving average (ARIMA)

Gleeson and Cullis (1987) proposed a method that assumes that the 'experimental error' is white noise and the 'spatial trend' can be regarded as random and represented by an autoregressive integrated moving average (ARIMA) model. The d-degree difference operation is used to simplify the ARIMA model. This model can be viewed as a two-step detrending process as well. The first step is to reduce the trend effect by differencing the original data (not necessarily first difference) and the second step is to model the remaining trend effect by regarding it as a random process with a covariance function. Because the difference operation and the covariance function can be selected based on the actual data set, this model is more flexible.

Cullis and Gleeson (1991) extended this procedure to two-dimensional spatial analyses. The two-dimensional analysis was shown to provide more efficient estimates of the genotypic means. Standard procedures are now available in S-Plus and Genstat software for 1- or 2-dimensional ARIMA analysis.

Using more than 1,000 variety trials, Cullis and Gleeson (1989) demonstrated that the use of this method resulted in a reduction of 42% in variances of variety yield differences compared with complete block analysis, whereas incomplete block analysis resulted in a reduction of 33%. Gleeson and Cullis, (1987) Kempton et al. (1994) and Grondona et al. (1996) showed that very simple ARIMA models usually worked very well in analysing field trials.

Random field models

Zimmerman and Harville (1991) proposed random field models to accommodate local trends. The local trend is modelled by including the "large-scale variation" and "small-scale" variation. The 'large-scale variation' is normally modelled through the mean structure (difference, smoothing operation or using a specific function); and the 'small-scale' variation is modelled through a spatially correlated structure (correlated error models). In this sense, many approaches mentioned above can be regarded as special forms of the random field models. For instance, the correlated error model takes account of only the 'small-scale' variation. But the methods using specific function and the first difference approach model only the 'large-scale' variation. Clearly, it would be better if both the 'large-scale' and 'small-scale' variations could be tak-

en into account. For instance, using a function or difference operation to account for the 'large-scale' variation, and a correlated error to account for the 'small-scare' variation (Brownie *et al.*, 1993). The ARIMA analysis discussed above assumed a specific class of covariance function. Other types of covariance functions may also be used.

Using observations from whole tests

All data in the multi-site testing data set can be used to estimate (predict) the genotypic means at each site. The additive main effects and multiplicative interaction model (AMMI) and the multi-site best linear unbiased prediction (BLUP) methods are introduced in the following sections.

AMMI Model

The AMMI model combines the additive model used in analysis of variance (ANOVA) with principle component analysis (PCA). The additive part of the AMMI model is estimated first with ANOVA, and the multiplicative part is estimated using the PCA on the ANOVA's residuals. The direct estimation of the GE interaction is generated by the multiplication of a genotype interaction PCA (IPCA) score by an environment IPCA score (Gauch, 1988). The AMMI model can be written as

$$y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{R} \lambda_k \alpha_{ik} \beta_{jk} + \theta_{ij} \ ,$$

where y_{ij} is the mean of *i*-th genotype in *j*-th environment, m is the overall mean, μ is the effect of *i*-th genotype, g_i is the effect of the *j*-th environment, λ_k is the *k*-th singular value of the GE interaction residual matrix; α_{ik} and β_{kj} are corresponding principal component scores for genotypes and environments, respectively, and θ_{ij} is the residual which contains both the unexplained interaction and the pure experiment error.

The α_{ik} and β_{kj} are obtained by multiplying the square root of the *k*-th singular value with its corresponding eigenvectors of the genotypes and environments, respectively.

To accurately estimate genotypic means the optimum number of interaction principal component axes needs to be determined. Gauch (1988) suggested a postdictive and a predictive assessment for this. The postdictive assessment uses an F-test to identify the

significance of each IPCA (root mean square difference between the observed and expected values, i.e., the square root of error mean square). The predictive assessment splits the data set into a part for model construction and a part for model validation and uses the cross validation technique. The root mean square of the predictive difference (RMSPD) and the mean square error (MSE) of the estimation [MSE(model)] are used to measure the success of the prediction. Smaller values of RMSPD and MSE(model) indicated good predictive success. RMSPD is calculated as follows: the differences between the prediction and validation observations are first squared and summed over all genotypes and environments and divided by the numbers of validation observations, and then its square root is taken. The MSE(model) can be computed approximately as

$$\sigma_M^2 = \sigma_{MV}^2 - \sigma_V^2 = (RMSPD)^2 - \sigma_\varepsilon^2 \ ,$$

where σ_M^2 is the variance of the model; σ_V^2 is the variance of validation observations and can be estimated empirically by the error mean square σ_ε^2. σ_{MV}^2 is the variance of the difference between the model and the validation observations and can be empirically estimated by the MSE(model – validation).

Piepho (1994) suggested that when data-splitting procedures were applied to RCB designs, the complete block rather than single observations should be randomised. In the case that only one replicate is used for validation, the estimate of MSE(model) can be approximated as

$$(RMSPD)^2 - \sigma_V^2 - b\hat{\phi}(b)/(b-1) \ ,$$

where b is the number of blocks in the design, and v is the number of genotypes, $\hat{\phi}(b) = (MSB - \sigma_V^2)/v$ with MSREP being the mean square of block.

All the necessary computations for the AMMI model can be done using SAS or other statistical software. See Appendix 2 for a SAS-based program for this model.

Cornelius *et al.* (1996) proposed a method which adjusts the least square estimators of the main effects and multiplicative components above by multiplying them by their respective shrinkage factors. The shrinkage factors for the genotypic and environment main effects are $s_g = \max (1 - F_g^{-1}, 0)$, $s_e = \max (1 - F_e^{-1}, 0)$,

where the F_g and F_e are the F-statistics for testing the genotypic and environment effects against the error mean square. The shrinkage factor for the k-th multiplicative component is $s_k = \max (1- F_k^{-1}, 0)$, with $F_k = \dfrac{b\lambda_k^2}{df_k \sigma_\varepsilon^2}$, where df_k is the degrees of freedom associated with k-th multiplicative component, and can be computed as $v + s - 1 - 2k$. This approach uses all the multiplicative components and consequently this difficult issue for the classical AMMI model is avoided. A simulation study showed that this approach is at least as good as the classical AMMI model (Cornelius and Crossa, 1999).

Multi-site BLUP

The Best Linear Unbiased Prediction (BLUP) method was developed for predicting the random effect when the working model is a mixed linear model (Henderson, 1984). To apply the BLUP method in estimating the genotypic means, it is necessary to define at least one of the main effects as random. In the traditional sense, the genotypic effect is fixed because the experimenters are only interested in the particular set of genotypes. However, White and Hodge (1990) argued that the genotypic effect could be regarded as a random effect if the set of genotypes can be regarded as a random sample of a single population. More generally, an effect may be regarded as a random effect if the levels of the effect may reasonably be assumed to come from a probability distribution (Maclean *et al.*, 1991; Robinson, 1991). If we assume that genotype effect is random, and the environment effect is fixed, then the multi-site linear model can be written as

$$y_{ij} = \tau_j + u_{ij} + \varepsilon_{ij} ,$$

where $\tau_j = \mu + e_j$ represents fixed effects and $u_{ij} = g_i + (ge)_{ij}$ represents random effects.

This model can be written in matrix notation as

$$y = X\beta + Zu + e ,$$

where X and Z are design matrixes which link the observation in y with the fixed and random parts, respectively.

The BLUP of the random effect u is

$$CV^{-1} (y - X\beta^0) .$$

The best linear unbiased estimation (BLUE) of the fixed part is

$$X\beta^0 = X(X'V^{-1}X)^- X'V^{-1}y .$$

The BLUP of the genotypic means at each site is

$$\text{BLUP (y)} = X\beta^0 + CV^{-1} (y - X\beta^0) .$$

where V is the covariance matrix between the observations in y, and C is the covariance matrix among the observations in y and the unobservable true genotypic effects.

The above equation is the general BLUP prediction equation in the sense that the C and V matrices can be of any type. It is clear that the BLUPs of genotypic means can be easily obtained once the matrixes V and C are determined. Therefore, the most important issue in the application of the BLUP method is to define V and C. In the following sections, several situations are considered

Case 1: Homogenous mean variances

Assume that the first s rows of the Y vector are the means of the first genotype at '*s*' sites. If the test is balanced and the mean variance is homogenous, C and V matrices are block diagonals with s × s sub matrices V* and C*, respectively. The non-diagonal elements of V^* and C^* are σ_g^2, and the diagonal elements are $\sigma_g^2 + \sigma_{ge}^2 + \sigma_\varepsilon^2$ and $\sigma_g^2 + \sigma_{ge}^2$, respectively, where σ_g^2 is the genotypic variance, σ_{ge}^2 is the genotype-by-environment variance and σ_ε^2 is pooled error.

Case 2: Heterogeneous GE variances and homogeneous within site error variances

In multi-site testing, the mean variances are normally heterogeneous since genotypes contribute differently to the GE interaction and/or the within site error variance is different. The contribution of each genotype to the GE interaction variance is proposed as a stability parameter and termed the stability variance (Shukla, 1972). Therefore, a more realistic model should take the different stability variances into account. Assume the within site error variance is homogeneous, genotypic effect is random and environment effect is fixed, and the V and C matrixes are block diagonal with s × s sub matrices V* and C*, respectively. The non-

diagonal elements of V* and C* are σ_g^2 and the diagonal elements are $\sigma_g^2 +_i \sigma_{ge}^2 + \sigma_\varepsilon^2$ and $\sigma_g^2 +_i \sigma_{ge}^2$, respectively.

Case 3: Both GE variances and within site error variances heterogeneous

If both the GE interaction variance and the within site error variance are heterogeneous, the V and C matrixes are the same as for case 2 except that the within site error variance varies from site to site.

If the sole objective of the analysis is to predict genotypic means, SAS procedure MIXED can be used to carry out the analysis (for statements see Appendix 3) using the observed genotypic means at each site as the observational units.

Piepho (1997; 1998) extended this mixed model-based method to the linear model with multiplicative component. Again it assumes that the genotypic effect is random and the environmental effect is fixed, and that V and C are still block-diagonal matrixes with submatrixes V* and C*, respectively. The non-diagonal elements of V^* and C^* are $\sigma_g^2 + \sum_{k=1}^{R} \beta_{jk} \beta_{j'k}$, and the diagonal elements are $\sigma_g^2 + \sum_{k=1}^{R} \beta_{jk}^2 + \sigma_p^2$ and $\sigma_g^2 + \sum_{k=1}^{R} \beta_{jk}^2 + \sigma_p^2 - \sigma_\varepsilon^2$, respectively, where σ_p^2 is the residual variance based on the cell mean model and contains the interaction variance and part of the error variance, $\sigma_\varepsilon^2 = \sigma_e^2 / r$, and β_{jk} is the k-th score of the j-th environment.

Conclusion

The classical methods for estimating the genotypic means at each site of a multi-site variety trial are simpler, but do not take the possible spatial heterogeneity among plots into consideration, and do not use all the information contained in multi-site test data. Spatial analysis using the information of other plots, particularly the neighbouring plots, can therefore significantly improve estimates of true means. However, there is no clear rule to guide the selection of the appropriate spatial model. In practice several of the models could be

used and the best one determined. Most of the methods can be easily implemented using SAS and/or Genstat, which is commonly accessible. Moreover, specialised software such as ASREML (Gilmour *et al.*, 1996) has been developed. Therefore, it is now time for practical breeders to become familiar with these methods and for those familiar with the methods to make simple "plug in the data" versions available so that full use can be made of their benefits.

The AMMI model is more flexible in the sense that there is no requirement for a large number of genotypes/sites. However, to determine the number of interaction principle component axes to use for the classical AMMI model is not a trivial problem. The shrinkage estimators seem to be better than the classical AMMI model. The multi-site BLUP is better than the classical AMMI when variance components can be estimated accurately. In addition, the missing GE combinations do not create serious problems. However, to obtain accurate estimates of the variance components, the number of tested genotypes and/or the number of sites must be large. As with the classical methods, neither AMMI nor multi-site BLUP takes possible systematic heterogeneity among plots into account. Using the spatially adjusted values as the raw data for the AMMI or multi-site BLUP analysis the advantages of both types of methods could be explored. Patterson and Nabugoomu (1990) outlined such a two-step procedure. Cullis *et al.* (1998) developed a method that combines spatial analysis and multi-site BLUP into a single step, and showed that it is more efficient than the two-step procedure. However, special software is required to carry out the analysis and, as stated above, simplified 'plug in the data' software and recognition of the advantages of the methods are needed.

References

Baird, D. and Mead, R. 1991. The empirical efficiency and validity for two neighbour models. *Biometrics 47*, 1473-1487.

Bartlett, M.S. 1978. Nearest neighbour models in the analysis of field experiments. *Journal of the Royal Statistics Society, B 40*, 147-174.

Besag, J.and Kempton, R. 1986. Statistical analysis of field experiments using neighbouring plots. *Biometrics 42*, 231-251.

Bhatti, A.U., Mulla, D.J., Koehler, F.E. and Gurmani, A.H. 1991. Identifying and removing spatial correlation from yield experiments. *Journal of the Soil Science Society of Americal 55*, 1523-1528.

Brownie, C., Bowman, D.T. and Burton, J.W. 1993. Estimating spatial variation in analysis of data from field trials: A comparison of methods. *Agronomy Journal 85*, 1244-1253.

Clarke, F.R., Baker, R.J. and DePauw, R.M. 1994. Moving mean and least squares smoothing for analysis of grain yield data. *Crop Science 34*, 1479-1483.

Cochran, W.G. and Cox, G.M. 1957. Experimental design 2nd ed. John Wiley & Son Inc. New York.

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

Cornelius, P.L, and Crossa J. 1999. Prediction assessment of shrinkage estimators of mutiplicative models for multienvironment cultivar trials. *Crop Science 39*, 998-1009.

Cullis, B.R. and Gleeson, A.C. 1989. The efficiency of neighbour analysis for replicated variety trials in Australia. *Journal of Agriculture Science, Cambridge 113*, 233-239.

Cullis, B.R.and Gleeson, A.C. 1991. Spatial analysis of field experiments- an extension to two dimensions. *Biometrics 47*, 1449-1460.

Cullis, B.R., Gogel, B., Verbyla, A. and Thompson, R. 1998. Spatial analysis of multi-environment early generation variety trials. *Biometrics 54*, 1-18.

Gauch, H.G. 1988. Model selection and validation for yield trials with interaction. *Biometrics 44*, 705-715.

Genstat 5 Committee 1993 Genstat 5, Release 3, reference manual. Clarendon Press, Oxford.

Gilmour, A.R., Thompson, R. and Cullis, B.R. 1996. ASREML. Biometric Bulletin 3. NSW Agriculture, Organge, NSW, Australia.

Gleeson, A.C. and Cullis, B.R. 1987. Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. *Biometrics 43*, 277-288.

Grondona, M.O., Crossa, J., Fox, P.N. and Pfeiffer, W.H. 1996. Analysis of variety trials using two-dimensional separable ARIMA processes. *Biometrics 52*, 763-770.

Green, P., Jenison, J. and Secheult, A.H. 1985. Analysis of field experiments by least squares smoothing. *Journal of the Royal Statistics Society, B 47*, 299-315.

Hackett, C.A., Reglinski, T. and Newton, A.C. 1995. Use of additive models to represent trends in barley field trial. *Annual Applied Biology 127*, 391-403.

Henderson, C.R. 1984. Application of Linear Models in Animal Breeding. University of Guelph. Guelph, Ontario.

Kempton, R.A. and Howes, C.W. 1980. The use of neighbouring plot values in the analysis of variety trials. *Applied Statistics 30*, 59-70.

Kempton, R.A., Seraphin, J.C. and Sword, A.M. 1994. Statistical analysis of two-dimensional variation in variety yield trials. *Journal of Agriculture Science, Cambridge 122*, 335-342.

Kirk, H.J., Hayes, F.L. and Monroe, R.J. 1980. Application of trend analysis to horticultural field trials. *Journal of the American Society of Horticultural Science 105*, 189-193.

Maclean, R.A., Sanders, W.L. and Stroup, W.W. 1991 A unified approach to mixed model theory. *American Statistician 45*, 54-64.

Mak, C., Harvey, B.L. and Berdahl, J.D. 1978. An evaluation of control plots and moving means for error control in barley nurseries. *Crop Science 18*, 870-873.

Papadakis, J.S. 1937. Methode statistique pour des experiences sur champ. Bulletin de I'Institute d' Amelioration des Plantes a Salonique No. 23.

Patterson, H.D. and Nabugoomu, F. 1992. REML and the analysis of series of crop variety trials. *In* Proceedings of the 25th International Biometric Conference, 77-93. Statistics Section AgResearch: Hamilton, New Zealand.

Pearce, S.G., Clarke, R.F. and Kempton, R.E. 1988. Manual of Crop Experimentation. London: Charles Griffin and Company Ltd, New York: Oxford University Press.

Piepho, H.P. 1994. Best linear unbiased prediction (BLUP) for regional yield trials: a comparison to additive main effects multiplicative interaction (AMMI) analysis. *Theoretical and Applied Genetics 89*, 647-654.

Piepho, H.P. 1997. Analysing 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. *Theoretical and Applied Genetics 97*, 195-201.

Qian, C.G, Basford, K.E, DeLacy, I.H. and Cooper, M, 2000. Evaluation of experimental designs and spatial analysis in wheat breeding trials. *Theoretical and Applied Genetics 100*, 9-16.

Robinson, G.K. 1991. That BLUP is a good thing: the estimation of random effects. *Statistical Science 6*, 15-51.

SAS Institute 1997. SAS/STAT software: changes and enhancements through release 6.12. SAS Institute Inc., Cary, North Carolina, USA.

Schwarzbach, E. 1985. Weighted nearest-neighbour analysis. *Biometrics 41*, 1088

Shukla, G.K. 1972. Some statistical aspects of partitioning genotype-environment components of variability. *Heredity 29*, 237- 245.

Snedecor, G.W.and Cochran, W.G. 1980. Statistical methods. 7th edn. Iowa State University Press, Ames, Iowa, USA.

Stroup, W.W., Baenziger, S.P. and Mulitze DK. 1994. Removing spatial variation from wheat yield trials: A comparison of methods. *Crop Science 86*, 62-66.

Tamura, R.N., Nelson, L.A. and Naderman, G.C. 1988. An investigation of the validity and usefulness of trend analysis for field plot data. *Agronomy Journal 80*, 712-718.

White, T. and Hodge, G. 1989. Predicting breeding values with applications in forest tree improvement. pp 367, Kluwer Academic Publishers; Dordrecht; Netherlands.

Wilkinson, G.N., Eckert, S.R., Hanock, T.W. and Mayo, O. 1983. Nearest neighbour (NN) analysis of field experiments (with discussion). *Journal of the Royal Statistics Society B 45*, 151-211.

Williams, E.R. 1986. A neighbour model for yield experiments. *Biometrika 73*, 279-287.

Wu T. and Dutilleul P. 1999. Validity and efficiency of neighbour analysis in comparison with classical complete and incomplete block analysis of field experiments. *Agronomy Journal 91*, 721-731.

Ye, G., McNeil, D.L. and Hill, G.D. 2001. Some methods for analysing multisite plant variety trials. II. Selection for yield and stablility. *Agronomy New Zealand 31*, 25-33

Zimmerman, D.L. and Harville, D.A. 1991. A random field approach to the analysis of field-plot experiments and other experiments. *Biometrics 47*, 223-239.

Appendix 1: SAS programs for implementing some spatial methods

In the following programs, A is a data set consisted of differenced observations, ROW and COL are variables to index the row and column positions of the plots, respectively.

1.1. Simple RCB design analysis

Using GLM	Using MIXED

PROC GLM;
  CLASS GEN BLOCK;
  MODEL YIELD = GEN BLOCK;
  LSMEANS GEN;
RUN;
	PROC MIXED;
	  CLASS GEN BLOCK;
	  MODEL YIELD=GEN;
	  RANDOM BLOCK
	  LSMEANS GEN;
	RUN;

1.2. Correlated error model

Assume that the error covariance structure can be modelled by a spherical covariance function.

Keep block structure	Ignore block structure

PROC MIXED NOPROFILE;
  CLASS GEN BLOCK;
  MODEL YIELD = GEN BLOCK;
  PARMS () () ()/NOITER;
  REPEATED /SUB = BLOCK TYPE
                = SP(SPH) (ROW COL);
LSMEANS GEN;
RUN;
	PROC MIXED NOPROFILE;
	  CLASS GEN;
	  MODEL YIELD = GEN;
	  PARMS () () ()/NOITER;
	  REPEATED /SUB = INTERCEPT TYPE
	                = SP(SPH) (ROW COL);
	LSMEANS GEN;
	RUN;

(Note: PARMS statement defines the parameters of covariance function estimated externally)
Other available covariance function are: EXP: exponential, GAU: Gaussian, and POW: Power.

1.3. First difference

The difference operation is done within block	Block structure is ignored

PROC MIXED DATA = A METHOD = REML;
  CLASS GEN BLOCK;
  MODEL YIELD = GEN;
  RANDOM BLOCK;
  LSMEANS GEN;
RUN;
	PROC GLM DATA = A;
	  CLASS GEN ;
	  MODEL YIELD = GEN;
	  LSMEANS GEN;
	RUN;

1.4. Trend analysis:
see Brownie *et al*.,(1993)

1.5. Nearest neighbour analysis:
see Brownie *et al*.,(1993)

Appendix 2. SAS program for AMMI model.

```
/* Obtain GE means*/
Proc Glm;
Class site gen;
Model yield = site block(site) gen gen* site;
Lsmeans gen*site;
Out = mset lsmean = my;
/* Obtain GE residual */
 Proc Glm set = mset;
   Class site gen;
    Model my = site gen;
    Output out = setR Residual = ry;
/* Get multivariate form of GE means*/
Proc sort data = mset;
  By gen;
Data mset (keep= y1-ys gen);
  Array yy(s) y1-ys;
  Do site = 1 to s;
  Set mset;
  By gen;
  yy(site) = myield;
  If last.gen then return;
Data mset (keep= y1-ys);
   Set mset;
/* Get multivariate form of GE residual*/
Proc sort data = setR;
  By gen;
Proc transpose;
  Out = RGE (rename (_1=yr1 _2=yr2   s=yrs¹));
  By gen; Id = site;
Data RGE (keep= y1-ys);
  Set RGE;
/*Singular value decomposition*/
Proc iml; s = ?¹;
v = ?²;
s = ?²;
Use RGE;
Read all into D;
```

```
E = D'*D;
V= D*D';
Call eigen (E1, E2, E);
Call eigen (V1, V2, V);
C = E1[1,1];
E = E2[,1]; V= V2[,1];
;/* first multiplicative component/
Y1=C'*V*E';
C= E1[2,1]; E= E2[,2]; V = V2[,2];
/* second multiplicative component*/
Y2=C'*V*E';
/* multiplicative effect*³/
GE = Y1+ Y2;
/*overall, genotypic & site means*/
Use mset;
Read all into M;
OM = M[:];/* overall mean*/
OM = [v,s,OM];
/*Genotypic mean*/
VM = M[,+];
Vm = VM/s;
Do i = 1 to s;
  VM = VM|| VM; End;
VM = VM [1:v, 1:s];
/*site means*/
SM = M[+,];
SM = SM/v;
Do i = 1 to v;
SM= SM//SM;
End;
SM = SM [1:v, 1:s]
Reset print;
Pre = Gm + SM +GE – OM;
/*Compute RMSPD*/
Y = Pre – Val⁴;
RMSPD = Y[##];RMSPD = RMSPD/(v*s);
RMSPD = RMSPD**0.5;
```

[1] the 's' should be replaced by the number of sites.

[2] the question marks are to be replaced by the number of genotypes and the number of sites.

[3] assume two multiplicative components are required.

[4] Val is the v×s matrix and consists of GE means obtained using validation part of data.

Appendix 3. Multisite BLUP

Assume that environment effect is RANDOM and genotype effect is FIXED. In following programs, the BLUE of fixed (GEN) effect and the BLUP of the random (SITE) are obtained; the BLUPs of genotypic means at each site can be computed in terms of the mixed linear model. If plot means are used, BLUP (y_{ij}) = BLUE (GEN_i) + BLUP $(SITE_j)$ + BLUP $[(BLOCK(SITE_j)]$+ BLUP $[(GEN*SITE)_{ij}]$. If genotypic means at each site are used, BLUP (y_{ij}) = BLUE (GEN_i) + BLUP $(SITE_j)$. Alternatively, the ESTIMATE statement in PROC MIXED can be used to obtain the BLUPs of genotypic means at each site.

3.1 Using plot means

Homogeneous within site error variances	Heterogeneous within site error variances

PROC MIXED;
   CLASS GEN SITE BLOCK;
   MODEL YIELD = GEN/DDFM
           = SATTERTH SOLUTION;
   RANDOM SITE BLOCK(SITE) GEN*SITE;
RUN;
	PROC MIXED;
	   CLASS GEN SITE BLOCK(SITE);
	   MODEL YIELD = GEN/DDFM = SATTERTH SOLUTION;
	
	   RANDOM SITE BLOCK(SITE) GEN*SITE/SOLUTION;
	   REPEATED /SUB = SITE TYPE = UN SOLUTION;
	RUN;

3.2. Using genotypic means at each site

Homogenous mean variances	Heterogenous mean variances

PROC MIXED METHOD = REML;
   CLASS GEN SITE;
   MODEL YIELD = SITE/SOLUTION;
   RANDOM INT GEN/SOLUTION;
RUN;
	PROC MIXED METHOD = REML;
	   CLASS GEN SITE;
	   MODEL YIELD = SITE/SOLUTION;
	   RANDOM INT /SUB = SITE /SOLUTION;
	   REPEATED / GROUP = GEN TYPE = UN(1); RUN;