UNIVERSITY OF SOUTHERN QUEENSLAND

HOW LOW CAN YOU GO? PERFORMANCE OF FACTOR ANALYTIC MODELS IN THE ANALYSIS OF MULTI-ENVIRONMENT TRIALS WITH SMALL NUMBERS OF VARIETIES

Bethany Macdonald
B.Sc.

Faculty of Sciences,
The University of Southern Queensland

March 2018

Submitted in partial fulfillment of the requirements of the award of
Bachelor of Science with Honours

CERTIFICATION OF DISSERTATION

I certify that the ideas, experimental work, results, analyses, software and conclusions reported in this dissertation are entirely my own effort, except where otherwise acknowledged. I also certify that the work is original and has not been previously submitted for any other award, except where otherwise acknowledged.

Bethany Macdonald Date

Endorsement

Signature of Supervisors Date

Abstract

Crop breeding programs test large numbers of crop varieties in field trials spanning a range of years and locations, with these groups of trials known as multi-environment trials (MET). In the early stages of crop breeding programs large numbers of new varieties are grown in a small number of field trials. The best varieties in each stage are selected to progress to the next stage so that in the final stages a small number of elite varieties are grown in a large number of field trials across the country. These trials are conducted to determine which varieties perform best in which environments and an appropriate statistical analysis resulting in accurate predictions of the variety by environment (VxE) effects is integral to this.

There have been many statistical approaches to the analysis of MET data, however all methods involve investigating the nature of the VxE effects. The factor analytic (FA) structure for the VxE effects allows heterogeneity of genetic variance for environments and heterogeneity of genetic covariance between pairs of environments, and is currently considered best practice in the analysis of MET data in Australia. The FA model has been shown to be the superior model for large numbers of varieties both in terms of goodness-of-fit and the selection of superior varieties. However, this superiority has not been demonstrated for small numbers of varieties, such as in the late stages of crop breeding programs, despite being regularly used in such scenarios. Five data sets with different underlying VxE patterns and numbers of trials, four numbers of varieties, and two levels of varietal concurrence were used to provide scenarios for a simulation study to investigate the adequacy of an FA variance structure for VxE effects. How the accuracy of the FA model changes as the number of crop varieties decrease, along with the implications the underlying VxE variance structure and level of varietal concurrence have on the accuracy of the FA model when dealing with small numbers of varieties were investigated. The comparisons were based on the mean square error of prediction of the VxE effects.

This study showed that 15 varieties per trial is sufficient in a MET data set to accurately estimate the VxE effects, and that in some cases MET data sets with

even as few as 10 varieties could be considered. It was found that the underlying patterns in the variance of the VxE effects impacted on how the accuracy of the FA model compared to the accuracy of other models, especially for very small numbers of varieties. In addition this study demonstrated that the FA model is affected by changes in concurrence more than the other models that were considered, however these changes in accuracy have minimal implications. Finally, this study highlighted the tendency of the log-likelihood ratio test to select overly complicated models in its test for a significant model improvement.

Acknowledgements

I would like to thank the following people and organisations for their contributions toward my Honours project:

A huge thank you to Dr. Rachel King and Dr. Alison Kelly for their supervision, guidance and advice. Their support was invaluable and made my year much less stressful than it could have been.

Col Douglas, Merrill Ryan, and Kristy Hobson for the use of their data.

The Grains Research and Development Corporation and Queensland Department of Agriculture and Fisheries for supporting this study.

Contents

List of Tables

List of Figures

Chapter 1

Literature review and introduction

Crop breeding programs conduct large numbers of field trials spanning multiple years and locations with the aim of breeding, selecting, and subsequently releasing crop varieties which outperform those currently available. These groups of trials are known as multi-environment trials (MET), where the term environment refers to year-location combinations. The aim of these trials is to investigate the performance of crop varieties in many environments in order to determine those which perform well across a range of environments and those which excel in specific environments (Smith *et al.*, 2001). These trials form the foundation of crop breeding programs in Australia and many other countries and are the basis for which a crop variety may be deemed suitable for commercial release (Smith *et al.*, 2005; Welham *et al.*, 2010). A variety is of the form below subspecies, having characteristics distinct from other varieties but able to be freely crossed with them.

Although the structure may differ slightly between countries, crop breeding programs typically contain multiple stages with the best varieties selected at each stage, so that the number of varieties tested decrease as the stages progress. The main focus of these breeding programs tends to be grain yield, although, many other traits may also be of interest and the concepts that will be discussed can easily be applied to other normally distributed traits. The early stages of the program consider early-generation material and large numbers of new breeding lines (usually greater than 500) are grown in a small number of field trials (usually less than three), while in the final stages small numbers of elite breeding lines (usually less than 40) are grown in a large number of field trials spanning the country and consequently capturing a range of geographical locations and growing seasons (Smith *et al.*, 2005). Due to the differences in the structure of the MET data originating from the early or late stages of the breeding programs, different considerations must be made for the analysis in each case.

1.1 Early methods for the analysis of MET data

The statistical analysis of MET data has been approached in many ways over the years. All of the methods involve investigating the nature of the variety

by environment (VxE) effects, which describe the yield performance of different varieties across multiple environments. The aim is to determine whether the observed differences in yield are due to variety (genetic) differences, environment differences, or the interaction between variety and environment. Typically a two-stage method of analysis has been used which involves estimating the mean yield for each variety at individual trials and then combining the means from each trial to form the data for the second stage of the analysis (Smith *et al.*, 2005). The second stage has been subject to a large range of statistical methods, with statisticians constantly attempting to better investigate the VxE interaction. Kempton (1984) discusses the classical analysis of variance (ANOVA) approach which fitted variety and environment main effects along with a VxE interaction effect, partitioning the sum of squares into components accounting for varieties, environments, and the VxE interaction to gain insight into the variation in varietal response across different environments. However, this traditional approach fails to provide substantial insight into the VxE interaction and in addition is difficult to interpret when there are large numbers of varieties and/or environments. Kempton (1984) mentions a number of methods that build on the traditional approach (ANOVA) with the aim of simplifying interpretation and deepening insight into the nature of the response of varieties across environments. These methods include further partitioning of the sum of squares through the classification of groups; regression analysis; principal components analysis; and the biplot technique.

Finlay & Wilkinson (1963) were some of the first authors to argue that the traditional ANOVA methods failed to adequately describe the VxE pattern and consequently proposed a linear regression method to compare the performance of varieties grown across a range of environments. The mean yield for each trial was used as an environmental index providing an evaluation of the environment. A linear regression was performed for each variety of individual yields on these environmental indices for each trial and the resulting regression coefficients used to categorise the sensitivity of a variety to environmental change. The regression coefficient along with the mean yield of a variety across environments provided an indication of a variety's phenotypic stability and adaptability; for example a variety with a small regression coefficient and high mean yield should yield consistently well in all environments, meaning that it is phenotypically stable and well adapted to all environments. This method has been shown to be informative and is useful for summarising the VxE effects when the VxE effects have a strong linear association with the environmental index, however, this linear relationship may not always hold (Byth *et al.*, 1976).

To address what they saw as the inadequacies of the regression method proposed by Finlay & Wilkinson (1963), Byth *et al*. (1976) proposed a two-way pattern analysis using numerical classification as an alternative option for investigating the VxE interaction. Pattern analysis had the advantage of not being dependent on the strength of the linear association between the VxE effects and an environmental index. Environments and varieties were each separately classified into 10 groups based on similarities in yield performance using the methods described by Mungomery *et al*. (1974). This resulted in reducing the size of the data matrix by 97 per cent but resulted in a loss of only 18 per cent of the total variation available in the full data set. Although this method summarised the data and meaning could be conveyed onto the groups in the example used by Byth *et al*. (1976), generally these groups are arbitrary and difficult to interpret biologically, conveying little information regarding the VxE pattern (Kempton, 1984).

Principal components analysis (PCA) has also been used to summarise the VxE effects. PCA is a popular multivariate technique producing linear combinations (components) of the variables which describe variation in the data. The components are uncorrelated and are ordered such that the first component explains the largest proportion of variance in the original data, the second component explains the second largest proportion of variance and so on. Ideally the majority of the variation in the data could be described using only a small number of components, increasing the ease of interpretation. Gabriel (1971) proposed the biplot technique which provides a method to represent graphically the response of a variety across different environments. The biplot has the advantage that the expected response of a variety in a particular environment can be determined through visual inspection of the biplot. However, this method is most useful when a high proportion of the variation in the VxE effects can be explained by only one or two components, allowing the majority of the variation to be described in two dimensions, or a single biplot, as opposed to needing to interpret multiple biplots.

Additive main effects and multiplicative interaction (AMMI) models combine the two earlier methods of ANOVA and PCA, using ANOVA to calculate additive main effects for varieties and environments and PCA to model the VxE interaction (Gauch, 1992). While there was some use of these models following their proposal in 1952 (Williams, 1952; Pike & Silverberg, 1952), Kempton (1984) provided the first sustained application of AMMI models to yield data (Gauch, 1992). Although these models provide insight into the VxE interaction, their use is restricted

by the requirement of balanced data in which every variety is grown at every environment, something that is frequently not the case within breeding programs.

1.2 Linear Mixed Models

In the approaches that have been discussed variety and environment main effects and VxE interaction effects were all treated as fixed effects. When there is only interest in the treatments considered in the experiment, the effects are called fixed effects, however, when a factor in an experiment is considered to be a random sample from a population and the specific levels of the factor are of no interest, such as the effects of individual mice in an experiment, the effects are called random effects (Searle, 1997). A model that contains only fixed effects, aside from the error term which is always random, is known as a linear model, or fixed effects model. Similarly when all the effects in a model are random effects, the model is known as a random effects model and models which contain both fixed and random effects are known as mixed models. Random effects are assumed to follow a Gaussian distribution with mean zero and constant variance. The linear mixed model (LMM) is an extension of the linear model, allowing for correlated error terms and additional random components. The three advantages of the LMM compared with linear models identified by Smith *et al.* (2005) for MET data are the ease with which unbalanced data is handled; the potential to model within-experiment error variation more realistically; and the ability to assume some effects to be random rather than fixed. These advantages make these models very flexible and underline why they have been embraced in the analysis of MET data.

In LMMs fixed effects are estimated using best linear unbiased estimation (BLUE) and random effects are estimated using best linear unbiased prediction (BLUP). There is a convention of "estimating" fixed effects and "predicting" random effects; however, Robinson (1991) states that "BLUP is a predictor only in the same way as most estimates are predictors". In general the variance parameters necessary for the estimation of the fixed and random effects are unknown and are estimated through restricted maximum likelihood (REML, Patterson & Thompson 1971). Consequently the fixed and random effects are estimated as empirical BLUEs (E-BLUEs) and empirical BLUPs (E-BLUPs) respectively, as they are based on estimated, rather than known, variance parameters.

The estimation method of REML consists of maximising the likelihood of a set of selected error contrasts, which is achieved using a system of score equations that are solved iteratively. Patterson & Thompson (1971) utilised a Fisher scoring

algorithm, however, less computer intensive methods have been derived, such as first and second order derivative free methods which employ sparse matrix methods. The average information (AI) algorithm is a second order scheme proposed by Gilmour *et al.* (1995), who showed it to be computationally convenient and efficient in the estimation of variance components when using REML. The AI algorithm is a modified Fisher scoring algorithm which uses an approximate average of the observed and expected information matrices rather than the expected information matrix. It is especially powerful for large data sets with complex variance models (Smith *et al.*, 2001).

Like early approaches to the analysis of MET data, early LMM applications tended to use a two-stage approach in which variety means were obtained from individual trial analyses in the first stage and then combined in an overall analysis in the second stage, which can be either weighted or unweighted. However, LMMs are not restricted to a two-stage approach, and allow individual plot data from multiple trials to be analysed in a single analysis, known as a one-stage analysis. Within this one-stage analysis, a LMM approach also allows for fitting separate covariance structures for the residual effects at each trial, where residual effects refer to all effects peripheral to the VxE effects (Smith *et al.*, 2001). Examples of these include experimental design terms or terms to model field trend within a trial (see for example Gilmour *et al.* (1997)).

Methods for modelling the residual effects fall into the broad categories of randomisation or model based approaches (Smith *et al.*, 2005). In the randomisation approach the model for these effects is determined by the experimental design, while the model based approach aims to offer the best fit to the data and focuses on accounting for spatial variation throughout the field. Variety trials tend to be arranged in a rectangular array with a number of rows and columns and this structure is utilised when modelling spatial variation. Gilmour *et al.* (1997), building on the approach of Cullis & Gleeson (1991), partitioned spatial variation into smooth trend, resulting from changes in response due to fertility and moisture status, both on a global and local scale, and extraneous variation, resulting from experimental procedures such as serpentine harvesting of rows. Gilmour *et al.* (1997) accommodate smooth global variation and extraneous variation through the inclusion of fixed and random effects where appropriate and model local stationary trend through the use of a correlation structure on plot residuals. A separable first order autoregressive model was found to generally be a robust option.

There is some criticism of spatial models in that the estimated treatment effects rely solely on the chosen model. The advantages of spatial models are thought to outweigh the disadvantages; however, an approach that merges the randomisation and model based approaches is considered to be more robust (Smith *et al.*, 2005). Using this merged approach the randomisation model is used as the base model and spatial models are used to explain remaining variation. When these hybrid models are applied to the analysis of MET data, they offer superior fits to the data compared to simple randomised complete block models which assume common block variance and plot variance for all trials and have rarely been found to provide good fits to Australian data (Smith *et al.*, 2005).

1.3 Models for the VxE effects in a LMM

Patterson *et al.* (1977) were among the first to analyse MET data using a LMM, with such models becoming increasingly popular in the last three decades. Most early models included the VxE interaction effect as a random effect, and each of the variety and environment main effects as either fixed or random effects (Smith *et al.*, 2005). Each of these random effects were assumed to follow a Gaussian distribution with mean zero and constant variance. These assumptions were somewhat limiting, assuming that environments had constant genetic variance, pairs of environments had constant genetic covariance, and environments had constant error variance. These assumptions have been acknowledged as questionable by a number of authors (including Patterson & Silvey 1980; Patterson & Nabugoomu 1992; Cullis *et al.* 1998) and consequently more complex models were proposed. However, with this added complexity comes added difficulty in fitting such models as more parameters must be estimated.

More complex mixed models made allowances for some heterogeneity of genetic variance between environments. Gogel *et al.* (1995) and Nabugoomu *et al.* (1999) proposed a regression approach similar to the method popularised by Finlay & Wilkinson (1963), but in a mixed model setting. In these models environment means provide a quantitative grading of the environment and can be used as a surrogate for potentially complex environmental variables. However, it is important to note that environment means must be estimated from the data and are consequently subject to error.

Piepho *et al.* (1998) proposed an alternate regression based approach to explore the VxE interactions. This method modelled variety performance using covariate information on environments, such as average rainfall and soil type, rather than environment means. Piepho *et al.* (1998) utilised a separable variance matrix for

the VxE interaction effects which allowed for correlations between varieties. The advantage of this method over the regression method which uses environment means is that for suitable covariates (eg. rainfall or soil type), predictions of varietal performance to special environmental conditions can be generated when information on the response of a variety to environmental conditions is available.

The regression approaches proposed by Gogel *et al.* (1995), Nabugoomu *et al.* (1999), and Piepho *et al.* (1998) have the advantage over the method popularised by Finlay & Wilkinson (1963) of being utilised in a mixed model setting. This allows unbalanced data to be analysed and complex covariance models to be used. However, like the Finlay & Wilkinson (1963) approach, these regression methods have the significant disadvantage of often explaining only a small proportion of the VxE interaction (Smith *et al.*, 2005).

The multiplicative models proposed by Piepho (1997) and Smith *et al.* (2001) can be regarded as a random effects analogue of AMMI. The multiplicative model applied to the VxE interaction effects was that associated with the multivariate technique of factor analysis. The variance structure for the VxE effects is known as the factor analytic (FA) structure of order k. When the FA model is applied to the variety effects in each environment, as in the case of Smith *et al.* (2001), they are decomposed into a regression of k hypothetical factors on variety scores along with a lack of fit term for the model. The FA model for the VxE effects differs from traditional random regression problems in that both the coefficients and covariates must be estimated from the data, where the covariates are known as variety scores and the coefficients as environmental loadings. This FA model results in heterogeneity of variety variance and covariance between environments, rather than constraining variance and covariance parameters to be equal as in earlier models. Consequently this model allows more realistic modelling of the VxE effects. Piepho (1997) proposed a similar model, however, with random environment effects rather than random variety effects, resulting in heterogeneity of VxE variance and covariance between varieties. The model proposed by Smith *et al.* (2001) also differed from that proposed by Piepho (1997) in that they allowed a separate variance to be estimated for each environment in the lack of fit component of the FA model, known as specific variances.

There are different trains of thought as to which of the environment and variety main effects, and VxE interaction term should be treated as random effects. Smith *et al.* (2005) hold that this choice should be dependent on the aim of the analysis given the properties of the estimation procedures used in either case. BLUPs best

predict the true variety effects and assuming that the estimates of the variance parameters are sufficiently precise, this also holds true for E-BLUPs. If the aim of the analysis is selection of the best varieties, E-BLUPs are most appropriate and varieties should be treated as random because the rankings of the estimated variety effects need to be as precise as possible with regard to the rankings of the true variety effects (Smith *et al.*, 2005). However, if the aim is to estimate the differences between specific variety effects as precisely as possible variety effects should be treated as fixed because the use of E-BLUPs are inappropriate given that the BLUP of a specific difference is biased. The aim of breeding trials is to select superior varieties and as a result the use of random variety effects is appropriate. Smith *et al.* (2005) highlight that with balanced data and orthogonal analyses, models with fixed or random variety effects would result in identical rankings of these effects; however, these authors prefer the use of random variety effects due to their advantage of more realistic estimates of genetic gain, as such estimates tend to be overly optimistic due to selection bias (Patterson & Silvey, 1980).

The flexibility of the FA model means that for large data sets a substantial number of variance parameters must be estimated. Despite its power, the AI algorithm falls short for FA models when one or more of the estimates of specific variances tend towards zero, resulting in the variance matrix for VxE effects being of less than full rank (termed reduced rank). As such a modified version of the AI algorithm was necessary. Thompson *et al.* (2003) presented a sparse implementation of the AI algorithm for REML estimation of FA variance parameters for the reduced rank case. In addition to allowing for the fitting of reduced rank variance models, this implementation also has the advantage of faster convergence compared to the algorithm proposed by Smith *et al.* (2001) when fitting FA models due to the use of sparse matrices in the estimation process.

Although a one-stage analysis is typically used in Australia for the analysis of early-generation MET analyses and short-term MET analyses, an approximate two-stage approach is used when analysing long-term METs in Australia, along with replicated late-stage MET data in the UK (Welham *et al.*, 2010). Welham *et al.* (2010) suggest that this is due to individual plot data traditionally being difficult to find due to it not being stored electronically, along with the computational difficulties involved with a single-stage analysis when complex variance models are used. Although the more efficient one-stage approach has been recommended (Smith *et al.*, 2005), Welham *et al.* (2010) formally evaluated the one- and two-stage (both weighted and unweighted) approaches using a simulation study. The study considered six statistical models and three different analysis methods, with the

three analysis methods consisting of a single-stage analysis, a weighted two-stage analysis, and an unweighted two-stage analysis. The MET data sets used in the study were simulated from the characteristics and estimated parameters from an Australian wheat breeding program and a set of UK recommended list wheat trials in order to be representative of actual data. The mean square error of prediction and relative genetic gain were used to assess the accuracy of the variety predictions in each environment compared to the effects used to simulate the data.

Welham *et al.* (2010) found that a one-stage approach resulted in the most accurate prediction of variety performance for a range of models. They also found that the unweighted two-stage analysis resulted in a loss of important information regarding estimates of variety performance, however, the weighted two-stage analysis provided an adequate approximation to the single-stage analysis, and may be used for large data sets when the one-stage analysis becomes computationally impractical (Welham *et al.*, 2010). The range of models used to analyse MET data are summarised in Figure 1.1, with distinctions for one- and two-stage analyses. This figure demonstrates how the models evolved and how they relate to each.

1.4 Extensions to the model for VxE effects

The FA model proposed by Smith *et al.* (2001) has been embraced in Australia due to its ability to effectively model the nature of the VxE interaction while also allowing for separate spatial covariance structures for each trial. This model has been applied to a wide range of applications and has been extended to allow for further complexity. The model proposed by Smith *et al.* (2001) made the assumption that varieties were independent, however, more recent work has allowed for the modelling of covariance between varieties. Oakey *et al.* (2006) partitioned the genetic effect of a variety into additive and non-additive effects using pedigree based relationships between varieties in the form of the additive relationship matrix. The additive effects, or breeding values, provide an indication of the potential of a variety as a parent. This analysis consequently allows the selection of varieties as potential parents through the use of additive effects, but also the selection of superior varieties through the combination of additive and non-additive effects (Oakey *et al.*, 2006). This method was considered in both a single trial scenario (Oakey *et al.*, 2006) and MET scenario (Oakey *et al.*, 2007). Oakey *et al.* (2006, 2007) further partitioned the non-additive effects into dominance and residual non-additive effects. However, when the majority of the

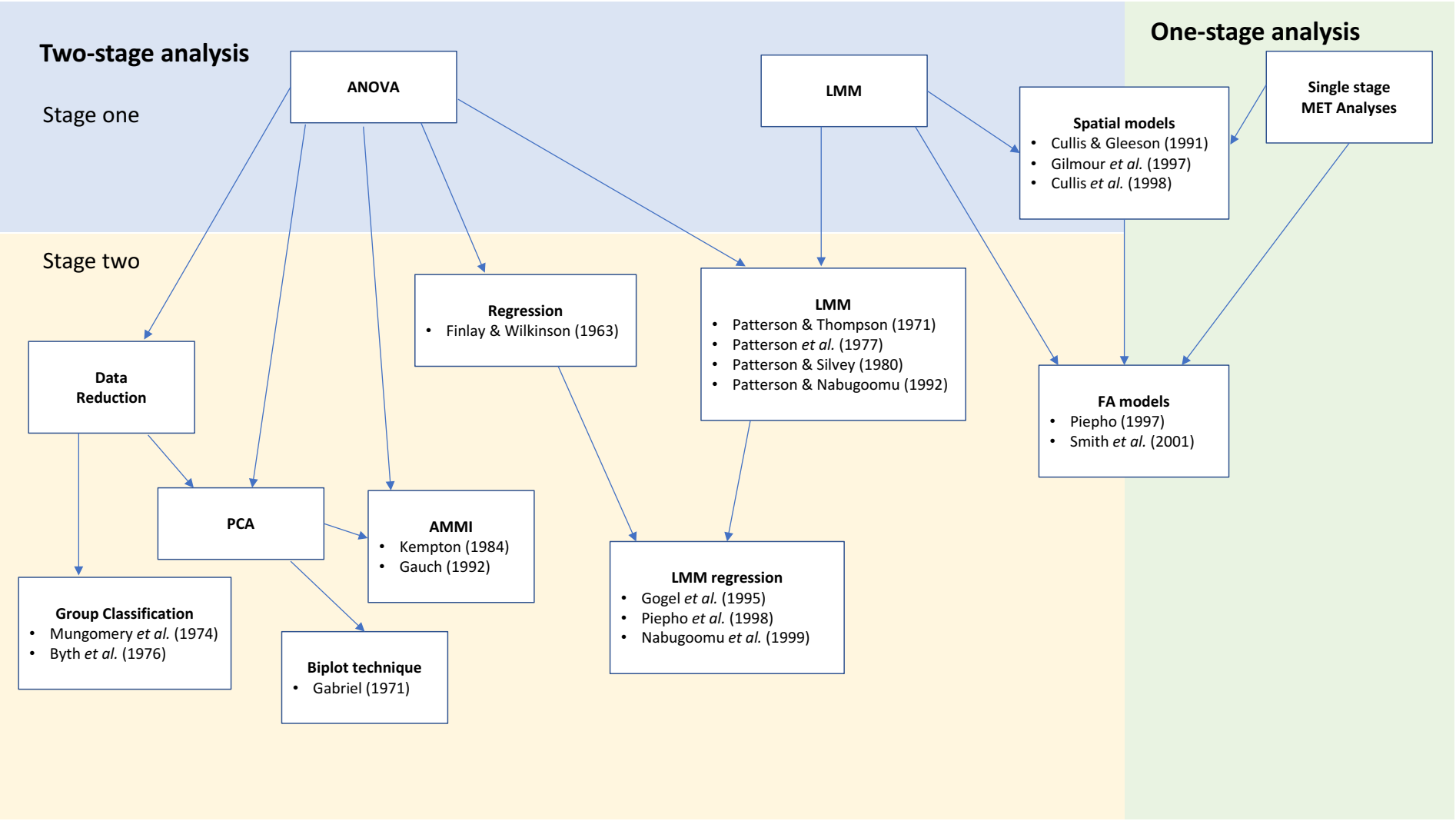


Figure 1.1: Flowchart showing the evolution of models used to analyse multi-environment trial data and how they relate to each other.

varieties are highly inbred, non-additive effects will reflect epistatic effects (interactions between genes within an individual) because inbreeding will largely eliminate dominance.

Although the inclusion of pedigree information has been shown to result in superior model fit (Oakey *et al.*, 2007; Beeck *et al.*, 2010), the elements of the relationship matrix are approximate to true relatedness based on an average proportion of genes in common, and in reality can be quite different to what is expected (Borgognone *et al.*, 2016). The benefits of including pedigree information in the analysis are seen to outweigh the limitations resulting from the approximations necessary in forming this matrix, however, an alternative relationship matrix can be derived from the molecular marker information. Borgognone *et al.* (2016) proposed using an FA model for the analysis of MET data with the genomic relationship matrix rather than the additive relationship matrix to model the relationship between varieties. The form of the genomic relationship matrix still allowed for the partitioning of genetic effects into additive genetic effects and non-additive genetic effects and resulted in lower average prediction error variance of genetic effects (Borgognone *et al.*, 2016).

The use of FA models in the analysis of MET data also allows for investigation into the varied nature of the VxE effects, exploring patterns and irregularities in the data, along with simplifying the results and interpretation. Cullis *et al.* (2010) proposed a number of statistical tools to explore the VxE interactions, including heatmaps which display visually the genetic correlations between environments and clustering methods which group environments in which varieties perform similarly in terms of rank position. These tools can simplify and aid in the interpretation of what can be large numbers of VxE effects. The use of an FA model allows for investigation into a variety's environmental stability for the environments considered in the data, through the regression form of the VxE effects. However, this regression is inherent within the FA model; no post-processing is necessary (Smith *et al.*, 2015).

While the FA model was developed in the context of analysing MET data originating from crop breeding programs, the application of the FA model has been wide and varied. Fox *et al.* (2006) fitted an FA model to plot data assessing grain size from Stage 3 barley trials grown at 25 sites over four years. Utilising these models allowed the authors to gain an improved understanding of VxE effects on expression of grain size, resulting in greater confidence in the selection

of barley varieties which maintain large, stable grain size across a range of environments. In a different approach, Christopher *et al.* (2014) used an FA structure to model variety effects for different traits, estimating the genetic correlations between yield, stay-green traits and normalised difference vegetative index measurements, where these traits were used as environments.

Stefanova & Buirchell (2010) analysed 39 trials of 25 historical lupin varieties using an FA model for the VxE effects. They found that the variety scores for the first two factors of the regression structure of the FA model were representative of genetic gain and stability of varieties. This analysis allowed the authors to identify the varieties which were adapted to low, medium and high rainfall zones and to assess genetic gain over a 31 year period.

Thompson *et al.* (2011) used an FA model to analyse MET data sets measuring the densities of root-lesion nematodes *Pratylenchus thornei* and *Pratylenchus neglectus* in chickpea. The aim of this experiment was to investigate the susceptibility of Australian and international chickpea varieties to these nematodes, allowing for more informed decisions in planning rotations in fields infested with either *P. thornei* or *P. neglectus*. Rodda *et al.* (2016) also modelled *P. thornei* density in chickpeas across a number of trials undertaken in the glasshouse and the field using an FA model. These models enabled the authors to determine that the relative differences in resistance to *P. thornei* identified were highly heritable and also that the genetic correlation between trials in the glasshouse and field were high, meaning that resistance to *P. thornei* in chickpea can be effectively selected in a limited set of environments, saving in labour and resources.

Kelly *et al.* (2007) investigated the accuracy of FA models for trials with large numbers of varieties. FA models were compared with LMMs fitting three models for the VxE effects. These models were a diagonal model, in which the genetic covariance between all pairs of environments is zero, a uniform model, which assumes constant genetic variance and constant genetic covariance across environments, and an unstructured model, which allows a large degree of flexibility in the genetic variance and covariance parameters across environments. The FA models were shown to generally be the model of best fit for a range of data sets taken from early-generation trials in a breeding program. Additionally the superiority of FA models in selection of varieties was shown through a simulation study. The number of varieties considered in this study were 500, 200, and 80, which are representative of the number of varieties per trial included in the early stages of a breeding program.

1.5 Research aims

The number of varieties per trial considered in the late stages of a breeding program are substantially smaller than in the earlier stages and the accuracy of FA models for small numbers of varieties in each trial has not been properly investigated. This prompts the question, does an FA model provide the best estimate of the VxE interaction effects for METs with smaller numbers of varieties? The aims of this project are

1. to determine whether the adequacy of an FA variance structure changes as the number of crop varieties within a trial decreases;
2. to investigate the implications the underlying VxE variance structure has on the accuracy of the FA model; and
3. to investigate the impact the level of varietal concurrence between environments has on the accuracy of the FA model.

Chapter 2

Methods

In the first section of this chapter, Section 2.1, the statistical theory behind linear mixed models will be discussed. This includes the derivation of the residual likelihood and REML score equations which are used to estimate variance parameters, along with the estimation of fixed and random effects. Following this, in Section 2.2 the selection and analysis of the primary data sets used to provide parameters for a simulation study will be detailed. The final section of this chapter, Section 2.3, explains the simulation study that was conducted to investigate the aims of this project.

2.1 Statistical method theory

2.1.1 Linear mixed models

Consider a series of t trials (synonymous with environments) in which m varieties have been grown. If n_j are the number of plots in the j^{th} trial, $n = \sum_{j=1}^{t} n_j$ is the total number of plots. A general linear mixed model for the $n \times 1$ vector of individual plot yields, $\boldsymbol{y}$, ordered as plots within trials, can be written as

$$\boldsymbol{y} = \mathbf{1}_n \mu + \boldsymbol{X}_e \boldsymbol{\tau}_e + \boldsymbol{X}_p \boldsymbol{\tau}_p + \boldsymbol{Z}_g \boldsymbol{u}_g + \boldsymbol{Z}_p \boldsymbol{u}_p + \boldsymbol{e} \tag{2.1}$$

where μ is the overall mean, $\boldsymbol{\tau}_e$ is a $t \times 1$ vector of fixed trial effects with design matrix $\boldsymbol{X}_e$, and $\boldsymbol{u}_g$ is a $mt \times 1$ vector of random variety effects for each trial (ordered as varieties within trials) with design matrix $\boldsymbol{Z}_g$. The vector $\boldsymbol{\tau}_p$ contains trial specific fixed effects with corresponding design matrix $\boldsymbol{X}_p$ and the vector $\boldsymbol{u}_p$ contains trial specific random effects with corresponding design matrix $\boldsymbol{Z}_p$. The $n \times 1$ vector $\boldsymbol{e}$ contains residual effects. The random effects are assumed to follow a Gaussian distribution with mean zero and variance matrix

$$\text{var}\begin{pmatrix} \boldsymbol{u}_g \\ \boldsymbol{u}_p \\ \boldsymbol{e} \end{pmatrix} = \begin{bmatrix} \boldsymbol{G}_g & 0 & 0 \\ 0 & \boldsymbol{G}_p & 0 \\ 0 & 0 & \boldsymbol{R} \end{bmatrix}.$$

The vector $\boldsymbol{u}_g$ represents a two-dimensional array of effects (environments and varieties) and it is assumed the variance structure has a separable form meaning the

variance of the VxE effects can be partitioned into variance due to environments and variance due to varieties such that

$$\boldsymbol{G}_g = \boldsymbol{G}_e \otimes \boldsymbol{G}_v,$$

where $\boldsymbol{G}_e$ and $\boldsymbol{G}_v$ are the $t \times t$ and $m \times m$ symmetric matrices for the variance for environments and varieties respectively. A common assumption is that the variety effects are independent (Smith *et al.*, 2001) such that $\boldsymbol{G}_v = \boldsymbol{I}_m$.

The trial specific effects and the residual effects are assumed to be independent for each trial such that $\boldsymbol{G}_p = \text{diag}(\boldsymbol{G}_{p_j})$ and $\boldsymbol{R} = \text{diag}(\boldsymbol{R}_j)$, where $\boldsymbol{G}_{p_j}$ is the variance matrix for the trial specific random effects at the jth trial and $\boldsymbol{R}_j$ is the residual variance matrix for trial j. The simplest form $\boldsymbol{R}_j$ can take is $\boldsymbol{R}_j = \sigma_j^2 \boldsymbol{I}_{n_j}$ which assumes plot residual effects are independent. However, Smith *et al.* (2001) utilised the approach of Gilmour *et al.* (1997) incorporating a spatial correlation matrix, such that $\boldsymbol{R}_j = \sigma_j^2 \boldsymbol{\Sigma}_{c_j} \otimes \boldsymbol{\Sigma}_{r_j}$, where $\boldsymbol{\Sigma}_{c_j}$ and $\boldsymbol{\Sigma}_{r_j}$ are the correlation matrices for columns and rows respectively, and σ_j^2 is the associated variance.

The traditional mixed model includes a variety main effect and a VxE interaction effect (Patterson *et al.*, 1977), such that

$$\boldsymbol{u}_g = (\boldsymbol{1}_t \otimes \boldsymbol{I}_m)\boldsymbol{u}_v + \boldsymbol{u}_{ge},$$

with these random effects following a Gaussian distribution with mean zero and variance matrix

$$\text{var}\begin{pmatrix} \boldsymbol{u}_v \\ \boldsymbol{u}_{ge} \end{pmatrix} = \begin{bmatrix} \sigma_g^2 \boldsymbol{I}_m & 0 \\ 0 & \sigma_{ge}^2 (\boldsymbol{I}_t \otimes \boldsymbol{I}_m) \end{bmatrix},$$

where σ_g^2 and σ_{ge}^2 are the estimated variance components for variety and the interaction between variety and environment respectively. This leads to

$$\begin{aligned} \text{var}(\boldsymbol{u}_g) &= \text{var}((\boldsymbol{1}_t \otimes \boldsymbol{I}_m)\boldsymbol{u}_v + \boldsymbol{u}_{ge}) \\ &= (\boldsymbol{1}_t \otimes \boldsymbol{I}_m)\text{var}(\boldsymbol{u}_v)(\boldsymbol{1}_t \otimes \boldsymbol{I}_m)' + \text{var}(\boldsymbol{u}_{ge}) \\ &= (\boldsymbol{1}_t \otimes \boldsymbol{I}_m)\sigma_g^2 \boldsymbol{I}_m(\boldsymbol{1}_t' \otimes \boldsymbol{I}_m') + \sigma_{ge}^2(\boldsymbol{I}_t \otimes \boldsymbol{I}_m) \\ &= \sigma_g^2(\boldsymbol{1}_t \otimes \boldsymbol{I}_m)(\boldsymbol{1}_t' \otimes \boldsymbol{I}_m') + \sigma_{ge}^2(\boldsymbol{I}_t \otimes \boldsymbol{I}_m) \\ &= \sigma_g^2(\boldsymbol{1}_t \boldsymbol{1}_t') \otimes (\boldsymbol{I}_m \boldsymbol{I}_m) + \sigma_{ge}^2(\boldsymbol{I}_t \otimes \boldsymbol{I}_m) \\ &= (\sigma_g^2 \boldsymbol{J}_t + \sigma_{ge}^2 \boldsymbol{I}_t) \otimes \boldsymbol{I}_m \qquad (2.2) \\ &\equiv \boldsymbol{G}_e \otimes \boldsymbol{I}_m, \end{aligned}$$

resulting in common genetic variance, $\sigma^2_g + \sigma^2_{ge}$, for all environments and a common genetic covariance, σ^2_g, between pairs of environments. This form of $\boldsymbol{G}_e$ is known as a uniform variance structure.

An alternative, and generally preliminary, model for $\boldsymbol{G}_e$ is an independent model, also known as a diagonal (DIAG) variance model. This model allows for heterogeneity of genetic variance for different environments and assumes zero covariance between pairs of environments, such that

$$\begin{aligned} \text{var}(\boldsymbol{u}_g) &= \begin{bmatrix} \sigma^2_{g_1} & & & \\ 0 & \sigma^2_{g_2} & & \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \sigma^2_{g_t} \end{bmatrix} \otimes \boldsymbol{I}_m \\ &\equiv \boldsymbol{G}_e \otimes \boldsymbol{I}_m, \end{aligned} \tag{2.3}$$

where $\sigma^2_{g_j}$ is the genetic variance for the jth trial.

The most general form of the genetic variance matrix, $\boldsymbol{G}_e$, is an unstructured matrix, which contains $t(t+1)/2$ parameters and can be expressed as

$$\begin{aligned} \text{var}(\boldsymbol{u}_g) &= \begin{bmatrix} \sigma^2_{g_1} & & & \\ \sigma_{g_{12}} & \sigma^2_{g_2} & & \\ \vdots & & \ddots & \\ \sigma_{g_{1t}} & \sigma_{g_{2t}} & \cdots & \sigma^2_{g_t} \end{bmatrix} \otimes \boldsymbol{I}_m \\ &\equiv \boldsymbol{G}_e \otimes \boldsymbol{I}_m, \end{aligned} \tag{2.4}$$

where $\sigma^2_{g_j}$ is the genetic variance at the jth trial, and $\sigma_{g_{ij}}$ is the genetic covariance between trials i and j. Although this model has desirable attributes, it is difficult to estimate from a computational perspective. Furthermore it may be inefficient or unstable for even moderately large numbers of environments (Smith *et al.*, 2001), and Smith *et al.* (2005) suggest this is also true for large numbers of varieties.

The factor analytic model proposed by Smith *et al.* (2001) handles these difficulties and has been shown to be a good approximation to the unstructured matrix (Smith *et al.*, 2005). Smith *et al.* (2001) applied the multiplicative model associated with the multivariate technique of factor analysis to the variety effects in each environment, so that the VxE interaction effects for an FA model of order

k can be written as

$$\boldsymbol{u}_g = (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_m)\boldsymbol{f} + \boldsymbol{\delta},$$

where $\boldsymbol{\Lambda}$ is a $t \times k$ matrix of environment loadings, $\boldsymbol{f}$ is a $mk \times 1$ vector of variety scores, and $\boldsymbol{\delta}$ is a $mt \times 1$ vector of residuals for the VxE model. The joint distribution of $\boldsymbol{f}$ and $\boldsymbol{\delta}$ are assumed to follow a Gaussian distribution with mean zero and variance matrix

$$\text{var}\begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{\delta} \end{pmatrix} = \begin{bmatrix} \boldsymbol{I}_k \otimes \boldsymbol{I}_m & 0 \\ 0 & \boldsymbol{\Psi} \otimes \boldsymbol{I}_m \end{bmatrix},$$

where $\boldsymbol{\Psi}$ is a diagonal $t \times t$ matrix of elements commonly referred to as specific variances. The variance matrix of the variety scores is an identity matrix which means that the scores have a constant variance of 1 and are all independent of each other. The variance of the VxE effects, $\boldsymbol{u}_g$, under this model results in heterogeneity of genetic variance for different environments and heterogeneity of genetic covariance between pairs of environments.

$$\begin{aligned}
\text{var}(\boldsymbol{u}_g) &= \text{var}((\boldsymbol{\Lambda} \otimes \boldsymbol{I}_m)\boldsymbol{f} + \boldsymbol{\delta}) \\
&= (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_m)(\boldsymbol{I}_k \otimes \boldsymbol{I}_m)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_m)' + \boldsymbol{\Psi} \otimes \boldsymbol{I}_m \\
&= (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_m)\boldsymbol{I}_{km}(\boldsymbol{\Lambda}' \otimes \boldsymbol{I}_m') + \boldsymbol{\Psi} \otimes \boldsymbol{I}_m \\
&= (\boldsymbol{\Lambda}\boldsymbol{\Lambda}') \otimes \boldsymbol{I}_m + \boldsymbol{\Psi} \otimes \boldsymbol{I}_m \\
&= (\boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}) \otimes \boldsymbol{I}_m \\
&\equiv \boldsymbol{G}_e \otimes \boldsymbol{I}_m,
\end{aligned} \tag{2.5}$$

where the genetic variance for each trial is given by the diagonal elements of $\boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$ and the genetic covariance between pairs of trials are the off-diagonal elements of $\boldsymbol{\Lambda}\boldsymbol{\Lambda}'$.

2.1.2 Estimation

The variance parameters of the linear mixed model are estimated using REML and the fixed and random effects are estimated as e-BLUEs and e-BLUPs respectively as discussed in Section 1.2. The following section derives the residual likelihood and the REML score equations, along with the estimates of the fixed and random effects.

The model in Equation 2.1 can be rewritten using the general form for a linear mixed model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}, \tag{2.6}$$

where $\boldsymbol{\tau} = \left(\mu, \boldsymbol{\tau}_e', \boldsymbol{\tau}_p'\right)'$ is a vector of fixed effects with design matrix $\boldsymbol{X} = \left[\boldsymbol{1}_n \; \boldsymbol{X}_e \; \boldsymbol{X}_p\right]$ and $\boldsymbol{u} = \left(\boldsymbol{u}_g', \boldsymbol{u}_p'\right)'$ is a vector of random effects with design matrix $\boldsymbol{Z} = \left[\boldsymbol{Z}_g \; \boldsymbol{Z}_p\right]$. The

random effects are assumed to follow a Gaussian distribution, with mean zero and variance matrix

$$\text{var}\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{pmatrix} = \begin{bmatrix} \boldsymbol{G} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R} \end{bmatrix},$$

where $\boldsymbol{G} = \text{diag}(\boldsymbol{G}_e, \boldsymbol{G}_p)$. The vectors of variance parameters associated with the random and residual effects are $\boldsymbol{\gamma} = \left(\boldsymbol{\gamma}_e', \boldsymbol{\gamma}_p'\right)'$ and $\boldsymbol{\phi}$ respectively, such that $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{\gamma})$, $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{\phi})$. The distribution of $\boldsymbol{y}$ is consequently Gaussian with mean $\boldsymbol{X\tau}$ and variance matrix $\boldsymbol{H} = \boldsymbol{R} + \boldsymbol{ZGZ}'$.

Residual maximum likelihood

Verbyla (1990) provided a useful derivation of the likelihood function in which it is partitioned into two independent parts, relating to the treatment contrasts and the residual contrasts. Verbyla (1990) utilised a matrix, $\boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_1 & \boldsymbol{L}_2 \end{bmatrix}$, such that $\boldsymbol{L}_1^{n\times p}$ and $\boldsymbol{L}_2^{n\times(n-p)}$ satisfy the conditions $\boldsymbol{L}_1'\boldsymbol{X} = \boldsymbol{I}_P$ and $\boldsymbol{L}_2'\boldsymbol{X} = \boldsymbol{0}$. $\boldsymbol{L}$ is then used to transform $\boldsymbol{y}$ to $\boldsymbol{L}'\boldsymbol{y}$ such that

$$\boldsymbol{L}'\boldsymbol{y} = \begin{bmatrix} \boldsymbol{L}_1' \\ \boldsymbol{L}_2' \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}.$$

The mean of $\boldsymbol{L}'\boldsymbol{y}$ is found by

$$\text{E}(\boldsymbol{L}'\boldsymbol{y}) = \boldsymbol{L}'\text{E}(\boldsymbol{y}) = \begin{bmatrix} \boldsymbol{L}_1' \\ \boldsymbol{L}_2' \end{bmatrix} \boldsymbol{X\tau} = \begin{bmatrix} \boldsymbol{L}_1'\boldsymbol{X\tau} \\ \boldsymbol{L}_2'\boldsymbol{X\tau} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_p\boldsymbol{\tau} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\tau} \\ \boldsymbol{0} \end{bmatrix}.$$

The variance of $\boldsymbol{L}'\boldsymbol{y}$ is found by

$$\begin{aligned} \text{var}(\boldsymbol{L}'\boldsymbol{y}) =& \boldsymbol{L}'\text{var}(\boldsymbol{y})(\boldsymbol{L}')' = \boldsymbol{L}'\boldsymbol{H}\boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_1' \\ \boldsymbol{L}_2' \end{bmatrix} \boldsymbol{H} \begin{bmatrix} \boldsymbol{L}_1 & \boldsymbol{L}_2 \end{bmatrix} \\ =& \begin{bmatrix} \boldsymbol{L}_1'\boldsymbol{H} \\ \boldsymbol{L}_2'\boldsymbol{H} \end{bmatrix} \begin{bmatrix} \boldsymbol{L}_1 & \boldsymbol{L}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_1 & \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2 \\ \boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_1 & \boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2 \end{bmatrix}. \end{aligned}$$

Consequently the distribution of $\boldsymbol{L}'\boldsymbol{y}$ is

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} \boldsymbol{\tau} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_1 & \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2 \\ \boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_1 & \boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2 \end{bmatrix}\right). \tag{2.7}$$

The likelihood of $\boldsymbol{L}'\boldsymbol{y}$ can be expressed as the product of the conditional likelihood of $\boldsymbol{y}_1$ given $\boldsymbol{y}_2$ and the marginal likelihood of $\boldsymbol{y}_2$. The log-likelihood of these distributions can be expressed similarly such that

$$l_F\left(\tau, \kappa; \boldsymbol{L}'\boldsymbol{y}\right) = l_T\left(\tau, \kappa; \boldsymbol{y}_1|\boldsymbol{y}_2\right) + l_R\left(\kappa; \boldsymbol{y}_2\right). \tag{2.8}$$

Using the standard results in Appendix A.1 and the work of Verbyla (1990) in Appendix A.2 the mean and variance of $\boldsymbol{y}_1|\boldsymbol{y}_2$ can be found

$$\begin{aligned}
\mathrm{E}(\boldsymbol{y}_1|\boldsymbol{y}_2) &= \boldsymbol{\tau} + \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\left(\boldsymbol{y}_2 - \boldsymbol{0}\right) \\
&= \boldsymbol{\tau} + \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{y}_2
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(\boldsymbol{y}_1|\boldsymbol{y}_2) &= \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_1 - \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_1 \\
&= \left[\boldsymbol{L}_1'\left(\boldsymbol{H} - \boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{L}_2'\boldsymbol{H}\right)\boldsymbol{L}_1\right] \\
&= \left[\boldsymbol{L}_1'\left(\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\right)\boldsymbol{L}_1\right] \\
&= \left[\boldsymbol{I}_p(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}\boldsymbol{I}_p'\right] \\
&= (\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}.
\end{aligned}$$

The conditional distribution of $\boldsymbol{y}_1|\boldsymbol{y}_2$ is consequently

$$\boldsymbol{y}_1|\boldsymbol{y}_2 \sim \mathrm{N}\left(\boldsymbol{\tau} + \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{y}_2, (\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}\right)$$

and its corresponding log-likelihood function (excluding constants) is

$$\begin{aligned}
l_T = &-\frac{1}{2}\log|(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}| - \frac{1}{2}\Big(\left(\boldsymbol{y}_1 - \boldsymbol{\tau} - \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{y}_2\right)' \\
&(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})\left(\boldsymbol{y}_1 - \boldsymbol{\tau} - \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{y}_2\right)\Big) \\
= &-\frac{1}{2}\Big(\log|(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}| + \Big(\left(\boldsymbol{L}_1'\boldsymbol{y} - \boldsymbol{\tau} - \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{L}_2'\boldsymbol{y}\right)' \\
&(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})\left(\boldsymbol{L}_1'\boldsymbol{y} - \boldsymbol{\tau} - \boldsymbol{L}_1'\boldsymbol{H}\boldsymbol{L}_2(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2)^{-1}\boldsymbol{L}_2'\boldsymbol{y}\right)\Big)\Big).
\end{aligned}$$

The marginal distribution of $\boldsymbol{y}_2$ is

$$\boldsymbol{y}_2 \sim \mathrm{N}\left(\boldsymbol{0}, \boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2\right)$$

and its associated log-likelihood (excluding constants) is

$$l_R = -\frac{1}{2}\left(\log|\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2| + \boldsymbol{y}'\boldsymbol{L}_2\left(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2\right)^{-1}\boldsymbol{L}_2'\boldsymbol{y}\right).$$

Given that the likelihood of $\boldsymbol{L}'\boldsymbol{y}$ can be expressed as the product of the conditional likelihood of $\boldsymbol{y}_1$ given $\boldsymbol{y}_2$ and the marginal likelihood of $\boldsymbol{y}_2$, its determinant

can be similarly partitioned using the determinant properties in Appendix B.3.

$$\begin{aligned}\log|\boldsymbol{L}'\boldsymbol{H}\boldsymbol{L}| &= \log|\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2| + \log|(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}| \\ \log|\boldsymbol{L}'\boldsymbol{L}| + \log|\boldsymbol{H}| &= \log|\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2| - \log|\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}| \\ \log|\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2| &= \log|\boldsymbol{L}'\boldsymbol{L}| + \log|\boldsymbol{H}| + \log|\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}|\end{aligned}$$

The log-likelihood of the marginal distribution of $\boldsymbol{y}_2$, excluding constants, can be rewritten as

$$\begin{aligned}l_R &= -\frac{1}{2}\left(\log|\boldsymbol{H}| + \log|\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}| + \left(\boldsymbol{y}'\boldsymbol{L}_2\left(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2\right)^{-1}\boldsymbol{L}_2'\boldsymbol{y}\right)\right) \\ &= -\frac{1}{2}\left(\log|\boldsymbol{H}| + \log|\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}| + (\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y})\right),\end{aligned}$$

where $\boldsymbol{P} = \boldsymbol{L}_2\left(\boldsymbol{L}_2'\boldsymbol{H}\boldsymbol{L}_2\right)^{-1}\boldsymbol{L}_2'$. This residual log-likelihood is used to estimate the variance parameters.

The REML solutions for $\boldsymbol{\kappa} = \left(\boldsymbol{\gamma}', \boldsymbol{\phi}'\right)'$ are obtained from the solution of the set of equations

$$U_R(\kappa_i) = \frac{\partial l_R}{\partial \kappa_i} = 0,$$

known as score equations, for $i = 1, \ldots, n_k$, where n_k is the number of variance parameters in $\boldsymbol{\kappa}$.

The score equation for κ_i is given by

$$U_R(\kappa_i) = -\frac{1}{2}\left\{\frac{\partial}{\partial \kappa_i}\left(\log|\boldsymbol{H}|\right) + \frac{\partial}{\partial \kappa_i}\left(\log|\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}|\right) + \frac{\partial}{\partial \kappa_i}\left(\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}\right)\right\}.$$

Using the derivative results in Section B.6,

$$\frac{\partial}{\partial \kappa_i}\left(\log|\boldsymbol{H}|\right) = \operatorname{tr}\left(\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i\right),$$

where $\dot{\boldsymbol{H}}_i = \frac{\partial \boldsymbol{H}}{\partial \kappa_i}$, and

$$\begin{aligned}\frac{\partial}{\partial \kappa_i}\left(\log|\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}|\right) =& \operatorname{tr}\left(\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\frac{\partial}{\partial \kappa_i}\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)\right) \\ =& \operatorname{tr}\left(\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\frac{\partial}{\partial \kappa_i}\left(\boldsymbol{H}^{-1}\right)\boldsymbol{X}\right) \\ =& \operatorname{tr}\left(\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\left(-\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i\boldsymbol{H}^{-1}\right)\boldsymbol{X}\right) \\ =& -\operatorname{tr}\left(\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i\boldsymbol{H}^{-1}\boldsymbol{X}\right).\end{aligned}$$

Using the results in Appendix A.3,

$$\begin{aligned}\frac{\partial}{\partial \kappa_i}\left(\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}\right) =& \boldsymbol{y}'\frac{\partial}{\partial \kappa_i}\left(\boldsymbol{P}\right)\boldsymbol{y} \\ =& \boldsymbol{y}'\boldsymbol{P}\dot{\boldsymbol{H}}_i\boldsymbol{P}\boldsymbol{y}.\end{aligned}$$

Combining these results, the score equation for κ_i is

$$\begin{aligned}U_R(\kappa_i) &= -\frac{1}{2}\left\{\mathrm{tr}\left(\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i\right) - \mathrm{tr}\left(\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i\boldsymbol{H}^{-1}\boldsymbol{X}\right) - \boldsymbol{y}'\boldsymbol{P}\dot{\boldsymbol{H}}_i\boldsymbol{P}\boldsymbol{y}\right\} \\ &= -\frac{1}{2}\left\{\mathrm{tr}\left(\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i - \boldsymbol{H}^{-1}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}\dot{\boldsymbol{H}}_i\right) - \boldsymbol{y}'\boldsymbol{P}\dot{\boldsymbol{H}}_i\boldsymbol{P}\boldsymbol{y}\right\} \\ &= -\frac{1}{2}\left\{\mathrm{tr}\left(\left(\boldsymbol{H}^{-1} - \boldsymbol{H}^{-1}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}\right)\dot{\boldsymbol{H}}_i\right) - \boldsymbol{y}'\boldsymbol{P}\dot{\boldsymbol{H}}_i\boldsymbol{P}\boldsymbol{y}\right\} \\ &= -\frac{1}{2}\left\{\mathrm{tr}\left(\boldsymbol{P}\dot{\boldsymbol{H}}_i\right) - \boldsymbol{y}'\boldsymbol{P}\dot{\boldsymbol{H}}_i\boldsymbol{P}\boldsymbol{y}\right\}. \qquad (2.9)\end{aligned}$$

Generally solving the system of equations $U_R(\boldsymbol{\kappa}) = 0$ requires an iterative method (Smith *et al.*, 2001). One such method is the average information (AI) algorithm (Gilmour *et al.*, 1995). The AI algorithm is a modified Fisher scoring algorithm which uses an approximate average of the observed and expected information matrix rather than the expected information matrix. As shown in Smith *et al.* (2005), given an estimate of $\boldsymbol{\kappa} = \boldsymbol{\kappa}^{(m)}$, it can be updated as

$$\boldsymbol{\kappa}^{(m+1)} = \boldsymbol{\kappa}^{(m)} + \left[\boldsymbol{I}^{(m)}\right]^{-1} U_R(\boldsymbol{\kappa}^{(m)}),$$

where $\boldsymbol{I}^{(m)}$ is the average information matrix, $\boldsymbol{I}$, for iteration (m) given by

$$\boldsymbol{I} = \frac{1}{2}\boldsymbol{Q}'\boldsymbol{P}\boldsymbol{Q}$$

and the columns of $\boldsymbol{Q}$ are working variables associated with $\boldsymbol{\kappa}$ given by

$$\boldsymbol{q}_{\boldsymbol{\kappa}_i} = \dot{\boldsymbol{H}}_i\boldsymbol{P}\boldsymbol{y}.$$

Mixed model equations

Estimates of the fixed and random effects, $\boldsymbol{\tau}$ and $\boldsymbol{u}$, can be found by maximising a function derived from the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{u}$, such that

$$\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{u}\end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix}\boldsymbol{X}\boldsymbol{\tau}\\ \boldsymbol{0}\end{bmatrix}, \begin{bmatrix}\boldsymbol{H} & \boldsymbol{Z}\boldsymbol{G}\\ \boldsymbol{G}\boldsymbol{Z}' & \boldsymbol{G}\end{bmatrix}\right).$$

The log-density function for the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{u}$ is given by

$$\log f_Y(\boldsymbol{y}|\boldsymbol{u}) + \log f_U(\boldsymbol{u}).$$

The distribution of $\boldsymbol{u}$ is given by

$$\boldsymbol{u} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{G})$$

and its associated log-density function is

$$\log f_u = -\frac{1}{2}\left(\log|\boldsymbol{G}| + \left(\boldsymbol{u}'\boldsymbol{G}^{-1}\boldsymbol{u}\right)\right).$$

Using the findings in Appendix A.1 the mean and variance of $\boldsymbol{y}|\boldsymbol{u}$ can be found

$$\begin{aligned}\mathrm{E}(\boldsymbol{y}|\boldsymbol{u}) &= \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{G}\boldsymbol{G}^{-1}(\boldsymbol{u} - \boldsymbol{0}) \\ &= \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u}\end{aligned}$$

$$\begin{aligned}\mathrm{var}(\boldsymbol{y}|\boldsymbol{u}) &= \boldsymbol{H} - \boldsymbol{Z}\boldsymbol{G}\boldsymbol{G}^{-1}\boldsymbol{G}\boldsymbol{Z}' \\ &= \boldsymbol{H} - \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' \\ &= \boldsymbol{R}.\end{aligned}$$

The conditional distribution of $\boldsymbol{y}|\boldsymbol{u}$ is consequently

$$\boldsymbol{y}|\boldsymbol{u} \sim \mathrm{N}(\boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u}, \boldsymbol{R})$$

and its associated log-density function is

$$\log f_Y = -\frac{1}{2}\left(\log|\boldsymbol{R}| + \left((\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\tau} - \boldsymbol{Z}\boldsymbol{u})'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\tau} - \boldsymbol{Z}\boldsymbol{u})\right)\right).$$

The log-density function for the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{u}$ is given by

$$\begin{aligned}\log f_u + \log f_Y = &-\frac{1}{2}\left(\log|\boldsymbol{G}| + \left(\boldsymbol{u}'\boldsymbol{G}^{-1}\boldsymbol{u}\right)\right) \\ &-\frac{1}{2}\left(\log|\boldsymbol{R}| + \left((\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\tau} - \boldsymbol{Z}\boldsymbol{u})'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\tau} - \boldsymbol{Z}\boldsymbol{u})\right)\right). \qquad (2.10)\end{aligned}$$

Differentiating Equation 2.10 with respect to $\boldsymbol{\tau}$ using results in Appendix B.6 and equating to zero results in

$$\begin{aligned}-2\boldsymbol{X}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\tau}} - \boldsymbol{Z}\tilde{\boldsymbol{u}}) &= 0 \\ -\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{y} + \boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X}\hat{\boldsymbol{\tau}} + \boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{Z}\tilde{\boldsymbol{u}} &= 0\end{aligned}$$

$$X'R^{-1}X\hat{\tau} + X'R^{-1}Z\tilde{u} = X'R^{-1}y, \tag{2.11}$$

where $\hat{\tau}$ and $\tilde{u}$ are the estimates of τ and u respectively. Differentiating Equation 2.10 with respect to u using results in Appendix B.6 and equating to zero results in

$$\begin{aligned} 2G^{-1}\tilde{u} - 2Z'R^{-1}(y - X\hat{\tau} - Z\tilde{u}) &= 0 \\ G^{-1}\tilde{u} - Z'R^{-1}y + Z'R^{-1}X\hat{\tau} + Z'R^{-1}Z\tilde{u} &= 0 \\ Z'R^{-1}X\hat{\tau} + \left(G^{-1} + Z'R^{-1}Z\right)\tilde{u} &= Z'R^{-1}y. \end{aligned} \tag{2.12}$$

Equations 2.11 and 2.12 are known as the mixed model equations and are more commonly expressed using matrix notation as

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & (G^{-1} + Z'R^{-1}Z) \end{bmatrix} \begin{bmatrix} \hat{\tau} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}. \tag{2.13}$$

Rearranging Equation 2.12 gives

$$\tilde{u} = \left(G^{-1} + Z'R^{-1}Z\right)^{-1}\left(Z'R^{-1}y - Z'R^{-1}X\hat{\tau}\right) \tag{2.14}$$

and substituting Equation 2.14 into 2.11 gives

$$\begin{aligned} X'R^{-1}X\hat{\tau} + X'R^{-1}Z\left(G^{-1} + Z'R^{-1}Z\right)^{-1}\left(Z'R^{-1}y - Z'R^{-1}X\hat{\tau}\right) &= X'R^{-1}y \\ X'R^{-1}X\hat{\tau} - X'R^{-1}Z\left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}X\hat{\tau} &= X'R^{-1}y - X'R^{-1}Z\left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}y \\ X'\left(R^{-1} - R^{-1}Z\left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}\right)X\hat{\tau} &= X'\left(R^{-1} - R^{-1}Z\left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}\right)y. \end{aligned}$$

Using the identity in Appendix B.4, this can be rewritten in terms of H^{-1}, such that

$$\begin{aligned} X'H^{-1}X\hat{\tau} &= X'H^{-1}y \\ \hat{\tau} &= \left(X'H^{-1}X\right)^{-1}X'H^{-1}y. \end{aligned} \tag{2.15}$$

Substituting Equation 2.15 into 2.14 gives

$$\begin{aligned} \tilde{u} &= \left(G^{-1} + Z'R^{-1}Z\right)^{-1}\left(Z'R^{-1}y - Z'R^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}y\right) \\ &= \left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}\left(I - X\left(X'H^{-1}X\right)^{-1}X'H^{-1}\right)y \\ &= \left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}HPy \\ &= \left[G - GZ'\left(R + ZGZ'\right)^{-1}ZG\right]Z'R^{-1}HPy \end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{GZ'}\left[\boldsymbol{I} - \boldsymbol{H}^{-1}\boldsymbol{ZGZ'}\right]\boldsymbol{R}^{-1}\boldsymbol{HPy} \\
&= \boldsymbol{GZ'}\left[\boldsymbol{I} - \boldsymbol{H}^{-1}(\boldsymbol{H} - \boldsymbol{R})\right]\boldsymbol{R}^{-1}\boldsymbol{HPy} \\
&= \boldsymbol{GZ'}\left[\boldsymbol{I} - \boldsymbol{I} + \boldsymbol{H}^{-1}\boldsymbol{R}\right]\boldsymbol{R}^{-1}\boldsymbol{HPy} \\
&= \boldsymbol{GZ'}\boldsymbol{H}^{-1}\boldsymbol{RR}^{-1}\boldsymbol{HPy} \\
&= \boldsymbol{GZ'}\boldsymbol{Py}. \qquad (2.16)
\end{aligned}
$$

Utilising Equations 2.15 and 2.16, in conjunction with the score equations shown in Equation 2.9, the fixed and random effects can be estimated. The VxE effects can be used to select the best performing varieties in each environment and the estimates of genetic variance and genetic covariance provide important information about the nature of the VxE effects and whether varieties perform similarly in certain environments. A practical example of the implementation of this process will be given in the following section.

2.2 Primary data sets and estimation of simulation parameters

Four primary data sets were selected and analysed to provide parameters to be used as the basis for a simulation study. Using these parameters 20 000 data sets were simulated to assess the performance of FA models when data sets contain trials with small to moderate numbers of varieties. The methods and results of this preliminary analysis are included in this section as they represent a necessary step towards the simulation of the data sets and the subsequent analysis of these simulated data sets, which is the focus of this research.

2.2.1 Selection of primary data sets

Four data sets originating from crop breeding programs were analysed, demonstrating the methods typically employed in the analysis of MET data. These data sets were chosen as they are representative of different crop types for late stage breeding programs in Australia. The selected data sets consist of data from the late stages of a wheat breeding program, two chickpea breeding programs (one considering Desi chickpeas and the other Kabuli chickpeas), and a mungbean breeding program. This section details the analysis of these data sets along with the results of these analyses.

The chosen data sets are summarised in Table 2.1 and vary in the number of years, trials and varieties they consider. The data set from the wheat breeding program is the largest, containing data from 52 trials and 128 unique varieties. In contrast the mungbean data set is much smaller, considering only 9 trials and 78

unique varieties. The data sets differ in the level of varietal concurrence between trials. A measure of the average varietal concurrence for each data set is given by the number of unique trial-by-variety combinations present in the data, divided by the potential number of unique combinations given the number of trials and varieties. Both the Desi chickpea and mungbean data sets have an average of approximately 60% of varieties in common between trials, while the remaining two data sets have much lower concurrence with an average of 38% and 28% for wheat and Kabuli chickpea respectively. The poor concurrence for these two data sets is driven by the number of years considered. All of the data sets have nearly perfect concurrence between trials from the same year, however, a number of varieties are often removed from the program each year and replaced with new varieties. This results in increasingly poor concurrence as more years are considered, and it is for this reason breeding programs tend to use a moving window of approximately five years for the analysis of multi-environment trial data. The analyses for all data sets were performed using yield as the dependent variable.

2.2.2 *Analysis of primary data sets*

The data sets were each analysed separately using standard analysis procedures for multi-environment trials. The yield data from each trial was initially analysed separately using a linear mixed model, equivalent to that shown in Equation 2.1. The model fitted an overall mean and a random variety effect, along with any significant trial specific covariates (such harvesting problems or bird damage). Using the approach discussed in Section 1.2 (Smith *et al.*, 2005), experimental design terms were included as random effects and spatial variation was modelled following the procedure of Gilmour *et al.* (1997). The residual (plot) effects were modelled using a separable variance structure, with a first order autoregressive model used in both the row and column directions. Diagnostic tools were used to assess spatial variation in the field, with formal tests used to determine whether terms accounting for this spatial variation should be included in the model.

Table 2.1: Summary of data sets used as sources of parameter estimates for data simulation.

Data set	Years	Trials	Varieties	Varieties per trial			Concurrence (%)	
				Min	Mean	Max	Overall	Per year
Desi Chickpea	2	18	50	30	30	30	60	100
Kabuli Chickpea	6	39	223	51	63	85	28	99
Mungbean	2	9	78	25	52	73	66	100
Wheat	4	52	128	42	49	60	38	99

Following the single-trial analyses, trials were combined in a multi-environment trial (MET) analysis. Spatial and design terms along with residual effects were modelled separately for each trial, using the models found in the single-trial analysis. Initially the variety by environment (trial) effects were modelled independently for each trial. This model was then extended to a factor analytic (FA) model of order 1. Higher order FA models were subsequently fitted to the VxE effects given they provided an improvement on the previous model and the inequality

$$k \leq \frac{2t + 1 - \sqrt{8t + 1}}{2} \tag{2.17}$$

was satisfied, where k is the order of FA model and t is the number of trials in the data. The Akaike information criterion (AIC) and log-likelihood ratio test were used to determine the order of the most parsimonious FA model within each data set, where all comparisons were made between nested models. Both of these are measures of goodness-of-fit, where smaller relative AIC values and a significant log-likelihood ratio test under a Chi-square distribution indicate a more parsimonious model. A summary of the models for each data set is shown in Table 2.2. All analyses were undertaken using ASReml-R (Butler *et al.*, 2009) in the R software environment (R Core Team, 2016).

2.2.3 *Analysis results of data sets*

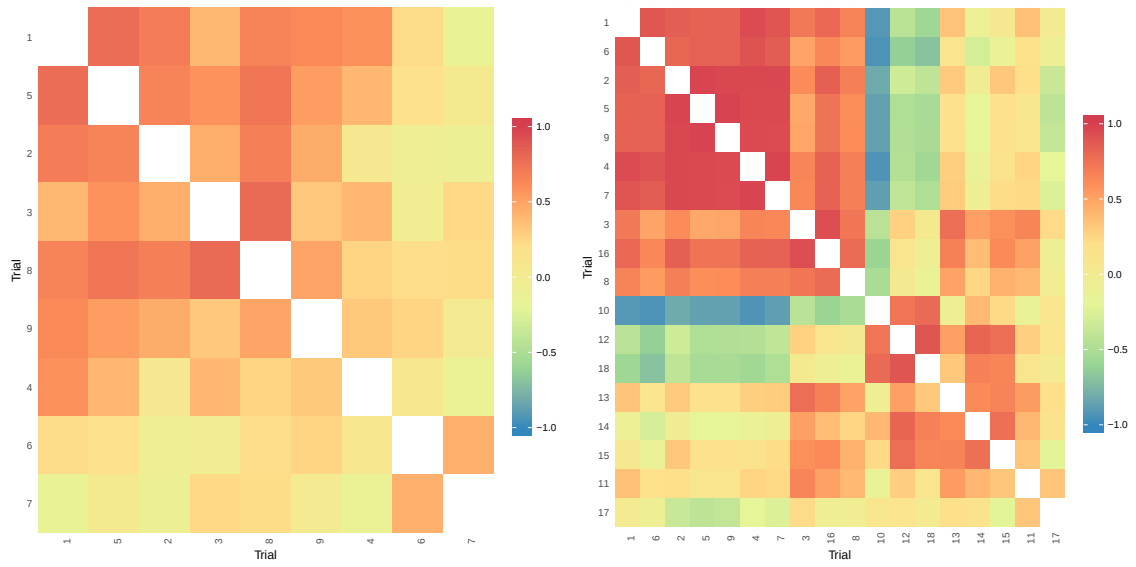
Predictions of varietal performance at each environment were produced from the analyses as e-BLUPs. Heatmaps were produced for each data set which display graphically the genetic correlations between environments as shown in Figures 2.1 and 2.2. These correlations range from 1 to -1, where 1 indicates the ranking of varieties between two environments is nearly identical, 0 indicates very little or no agreement in the ranking of varieties between two environments, and -1 indicates a reversal in the ranking of varieties between two environments. As evident in Figures 2.1 and 2.2, the nature of the VxE effects differ between the data sets. The mungbean and Kabuli chickpea data sets contain mostly low to high positive correlations, while the Desi chickpea and wheat data sets contain high negative correlations as well. Summaries of the trial means, genetic variance and error variance are summarised in Tables 2.3, 2.4, 2.5, and 2.6. The wheat data set had a much larger range of trial mean yields than the other data sets, ranging from 1.3 – 7.3 t/ha. The Desi chickpea trial mean yields ranged from 1.1 – 3.8 t/ha, and the Kabuli chickpea trial mean yields ranged from 0.4 – 3.8 t/ha, while the mungbean trial mean yields ranged from 0.5 – 1.8 t/ha. Only one mungbean trial and three Desi chickpea trials had less genetic variance than error variance, while 32 and 24 of the wheat and Kabuli chickpea trials respectively had less genetic

Table 2.2: Summary of models used to analyse selected data sets, showing the number of parameters estimated in the model (n), the Akaike information criterion (AIC), given here as the difference between the model and the model with the smallest AIC in each data set, the log-likelihood (Logl), and percent of genetic variance accounted for by the FA components in the model (% vaf).

Data set	Model	n	AIC	Logl	% vaf
Mungbean	DIAG	48	76.85	1908.30	
	FA1	57	8.07	1951.69	45.10
	FA2	64	14.24	1955.60	62.38
	FA3	70	9.26	1964.09	60.05
	FA4*	75	**0.00**	1973.72	85.95
	FA5	78	1.54	1975.95	85.73
Kabuli chickpea	DIAG	201	672.79	7813.47	
	FA1	239	147.99	8113.86	62.05
	FA2	275	75.04	8186.34	68.68
	FA3	311	51.29	8234.21	74.97
	FA4	343	17.75	8282.98	79.63
	FA5	374	12.12	8316.80	81.95
	FA6*	402	**0.00**	8350.86	87.14
Desi chickpea	DIAG	84	372.98	2149.25	
	FA1	102	40.87	2333.31	73.14
	FA2	116	26.64	2354.43	86.87
	FA3*	129	**0.00**	2380.75	92.22
Wheat	DIAG	209	1144.51	6374.67	
	FA1	260	512.02	6741.92	39.63
	FA2	305	327.82	6879.01	50.07
	FA3	353	290.69	6945.58	61.43
	FA4	398	158.8	7056.52	70.97
	FA5	438	68.94	7141.45	85.48
	FA6*	481	**0.00**	7218.92	90.97

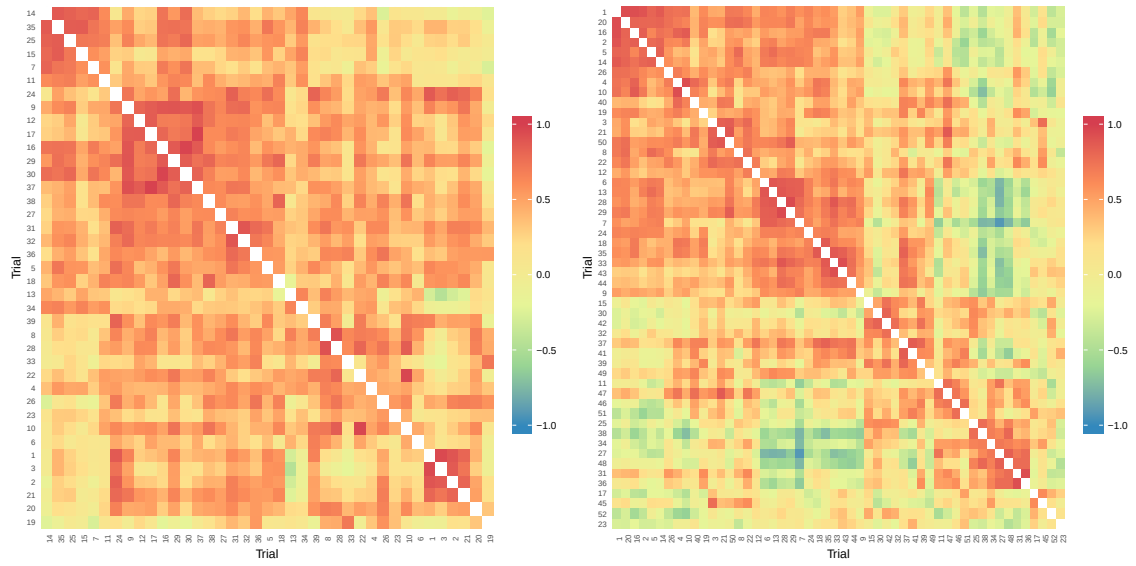
* Model which offered the best fit to the data according to a log-likelihood ratio test.

variance than error variance. There were also two Desi chickpea trials that had much more (over 10 times) genetic variance than error variance. This preliminary analysis of these four data sets resulted in the parameters that formed the basis of the simulation study as discussed in the next section.



(a) **(b)**

Figure 2.1: Heatmap showing the genetic correlations between trials from the analysis of the (a) mungbean and (b) Desi chickpea data sets.



(a) **(b)**

Figure 2.2: Heatmap showing the genetic correlations between trials from the analysis of the (a) Kabuli chickpea and (b) wheat data sets.

Table 2.3: Analysis summary of mungbean data set, showing estimated trial means, genetic variances and error variances.

Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j	Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j
1	0.792	0.032	0.006	6	1.699	0.028	0.023
2	1.397	0.070	0.033	7	0.854	0.015	0.018
3	0.848	0.042	0.014	8	1.752	0.063	0.038
4	0.952	0.017	0.007	9	0.495	0.010	0.009
5	0.714	0.010	0.007				

Table 2.4: Analysis summary of Desi chickpea data set, showing estimated trial means, genetic variances and error variances.

Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j	Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j
1	1.797	0.025	0.016	10	1.560	0.046	0.228
2	1.269	0.034	0.003	11	2.077	0.010	0.003
3	2.304	0.014	0.015	12	3.761	0.035	0.020
4	1.080	0.007	0.005	13	1.759	0.009	0.005
5	1.958	0.087	0.030	14	2.570	0.013	0.011
6	1.926	0.085	0.027	15	3.039	0.020	0.009
7	1.626	0.059	0.003	16	1.349	0.008	0.062
8	2.183	0.020	0.009	17	1.727	0.015	0.006
9	1.359	0.048	0.039	18	3.837	0.030	0.018

Table 2.5: Analysis summary of Kabuli chickpea data set, showing estimated trial means, genetic variances and error variances.

Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j	Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j
1	1.706	0.057	0.093	21	2.961	0.036	0.047
2	2.261	0.200	0.103	22	1.024	0.049	0.010
3	2.923	0.337	0.157	23	1.759	0.013	0.016
4	1.796	0.013	0.016	24	3.689	0.078	0.025
5	3.849	0.077	0.079	25	1.187	0.021	0.013
6	2.820	0.023	0.064	26	1.415	0.013	0.007
7	2.479	0.033	0.076	27	2.732	0.032	0.046
8	2.491	0.048	0.143	28	1.764	0.029	0.050
9	3.677	0.074	0.035	29	2.398	0.012	0.055
10	1.129	0.006	0.012	30	2.017	0.012	0.134
11	1.482	0.011	0.010	31	2.577	0.016	0.016
12	3.579	0.028	0.032	32	1.329	0.007	0.014
13	0.972	0.010	0.015	33	2.662	0.047	0.051
14	1.359	0.013	0.035	34	0.424	0.011	0.013
15	1.045	0.019	0.017	35	0.964	0.021	0.019
16	2.463	0.028	0.028	36	2.374	0.055	0.033
17	1.062	0.006	0.012	37	2.454	0.013	0.030
18	1.257	0.010	0.010	38	1.833	0.027	0.015
19	2.192	0.055	0.140	39	1.282	0.011	0.013
20	2.380	0.220	0.086				

Table 2.6: Analysis summary of wheat data set, showing estimated trial means, genetic variances and error variances.

Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j	Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j
1	2.553	0.015	0.019	27	1.700	0.021	0.026
2	2.149	0.014	0.024	28	2.981	0.191	0.073
3	7.347	0.212	0.069	29	5.153	0.477	0.108
4	1.969	0.010	0.058	30	1.760	0.044	0.123
5	3.921	0.084	0.045	31	1.324	0.004	0.059
6	4.802	0.037	0.023	32	1.691	0.177	0.096
7	4.180	0.029	0.021	33	2.603	0.042	0.045
8	3.156	0.052	0.073	34	2.901	0.020	0.043
9	2.399	0.032	0.017	35	2.124	0.019	0.048
10	2.467	0.011	0.017	36	1.626	0.004	0.167
11	1.303	0.005	0.062	37	2.392	0.009	0.026
12	4.063	0.017	0.035	38	2.231	0.048	0.068
13	3.178	0.027	0.225	39	2.437	0.016	0.047
14	2.323	0.054	0.031	40	2.386	0.027	0.237
15	2.866	0.014	0.052	41	2.618	0.014	0.168
16	2.099	0.016	0.021	42	3.406	0.024	0.053
17	2.711	0.155	0.083	43	2.111	0.024	0.074
18	2.423	0.011	0.040	44	1.380	0.015	0.009
19	3.120	0.027	0.126	45	3.860	0.293	0.080
20	5.556	0.170	0.197	46	3.086	0.030	0.019
21	6.844	0.649	0.426	47	3.228	0.032	0.035
22	4.223	0.044	0.104	48	1.733	0.026	0.035
23	3.536	0.139	0.144	49	4.205	0.063	0.048
24	2.802	0.023	0.021	50	4.514	0.256	0.124
25	1.900	0.335	0.055	51	2.816	0.031	0.030
26	1.925	0.012	0.025	52	4.235	0.104	0.097

2.3 Simulation study

The focus of this research is a simulation study which was conducted to investigate the adequacy of the FA model structure in comparison to other model structures for estimating genetic performance in field trials with small numbers of varieties.

2.3.1 Simulation of data

Forty data generation models were considered, comprising the factorial combination of **five** variance-covariance structures typical of MET data sets, **four** numbers of varieties per trial (10, 15, 25, and 50), and **two** levels of varietal concurrence (perfect and partial). The five variance-covariance structures were drawn from the analyses of the four data sets discussed in Section 2.2.1 and shown in Table 2.1 and in addition the structure from the South Australia barley MET used by Kelly *et al.* (2007) was also used. This barley MET data set is drawn from the early stages of the South Australian barley breeding program and consisted of 10 trials and 480 varieties. Perfect concurrence (in the simulation study) meant that all variety by trial combinations were present, while partial concurrence meant that half of the variety by trial combinations were present. For each of the 40 data generation models, 500 simulations were conducted. The structure of the data generation models is illustrated in the flowchart shown in Figure 2.3.

The number of varieties per trial in the data sets that four of the variance-covariance structures are taken from range from 25 to 85 which is higher than the numbers of varieties considered in this simulation study. However, this is not considered to be a limiting factor given the interest in this study is how well the parameters in the study are estimated, not how well the true parameters in the original data set are estimated. These estimates are merely to set up a realistic simulation study, but in practice these numbers could have been determined using other methods. The models used to analyse these data sets have been shown to be adequate for larger numbers of varieties per trial (>50) (Kelly *et al.*, 2007) and as such it is assumed that these parameters are a reasonably good representation of the underlying patterns in the data sets.

Eight designs were generated for each data set as completely randomised designs with three replicates. Each design corresponded to a variety number per trial (10, 15, 25, and 50) by varietal concurrence level (perfect and partial) combination. For the designs with partial concurrence, the varieties present at each trial were selected so that 50% of the total variety by trial combinations were

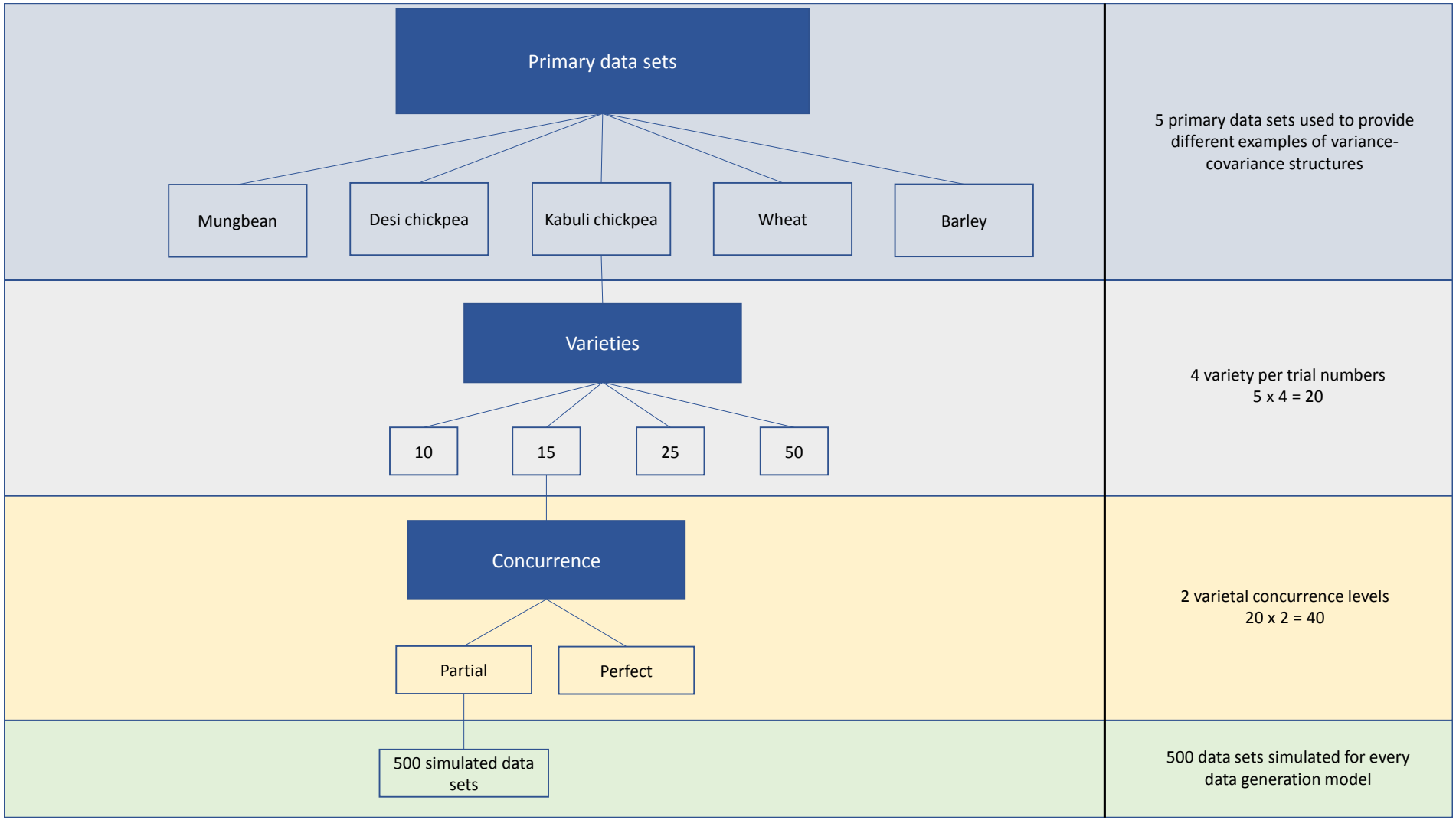


Figure 2.3: Flowchart demonstrating how the 40 data generation models were formed for the simulation study.

present.

Yield data was simulated according to the equation

$$\boldsymbol{y} = \mathbf{1}_n \mu + \boldsymbol{X}_e \boldsymbol{\tau}_e + \boldsymbol{Z}_g \boldsymbol{u}_g + \boldsymbol{e}, \tag{2.18}$$

where μ is the overall mean, $\boldsymbol{\tau}_e$ is a $t \times 1$ vector of trial effects, and $\boldsymbol{u}_g$ is a $mt \times 1$ vector of variety effects for each trial, generated from a Gaussian distribution with mean zero and variance matrix $\boldsymbol{G}_e \otimes \boldsymbol{I}_m$. The $n \times 1$ vector $\boldsymbol{e}$ contains residual effects which were generated from a Gaussian distribution with mean zero and variance matrix $\oplus \sigma_j^2 \boldsymbol{I}_{n_j}$. The matrix $\boldsymbol{X}_e$ is a $n \times t$ design matrix for the vector of trial effects and the matrix $\boldsymbol{Z}_g$ is a $n \times mt$ design matrix for the vector of variety effects for each trial, both formed from the generated designs.

The parameters estimated from the analysis of the four primary data sets and the parameters from the barley data set in Kelly *et al*. (2007) (trial means, genetic variance matrices and error variances) were used in Equation 2.18 to simulate the yield data. These parameters are shown in Tables 2.3, 2.4, 2.6, 2.5, and 2.7 and the genetic covariance matrix is shown in graphical form as a correlation matrix in Figures 2.1, 2.2, and 2.4. For each data set by number of varieties per trial combination, the same VxE effects were used when generating the yield data for both the perfect and partial concurrence. The code used to simulate the data is given in Appendix C.1.

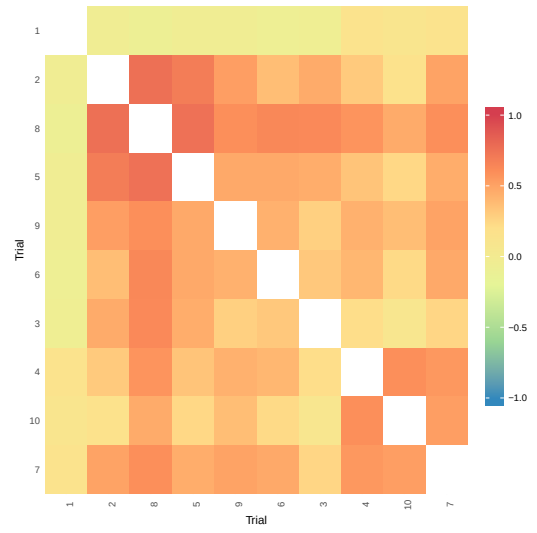


Figure 2.4: Heatmap showing the genetic correlations between trials from the barley data set from Kelly *et al.* (2007).

Table 2.7: Summary of barley data set from Kelly *et al.* (2007), showing estimated trial means, genetic variances and error variances.

Trial	Mean $\mu + \tau_{ej}$	Genetic variance $\sigma^2_{g_j}$	Error variance σ^2_j
1	1.300	0.039	0.011
2	1.900	0.061	0.020
3	0.830	0.019	0.014
4	2.100	0.038	0.032
5	2.100	0.089	0.067
6	1.600	0.098	0.050
7	3.500	0.093	0.055
8	0.850	0.010	0.022
9	2.100	0.064	0.043
10	2.800	0.101	0.056

2.3.2 *Analysis of simulated data*

The simulated data sets were analysed using four model types for the variance of the VxE effects ($\boldsymbol{G}_e \otimes \boldsymbol{I}_m$). The **four model types** used were a **uniform model**, a **diagonal model**, an **unstructured model**, and **FA models** as discussed in Section 2.1.1 with the forms shown in Equations 2.2, 2.3, 2.4, and 2.5 respectively. All analyses were undertaken using ASReml-R (Butler *et al.*, 2009) in the R software environment (R Core Team, 2016).

A set of VxE effects were generated from each $\boldsymbol{G}_e$ model, resulting in 500 sets of effects for each of the 40 data generation models. The mean square error of prediction (MSEP) was calculated for each set of VxE effects,

$$\text{MSEP} = \frac{1}{mt}\sum_{j=1}^{t}\sum_{i=1}^{m}(\tilde{\boldsymbol{u}}_{g_{ij}} - \boldsymbol{u}_{g_{ij}})^2,$$

where $\tilde{\boldsymbol{u}}_{g_{ij}}$ are the predicted VxE interaction effects from each model and $\boldsymbol{u}_{g_{ij}}$ are the VxE effects used to simulate the data. The correlation between $\tilde{\boldsymbol{u}}_{g_{ij}}$ and $\boldsymbol{u}_{g_{ij}}$, the AIC, and log-likelihood (in addition to the MSEP) were also recorded for each analysis.

Although there are four model types applied to each simulated data set, in practice there are actually a number of FA models sequentially fitted to the data, each fitting a different number of factors. The number of FA models fitted to a given simulated data set depended on the number of trials in the data set and how well the FA models fitted the data. Higher order FA models were sequentially fitted to the data set in question until either the model did not offer a significant improvement on the previous FA model according to a log-likelihood ratio test, or the condition in Equation 2.17 was not satisfied. This means that the highest order of FA model differed between data-generation models, but also between simulations of the same data-generation model.

To best compare the FA models to the other models, firstly the MSEPs and correlations were each averaged over the simulations for each FA model, resulting in an average MSEP and an average correlation for each data generation model for each FA model. Following this, the most parsimonious FA model for each simulation according to a log-likelihood ratio test was selected and the MSEPs and correlations for only these models were averaged over simulations, resulting in one average MSEP and correlation for each data generation model for FA models as a whole. These average MSEPs and average correlations for FA models as a single category were then compared to MSEPs and correlations averaged over the simulations for the diagonal, uniform and unstructured models for each data generation model. The average MSEPs and correlations were used to compare $\boldsymbol{G}_e$ models and assess their accuracy for different numbers of varieties and concurrence levels. The code used to analyse the simulated data sets is given in Appendix C.2.

Chapter 3

Results

The results of the simulation study undertaken to investigate the adequacy of variance models for the VxE effects for small numbers of varieties, specifically the FA structure, are given in this chapter. The accuracy of the estimation of four models for the variance of the VxE effects, also referred to as genetic variance, are compared for 40 data generation models using the average MSEP. The four model types considered for the variance of the VxE effects, or VxE models, were a uniform model, a diagonal model, FA models of order k, and an unstructured model. As discussed in Section 2.3.1, the 40 data generation models consist of the factorial combination of five variance-covariance structures which characterise the data sets from which they were drawn, four numbers of varieties per trial (10, 15, 25, and 50), and two levels of concurrence (partial and perfect). The average MSEPs for the VxE models are compared within a data set; however, the trends, rather than the individual numbers, can be compared across data sets. The order of FA models fitted is heavily influenced by the number of trials contained in a MET data set and for this reason the number of FA models produced for each data set differ.

In Section 3.1, the different order of FA models are compared for each data generation model. In Section 3.2, the FA models are compared to the other VxE models considered in this study and in Section 3.3, the MSEP and log-likelihood ratio test (LLRT) are compared in their selection of superior models.

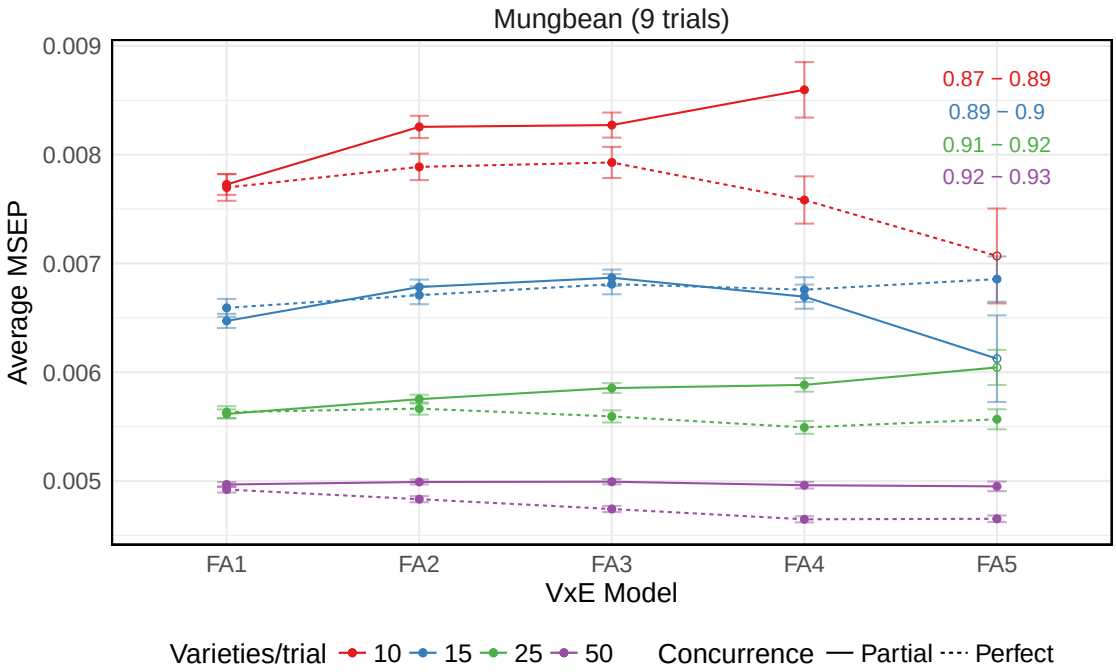
3.1 Comparison of FA models

This section considers the different orders of FA models used to estimate the variance of the VxE effects. The average MSEP for these VxE models from the 500 simulations are shown in Figures 3.1, 3.2, and 3.3 for the different data generation models. The data generation models from each data set are shown in separate plots in these figures. The x-axis shows the dimension of FA model, ordered from the lowest order FA model to the highest order FA model. The y-axis shows the average MSEP, with this scale differing for each data set. Each colour represents a different variety per trial number (10, 15, 25, 50) and the two line types represent

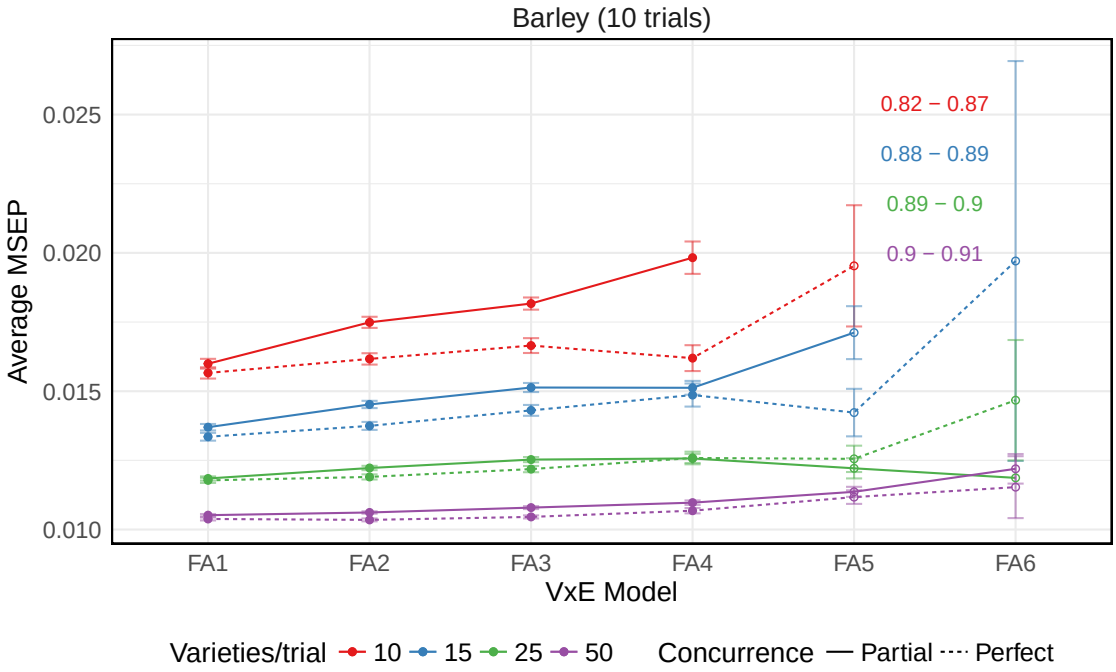
the two concurrence levels (perfect and partial). Standard error bars are shown for each average MSEP, with the number of times the model converged used as the sample size ($n \leq 500$). The full tables of average MSEPs are given in Appendix D.1.

Within a data set, the MSEP gives an indication of whether one type of VxE model for a given scenario is an improvement over another VxE model; however, it is difficult to determine how accurate a model is independently of other models. This is because the MSEP is a relative measure of accuracy and as such its scale is dependent on the data. It is important to recognise this because if the most accurate model for a scenario is not accurate enough to adequately estimate the VxE effects it is of no practical use regardless of its relative superior accuracy. The range of average correlations between the true VxE effects and the estimated VxE effects for the FA models for each variety number are shown on the Figures 3.1, 3.2 and 3.3 in the same colour as their respective variety number lines and points. These correlations indicate whether the accuracy level of a model is practical, regardless of how its accuracy compares to another model. Practical is defined here as reaching a level of accuracy that a breeding program would be satisfied with. This practical level could change substantially depending on the stage of the program or the specific aims of the program, but for the purpose of this study, a correlation of or above 0.85 is considered to be a practical level of correlation. The correlations also allow comparison across data generation models from different data sets. The full tables of correlations are shown in Appendix D.2.

As discussed in Section 2.3.2, a higher order FA model was only fitted to the data provided the previous model offered a significant improvement based on a log-likelihood ratio test and it fulfilled the inequality in Equation 2.17. Due to these conditions some of the higher order models for a given data generation model were only fitted to the data a small number of times. Tables showing the number of simulations in which an FA model was fitted to the data are given in Appendix D.3. In addition to this, a number of FA models were not able to converge in all simulations which further reduces the sample size for some of the FA models. The full tables of convergence information are given in Appendix D.4. Both of these factors need to be taken into consideration when comparing the FA models. Open circles (as opposed to solid circles) are used in Figures 3.1, 3.2, and 3.3 for any average MSEPs that have an effective sample size of 50 simulations or less, taking into account both the number of times the model was fitted and the percentage of time it converged. It is interesting to note that although higher order FA models were fitted (based on improved log-likelihood) in data sets with

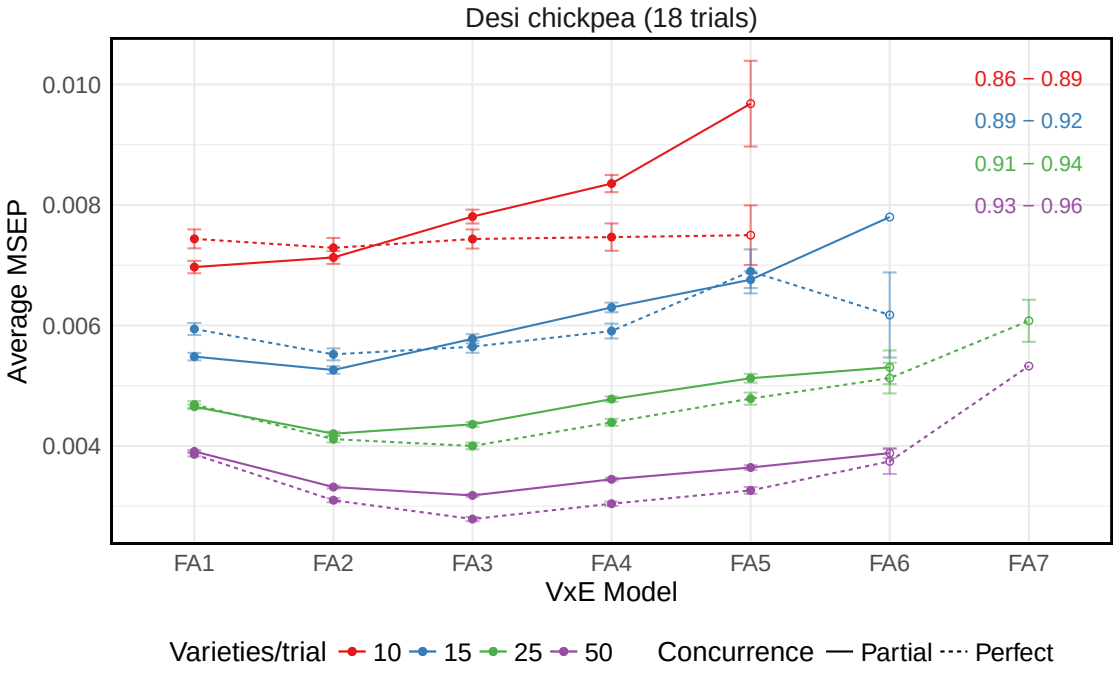


(a) Mungbean data set

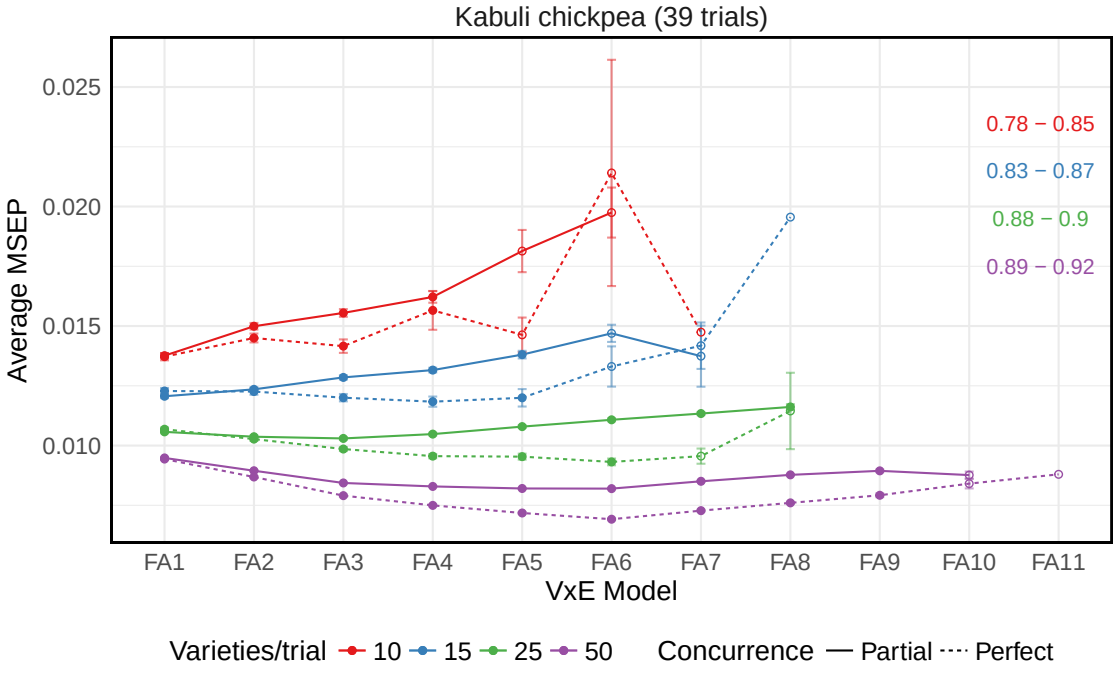


(b) Barley data set

Figure 3.1: Average mean square error of prediction (MSEP) for the FA models from 500 simulations for the data generation models from the (a) mungbean and (b) barley variance-covariance structures. The average correlation range for the models for each variety number are shown on each plot in the same colour as their respective variety number lines and points. Note the different y-axis scales.



(a) Desi chickpea data set



(b) Kabuli chickpea data set

Figure 3.2: Average mean square error of prediction (MSEP) for the FA models from 500 simulations for the data generation models from the (a) Desi chickpea and (b) Kabuli chickpea variance-covariance structures. The average correlation range for the models for each variety number are shown on each plot in the same colour as their respective variety number lines and points. Note the different y-axis scales.

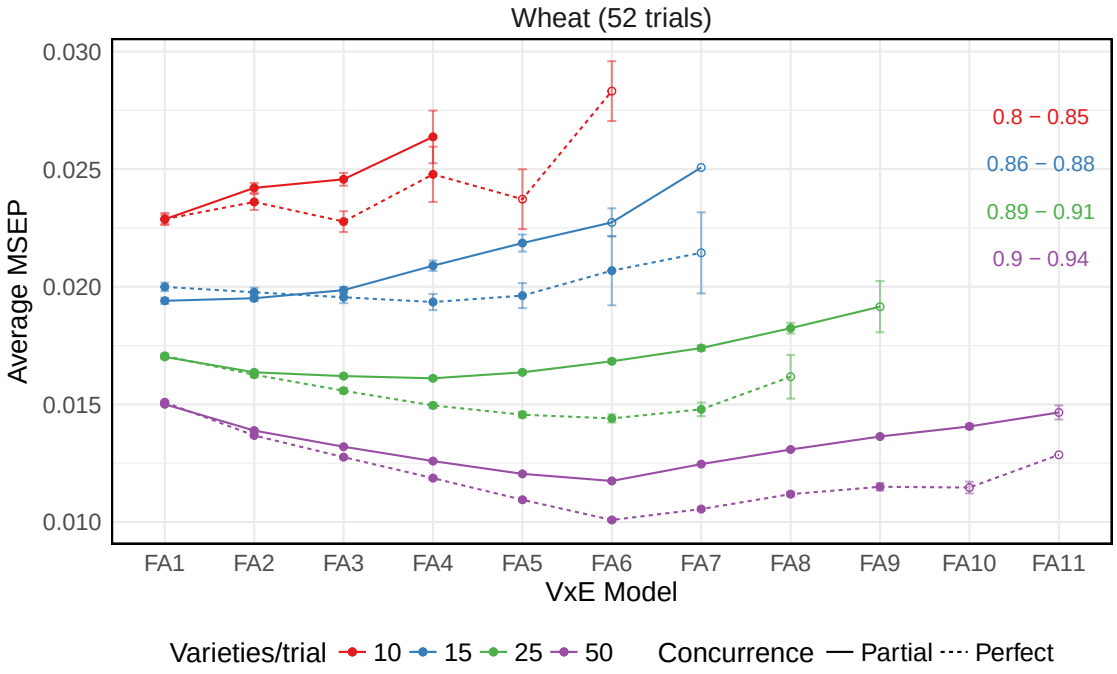


Figure 3.3: Average mean square error of prediction (MSEP) for the FA models from 500 simulations for the data generation model from the wheat variance-covariance structures. The average correlation range for the models for each variety number are shown on the figure in the same colour as their respective variety number lines and points.

larger numbers of trials, they rarely resulted in the most accurate predictions (based on MSEP). For example, when considering 50 varieties in the wheat data set, up to an FA11 model was fitted for both concurrence levels, however, in both cases an FA6 was the most accurate model.

For the mungbean data set (Figure 3.1a) the FA1 model was the best performing model for 10 varieties per trial with partial concurrence (0.0077), while the FA4 model had the highest MSEP (0.0086). For the same variety number with perfect concurrence the FA2 and FA3 models had a larger MSEP than the FA1 model. However, the FA4 and FA5 models had smaller MSEPs than the FA1 model, with the FA5 model being the most accurate (0.0071), though it should be noted this value has a sample size of less than 50. The MSEPs for 15 varieties per trial were very similar for partial and perfect concurrence, aside from the FA5 model. For both perfect and partial concurrence, the average MSEPs for the FA1 model are slightly smaller than for the FA2, FA3, and FA4 models, however, the MSEP for the FA5 model increases slightly for perfect concurrence and decreases substantially for partial concurrence. Consequently, the most accurate model for 15 varieties with perfect concurrence was an FA1 model (0.0066), and for partial concurrence it was an FA5 model (0.0061), however, given $n < 50$, this value may

not be reliable. For 25 varieties per trial with partial concurrence the average MSEP increased slightly as the order of FA model increased, while it decreased for perfect concurrence between an FA1 and FA4 model, however, remained consistent between an FA4 and FA5 model. For 50 varieties per trial with partial concurrence the average MSEP remained consistent for all FA models (0.005), while for perfect concurrence the average MSEP decreased consistently between FA1 and FA4 models and remained consistent between FA4 and FA5 models ($\sim$ 0.0047). For both 25 and 50 varieties per trial, perfect and partial concurrence had similar average MSEPs for FA1 and FA2 models; however, for both variety numbers, the difference between the average MSEPs for perfect and partial concurrence increased for the higher order FA models.

The average MSEPs for the data generation models from the barley data set (Figure 3.1b) are fairly consistent across FA models for a given variety number and concurrence level, with the highest order FA model fitted to a data generation model occasionally being the exception to this, although it should be noted these higher order models tend to have a small sample size. The average MSEPs for the data generation models from the Desi chickpea data set (Figure 3.2a) mostly followed the same pattern for each variety number with the average MSEP decreasing between FA1 and FA2 or FA3 and then increasing as the order of FA model rose. For the data generation models from the Kabuli chickpea (Figure 3.2b) and wheat data sets (Figure 3.3) with 10 and 15 varieties per trial, the average MSEP tended to be smallest for the FA1 model. For 25 and 50 varieties per trial, higher order FA models, such as an FA6 model, tended to have the smallest average MSEP, aside from 25 varieties with partial concurrence, in which an FA2 or an FA3 model had the smallest average MSEP. Despite an FA11 model being fitted for some data generation models from the Kabuli chickpea and wheat data sets, an FA6 had the smallest average MSEP in these cases.

The consistency of the average MSEP across FA models changes depending on the data generation model. The data generation models for perfect and partial concurrence resulted in a similar average MSEP for their respective data sets, numbers of varieties per trial and FA model, however, perfect concurrence tended to result in a lower average MSEP. The average MSEP decreased as the number of varieties per trial increased, regardless of the level of concurrence or the data set considered. For most of the data generation models, the highest order FA model that had been fitted to the data tended not to result in the smallest average MSEP.

3.2 Comparison of FA models with other models

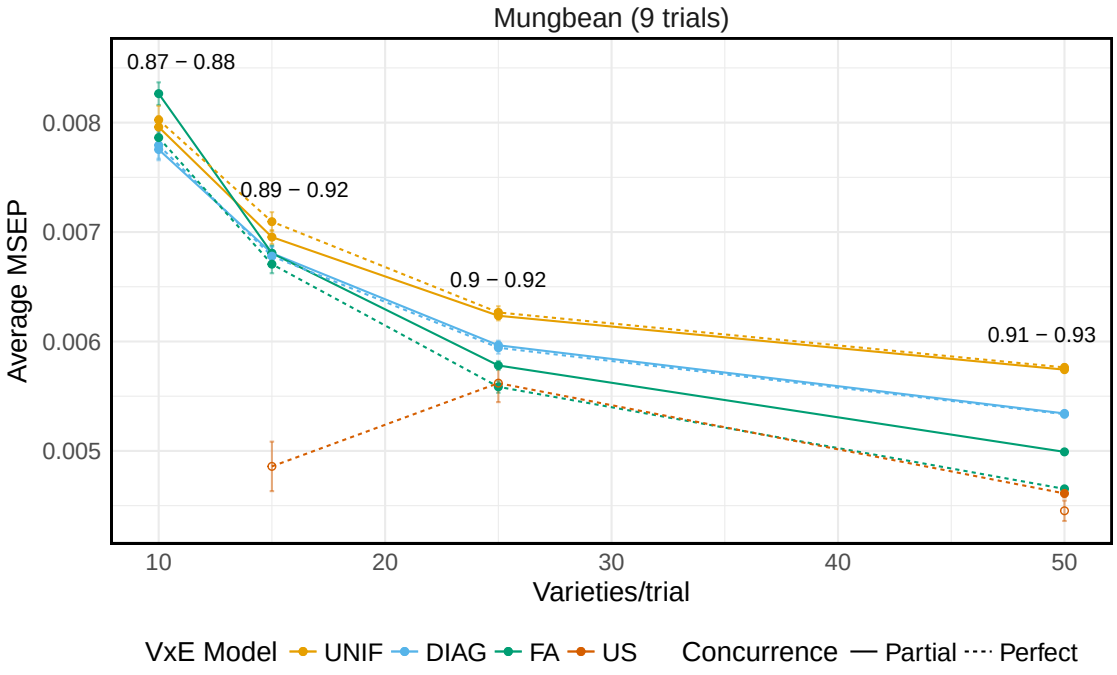
This section compares the FA model to the other models used to estimate the variance of the VxE effects. The average MSEP from the 500 simulations are shown in Figures 3.4, 3.5, and 3.6 for each data generation model. The average MSEP is shown for the uniform model, diagonal model, the FA model which offered the best fit to the data in each simulation (using a log-likelihood ratio test), and the unstructured model. In these figures the average MSEPs for each data set are shown in separate plots. The x-axis shows the number of varieties per trial and the y-axis shows the average MSEP, with the scale of this axis differing for each data set. Each colour represents a different VxE model, while the line types represent the two concurrence levels. Standard error bars are shown for each average MSEP. The average correlation range for the models for each variety number are shown on the figures above their respective variety number.

In interpreting these results it should be noted that the unstructured model failed to converge in all simulations for all of the data generation models from the Desi chickpea, Kabuli chickpea and wheat data sets. The percentage of simulations in which the unstructured model converged for the remaining data generation models are shown in Table 3.1. Open circles (as opposed to solid circles) are used in Figures 3.4, 3.5, and 3.6 for any average MSEPs that had a sample size of 50 simulations or less.

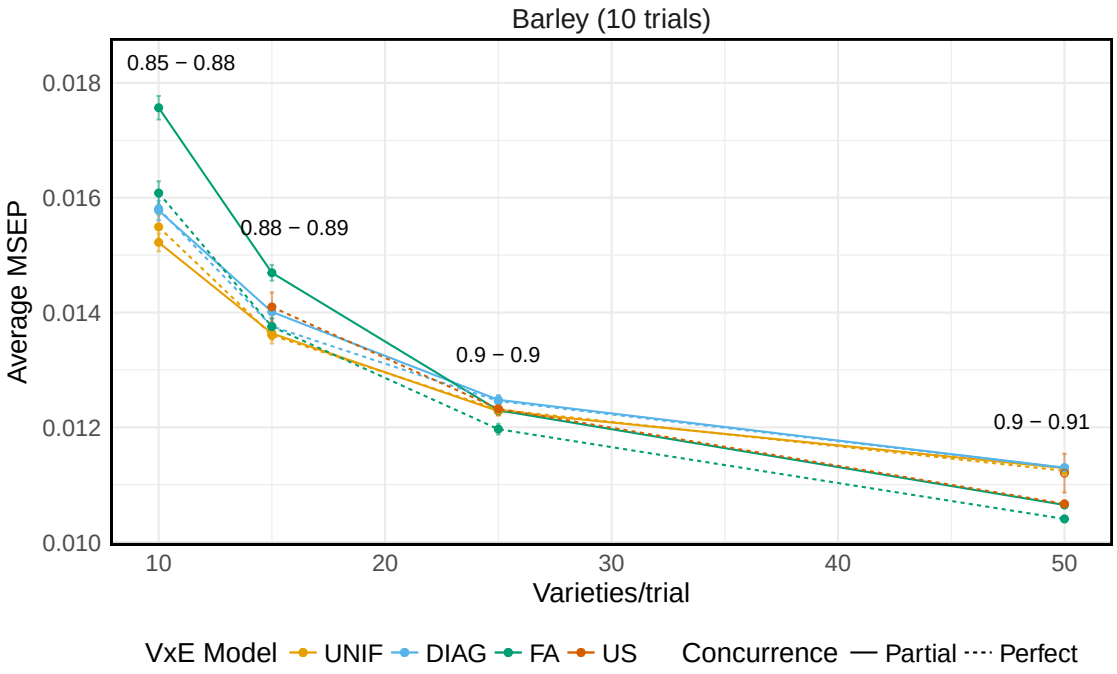
Table 3.1: Percentage of simulations in which the unstructured model converged in 500 simulations for mungbean and barley data sets.

Variance-covariance structure	Concurrence	Varieties per trial			
		10	15	25	50
Mungbean (9 trials)	Partial	0.0	0.0	0.0	0.6
	Perfect	0.0	0.6	9.8	52.6
Barley (10 trials)	Partial	0.0	0.0	0.0	0.8
	Perfect	0.0	40.4	77.6	89.6

For the mungbean data set (Figure 3.4a), the average MSEP for all of the VxE models, aside from the unstructured model, were quite similar for 10 and 15 varieties per trial for both perfect and partial concurrence. The average MSEP for the FA model for partial concurrence had a larger MSEP (0.0083) than the other models for 10 varieties per trial for both concurrence levels, with all remaining MSEPs for 10 varieties having very similar MSEPs (0.0078 - 0.008), however, this difference in MSEP has limited practical implications as the correlations only change from 0.87 to 0.88. For 15 varieties per trial the FA model with partial

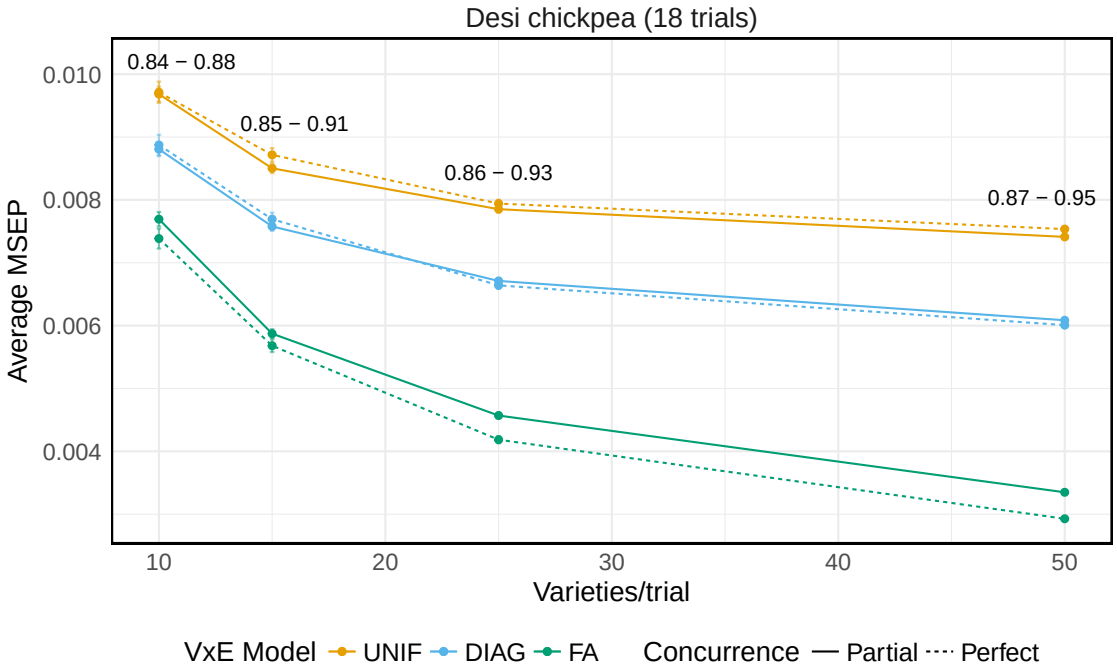


(a) Mungbean data set

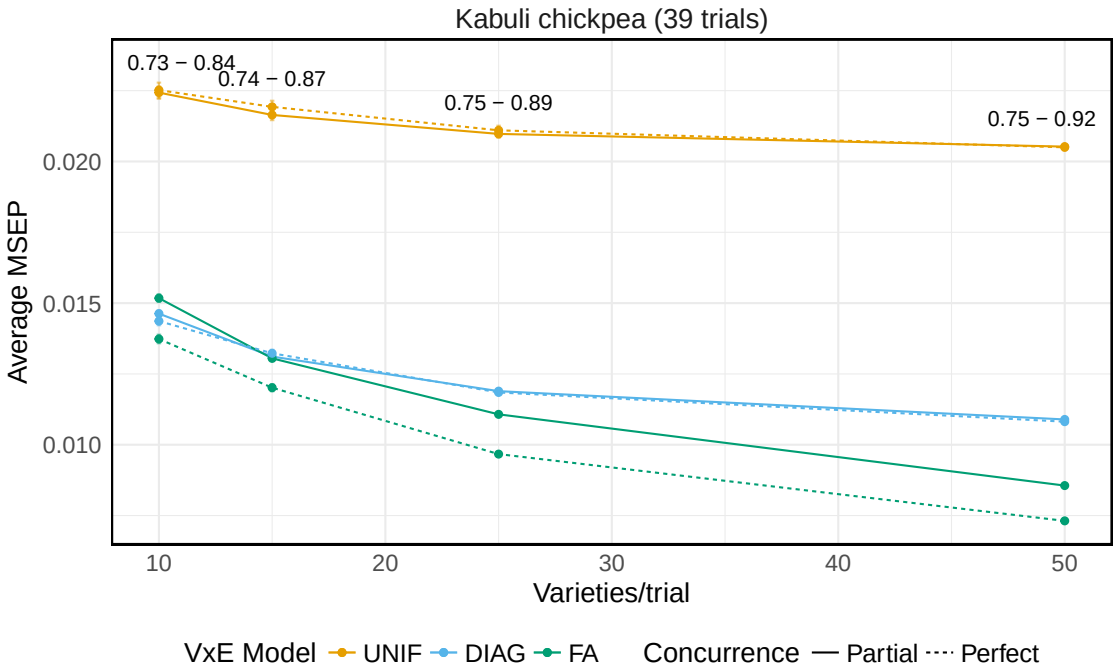


(b) Barley data set

Figure 3.4: Average mean square error of prediction (MSEP) from 500 simulations for the data generation models from the (a) mungbean and (b) barley variance-covariance structures for the uniform model (UNIF), diagonal model (DIAG), factor analytic model (FA), and unstructured model (US). The average correlation range for the models for each variety number are shown on each plot in black. Note the different y-axis scales.



(a) Desi chickpea data set



(b) Kabuli chickpea data set

Figure 3.5: Average mean square error of prediction (MSEP) from 500 simulations for the data generation models from the (a) Desi chickpea and (b) Kabuli chickpea variance-covariance structures for the uniform model (UNIF), diagonal model (DIAG), factor analytic model (FA), and unstructured model (US). The average correlation range for the models for each variety number are shown on each plot in black. Note the different y-axis scales.

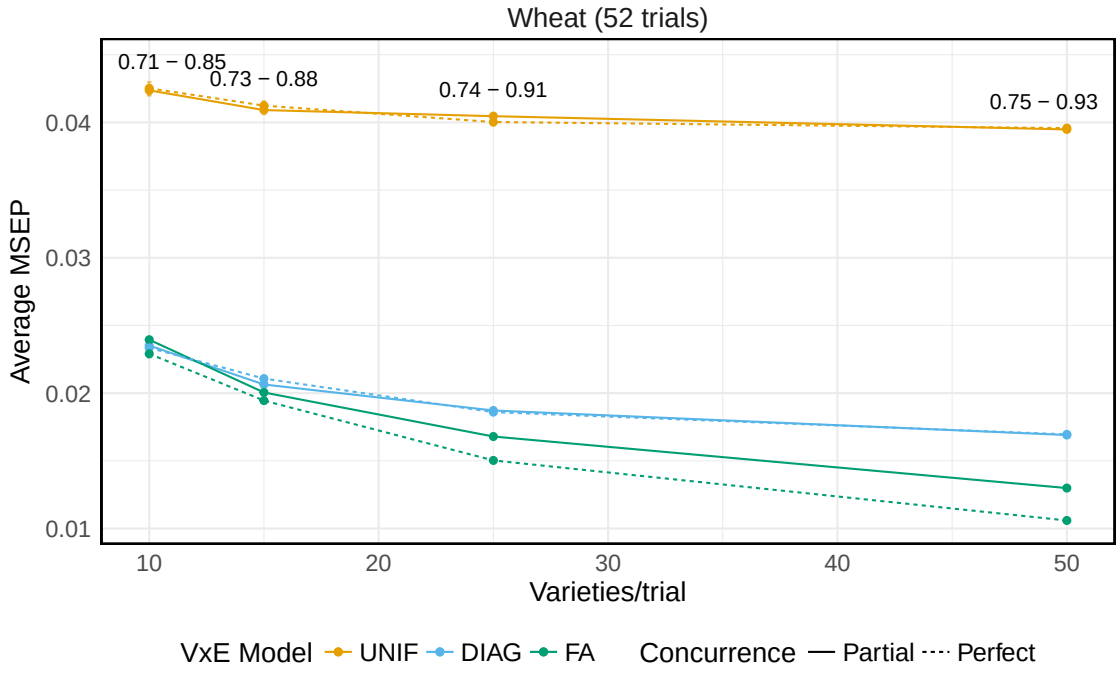


Figure 3.6: Average mean square error of prediction (MSEP) from 500 simulations for the data generation model from the wheat variance-covariance structures for the uniform model (UNIF), diagonal model (DIAG), factor analytic model (FA), and unstructured model (US). The average correlation range for the models for each variety number are shown on the figure in black.

concurrence improved with an average MSEP identical to that for the diagonal model for each concurrence level (0.0068). The average MSEP for the unstructured model for 15 varieties per trial with perfect concurrence was much smaller than the other models (0.0049); however, given that this estimate is based on only three simulations, limited importance will be placed on this result. The differences between the models become more pronounced for 25 and 50 varieties per trial. For 25 and 50 varieties per trial for the uniform and diagonal models there was no difference in the average MSEP between perfect and partial concurrence. The average MSEP for the FA model with perfect concurrence was consistently smaller than that for partial concurrence for both 25 and 50 varieties per trial. The average MSEPs for the unstructured model with perfect concurrence for 25 and 50 varieties were very close to those for the FA model with perfect concurrence. For both 25 and 50 varieties per trial the FA model had a smaller average MSEP than both the uniform and diagonal model, and the diagonal model had a smaller average MSEP than the uniform model. However, it is important to note that for both 25 and 50 varieties the correlations between the worst models for partial or perfect concurrence and the best models only differed by 2%.

For the barley data set (Figure 3.4b), the FA model for 10 varieties per trial, and to a lesser extent for 15 varieties per trial, with partial concurrence performed considerably poorer than the other analysis models for 10 and 15 varieties per trial, however, the models performed similarly to each other for both 25 and 50 varieties per trial for both perfect and partial concurrence. For the Desi chickpea (Figure 3.5a), Kabuli chickpea (Figure 3.5b), and wheat (Figure 3.6) data sets, the differences between the models were much more pronounced. The uniform model consistently had a higher average MSEP across variety numbers, especially for the Kabuli chickpea and wheat data sets. The diagonal model and FA model performed similarly to each other for the Kabuli chickpea and wheat data generation models for 10 and 15 varieties per trial, while for the Desi chickpea data set there were clear differences between the VxE models for 10 and 15 varieties. However, for all three of these data sets, the FA model for 25 and 50 varieties per trial was noticeably better than the other VxE models for both partial and perfect concurrence.

For the uniform and diagonal models in all data sets, partial and perfect concurrence resulted in similar average MSEP values for their respective variety numbers (Figures 3.4, 3.5, 3.6). The average MSEP for partial and perfect concurrence differed more for the FA models, with perfect concurrence resulting in a smaller average MSEP for the same number of varieties per trial. The average MSEP for each data set decreased as the number of varieties per trial increased as expected, decreasing most rapidly between 10 and 15 varieties, then 15 and 25, followed by 25 and 50 varieties.

3.3 Model selection

Given that the parameters used to simulate the data in this study are known, the MSEP can be used to determine the model which best estimates the true effects. In practice the true effects are unknown and a different method of choosing the best model which does not rely on this knowledge must be used, such as the log-likelihood ratio test (LLRT) or AIC. The percentage of simulations in which each type of model had the smallest MSEP and the percentage of simulations in which each model type was determined to be the most parsimonious model according to a LLRT are summarised in Tables 3.2, 3.3, 3.4, 3.5, and 3.6. The percentage of simulations in which models were selected to be the most parsimonious using the AIC were very similar to those selected using the LLRT and are consequently not presented. These tables demonstrate the differences between the model which would be selected as the best model using these two methods.

These tables show that in selecting the best model, the MSEP tended to lead to a choice of a simpler model than the LLRT. For the mungbean data set (Table 3.2) for 10 varieties with partial concurrence the uniform model was selected most frequently (26% of simulations), while the LLRT selected the FA2 model most often (57.2% of simulations) and only selected the uniform model in 0.2% of simulations. Similarly for 10 varieties with perfect concurrence the MSEP selected the FA1 model and uniform model most frequently (27% and 26.6% of simulations respectively), while the LLRT selected the FA2 model most frequently (49.4% of simulations). For 15 varieties per trial the FA2 model was selected approximately a third of the time for both concurrence levels using the MSEP, while the LLRT predominantly selected the FA2 and FA3 models (31.6 - 41.6% of simulations). For 25 varieties per trial the MSEP selected the FA1 model most often (43.6% of simulations) for partial concurrence, while for perfect concurrence the FA1 and FA3 model were both selected in 22% of simulations. The LLRT for 25 varieties selected the FA1, FA2, and FA3 models most frequently for partial concurrence and the FA2, FA3, and FA4 models most often for perfect concurrence. Using the MSEP the FA1 model was selected as the best model approximately a third of the time for 50 varieties with partial concurrence, while the LLRT selected the FA3 model approximately a third of the time. For 50 varieties with perfect concurrence the MSEP and LLRT both selected the FA4 model most frequently, however, the percentage of simulations in which this model was selected as the best varied considerably, being selected by the MSEP in 39.8% of simulations, while the LLRT selected it in 62.8% of simulations.

For the barley data set the uniform and diagonal models had the smallest MSEP in a large proportion of simulations for 10 and 15 varieties per trial regardless of concurrence, while these models were barely ever selected as the best using the LLRT. For 25 varieties with perfect and partial concurrence and 50 varieties with partial concurrence the FA1 model was selected most frequently using the MSEP, while the LLRT selected the FA2 model most frequently. Both methods selected the FA2 model for 50 varieties with perfect concurrence. For all of the data generation models from the Desi chickpea data set, aside from 10 varieties with perfect concurrence, the LLRT most frequently selected the FA3 model as the best model. The model selected by the MSEP varied much more according to variety number and concurrence, with an FA1 model selected most frequently for 10 varieties with partial concurrence, an FA2 model selected most often for 10 varieties with perfect concurrence, 15 varieties with both perfect and partial concurrence, and 25 varieties with partial concurrence, and an FA3 model selected most frequently for 25 varieties with perfect concurrence, and 50 varieties

for both concurrence levels.

For the Kabuli data set, the MSEP and LLRT selected the same model most frequently for 15 and 25 varieties per trial with perfect concurrence. For 25 varieties with partial concurrence the MSEP selected the FA3 model most often while the LLRT selected the FA6 model most often, however, for all remaining data generation models from the Kabuli data set the LLRT selected the FA model of and equal to orhigher than that selected using the MSEP. The MSEP and LLRT selected similar models for the data generation models from the wheat data set with the LLRT most often selecting an FA model one order higher than that selected most often by the MSEP. However, the diagonal model was selected as the best model using the MSEP in 34% of simulations for 10 varieties with partial concurrence and in 26.6% of simulations with perfect concurrence, while the LLRT selected this model in only 1.41% and 2.4% of simulations respectively.

Table 3.2: Percent of simulations in which model was best according to the MSEP and log-likelihood ratio test (LLRT) for 500 simulations for the mungbean data set.

Method	Varieties	Concurrence	UNIF	DIAG	FA1	FA2	FA3	FA4	FA5	US
MSEP	10	Partial	26.0	23.8	20.4	13.4	15.6	0.8		
	10	Perfect	26.6	18.8	27.0	13.4	12.2	2.0		
	15	Partial	14.4	18.4	34.6	15.8	11.4	5.4		
	15	Perfect	14.2	13.4	31.2	18.4	11.4	9.8	1.6	
	25	Partial	5.8	16.4	43.6	19.6	8.2	5.4	1.0	
	25	Perfect	4.2	4.8	22.0	16.6	22.6	18.0	9.0	2.8
	50	Partial	0.6	5.6	34.4	22.8	18.8	10.8	7.0	
	50	Perfect		0.2	4.6	7.4	14.6	39.8	20.4	13.0
LLRT	10	Partial	0.2	0.4	21.2	57.2	21.0			
	10	Perfect	1.2	0.2	27.2	49.4	18.6	3.2	0.2	
	15	Partial		0.2	21.6	41.6	35.8	0.8		
	15	Perfect			19.0	31.6	33.6	14.8	1.0	
	25	Partial			26.8	31.6	33.6	8.0		
	25	Perfect			9.8	21.0	33.0	29.0	6.4	0.8
	50	Partial			11.6	27.4	34.6	22.8	3.6	
	50	Perfect				2.0	13.4	62.8	15.8	6.0

Table 3.3: Percent of simulations in which model was best according to the MSEP and a log-likelihood ratio test (LLRT) for 500 simulations for the barley data set.

Method	Varieties	Concurrence	UNIF	DIAG	FA1	FA2	FA3	FA4	FA5	US
MSEP	10	Partial	40.3	23.8	21.2	6.8	7.0	0.8		
	10	Perfect	43.4	18.2	21.8	12.6	3.8	0.2		
	15	Partial	33.0	19.6	27.8	13.0	4.0	2.6		
	15	Perfect	29.8	13.4	35.2	14.6	3.6	0.4		3.0
	25	Partial	19.2	14.6	43.4	16.6	4.8	1.0	0.4	
	25	Perfect	15.0	4.8	41.2	28.6	5.8	1.0	0.2	3.4
	50	Partial	4.0	2.2	56.4	27.2	8.4	1.0	0.6	0.2
	50	Perfect	0.6	0.4	37.0	40.6	15.6	2.4	0.2	3.2
LLRT	10	Partial	0.4	1.8	25.3	50.9	21.6			
	10	Perfect	0.6	1.8	37.4	48.2	11.2	0.8		
	15	Partial		0.6	34.4	32.2	30.2	2.6		
	15	Perfect	0.4	0.2	39.0	40.0	15.2	3.4	0.6	1.2
	25	Partial			31.8	38.6	21.8	7.6	0.2	
	25	Perfect			26.4	49.2	18.0	4.2	0.4	1.8
	50	Partial			24.2	50.0	20.8	4.6	0.4	
	50	Perfect			7.0	51.2	30.4	6.6	1.2	3.6

Table 3.4: Percent of simulations in which model was best according to a log-likelihood ratio test (LLRT) for 500 simulations for the Desi chickpea data set.

Method	Varieties	Concurrence	UNIF	DIAG	FA1	FA2	FA3	FA4	FA5	FA6
MSEP	10	Partial	4.0	5.8	47.4	30.2	9.2	3.4		
	10	Perfect	3.8	2.4	31.4	39.0	18.6	4.8		
	15	Partial	1.2	0.8	31.4	53.0	11.8	1.8		
	15	Perfect	0.4	0.4	14.2	49.8	29.2	6.0		
	25	Partial	0.2	0.2	10.0	56.0	31.8	1.8		
	25	Perfect			2.6	26.0	69.2	2.2		
	50	Partial				16.2	83.0	0.8		
	50	Perfect				3.0	96.8	0.2		
LLRT	10	Partial			0.6	40.2	57.8	1.4		
	10	Perfect			2.8	57.6	37.0	2.6		
	15	Partial				15.6	58.2	26.0	0.2	
	15	Perfect			0.2	33.4	55.4	10.0	1.0	
	25	Partial				4.6	50.2	40.2	5.0	
	25	Perfect				3.8	61.4	29.6	4.4	0.8
	50	Partial					49.5	42.2	8.2	
	50	Perfect					60.2	33.8	5.8	0.2

Table 3.5: Percent of simulations in which model was best according to a log-likelihood ratio test (LLRT) for 500 simulations for the Kabuli chickpea data set.

Method	Varieties	Concurrence	UNIF	DIAG	FA1	FA2	FA3	FA4	FA5	FA6	FA7	FA8	FA9	FA10
MSEP	10	Partial		22.8	57.2	13.3	5.9	0.8						
	10	Perfect	0.2	12.6	51.1	25.7	8.0	2.4						
	15	Partial		9.8	51.0	24.7	9.8	3.6	0.8	0.2				
	15	Perfect	0.2	1.6	31.6	33.4	22.8	9.2	1.2					
	25	Partial		0.4	14.4	28.0	33.2	15.0	5.6	2.4	0.8	0.2		
	25	Perfect			6.6	19.2	31.8	21.8	13.0	7.0	0.6			
	50	Partial				0.2	9.0	18.6	28.0	42.6	1.4	0.2		
	50	Perfect				0.6	0.8	0.8	9.4	88.2	0.2			
LLRT	10	Partial			7.3	64.4	25.9	2.0	0.4					
	10	Perfect		0.2	38.1	38.9	15.6	5.2	1.8	0.2				
	15	Partial			1.2	23.7	43.2	25.3	6.2	0.4				
	15	Perfect			18.4	43.2	22.2	10.8	4.0	1.2	0.2			
	25	Partial				1.2	1.0	4.0	16.2	52.0	25.6			
	25	Perfect			3.8	24.8	27.0	21.0	16.2	6.6	0.6			
	50	Partial							0.2	17.2	50.4	30.2	2.0	
	50	Perfect				1.2	0.2		1.2	28.0	42.8	22.4	4.0	0.2

Table 3.6: Percent of simulations in which model was best according to a log-likelihood ratio test (LLRT) for 500 simulations for the wheat data set.

Method	Varieties	Concurrence	DIAG	FA1	FA2	FA3	FA4	FA5	FA6	FA7	FA8	FA9	FA10
MSEP	10	Partial	34.0	47.9	14.3	3.4	0.4						
	10	Perfect	26.6	43.4	21.6	6.2	1.8	0.4					
	15	Partial	12.5	39.0	28.1	16.0	3.6	0.8					
	15	Perfect	6.0	29.0	30.0	21.8	9.8	3.4					
	25	Partial	0.2	6.8	24.2	23.4	25.6	15.2	3.8	0.8			
	25	Perfect	0.6	5.0	12.8	17.6	26.0	24.2	11.6	2.2			
	50	Partial		0.2	0.8	0.8	3.2	21.6	73.2	0.2			
	50	Perfect	0.4	0.2	0.6	0.8	1.8	3.8	86.6	5.8			
LLRT	10	Partial	1.4	20.3	65.8	12.3	0.2						
	10	Perfect	2.4	37.8	34.8	15.4	7.0	2.2	0.4				
	15	Partial		2.8	43.6	36.6	13.9	2.8	0.2				
	15	Perfect	1.4	20.0	32.2	18.6	15.4	9.8	2.6				
	25	Partial			2.6	5.8	13.6	29.4	35.4	11.8	1.4		
	25	Perfect	0.6	1.6	16.6	15.2	22.0	21.4	20.4	2.2			
	50	Partial			1.0	0.8	1.4	0.2	3.6	29.2	45.6	17.2	1.0
	50	Perfect	0.4	0.2	0.6	0.8	1.8	2.4	41.8	35.6	14.2	2.0	0.2

Chapter 4

Discussion

In the early stages of a crop breeding program large numbers of varieties are grown at each trial, however, in the later stages of the program these numbers are dramatically reduced. Accuracy is important at all stages of a program, but in these final stages, where final decisions are made regarding the varieties to be recommended for commercial release, it is of paramount importance. There has been limited research regarding the accuracy of FA models for late stage breeding trials, despite FA models being commonly used for the analysis of such data. The aims of this project were to determine whether the adequacy of an FA variance structure changes as the number of crop varieties within a trial decreases; and, when dealing with these small numbers, to investigate the impact the underlying VxE variance structure and level of varietal concurrence between environments have on the accuracy of the FA model.

The main focus of this study was to investigate whether the accuracy of commonly used models is influenced by decreasing the number of varieties per trial in a MET data set. There is limited research regarding how accurate the models investigated are in the analysis of MET data for small numbers of varieties. Welham *et al.* (2010) undertook a simulation study dealing with small numbers of varieties per trial, however, variety number was not a parameter being altered in this study and consequently conclusions cannot be made regarding how changing the variety number impacts on the accuracy of the models. Furthermore the focus of the study undertaken by Welham *et al.* (2010) was in comparing one- and two-stage analyses of MET data and only investigated an FA2 model, a uniform model, and a modified uniform model that fitted fixed variety effects and random environment effects.

The results of this study showed that changing the number of varieties per trial impacts on the accuracy of models for the variance of the VxE effects. In Figures 3.1, 3.2, and 3.3, where each colour represents a variety number, and in Figures 3.4, 3.5, and 3.6, where the x-axis represents variety number, it can be seen

that as the number of varieties decrease, the accuracy also decreases (increased MSEP). This trend is true for all of the VxE models investigated, and is consistent regardless of data set or varietal concurrence. The rate at which the accuracy of the models improves is greater when dealing with small numbers of varieties per trial, with the change in accuracy between 10 and 15 varieties being greater than or equivalent to the change in accuracy between 25 and 50 varieties.

There appears to be some distinction between 15 and 25 varieties per trial, with the patterns in the results very similar for 10 and 15 varieties and very similar for 25 and 50 varieties. This suggests that although all variety numbers considered in this study are "small" by crop breeding program standards, 10 and 15 varieties fall into a different "small" category compared to 25 and 50 varieties. For the scenarios which had 10 or 15 varieties per trial, in two of the data sets the uniform model performed better than or equivalent to the other VxE models, however, in the remaining three data sets the uniform model was substantially less accurate than the other VxE models. The poor accuracy of the uniform model in most cases is not surprising with many authors (including Patterson & Silvey 1980; Patterson & Nabugoomu 1992; Cullis *et al*. 1998; Smith *et al*. 2005) noting that the assumptions inherent in this model are questionable. However, these results also indicate that in some cases (perhaps those with small numbers of trials) a uniform model can accurately predict the VxE effects, even with the limiting constraints on the underlying variance structure fitted by this model. In contrast, for these same scenarios the diagonal model performs very similarly to the FA model in all but one of the data sets. This suggests that for many data sets a diagonal model, which is equivalent to analysing the trials separately, is just as accurate as an FA model for very small numbers (10 - 15) of varieties per trial, despite being a much simpler model and ignoring all covariance.

There is no data on the unstructured model for 10 varieties because it did not converge for any scenarios and there are only two average MSEPs for the unstructured model for 15 varieties and one of these had a sample size of three. Given these limited results no conclusions will be made regarding the accuracy of the unstructured model for very small numbers of varieties. However, its failure to converge in this study indicates a potentially severely limiting factor in the use of this model for such scenarios. The computational difficulties involved in fitting the unstructured model have been well documented particularly for large data sets (Smith *et al*. 2001, 2005; Kelly *et al*. 2007; Smith *et al*. 2015), however, in this study it was found that in many cases even for relatively small data sets (9 or 10 trials with 10 varieties) the unstructured model failed to converge even once

in 500 simulations.

In all scenarios with 25 or 50 varieties per trial either the FA or unstructured model resulted in the most accurate predictions of the VxE effects, and in all of these cases the average MSEPs for these models were very similar. It is difficult to compare the unstructured model due to its poor convergence, however, for the scenarios with an adequate sample size ($n > 50$) the unstructured model always resulted in equally accurate or less accurate estimation of the VxE effects than the FA model. When considering 25 and 50 varieties, in four of the five data sets (mungbean, Desi chickpea, Kabuli chickpea and wheat data sets) there is a noticeable difference in accuracy between the uniform model and the other VxE models, and in three of these data sets the difference is very pronounced (Desi chickpea, Kabuli chickpea and wheat data sets). There is also a distinction between the accuracy of the diagonal model and the FA model in these four data sets, however, the difference is less than that between the FA and uniform model. These results suggest that for moderately small numbers of varieties (25 - 50), the FA model results in the most accurate estimate of the VxE effects. The results for the barley data set show that there are situations in which the uniform and diagonal models can result in a similar accuracy to the FA model. The findings for 25 and 50 varieties per trial are mostly consistent with the findings of Kelly *et al.* (2007) for large numbers of varieties.

The study undertaken by Kelly *et al.* (2007) investigated the accuracy of FA models for trials with 80, 200, and 500 varieties with data sets containing eight to 10 trials. The study showed that the uniform and diagonal model performed poorly compared to the FA and unstructured models for all variety numbers investigated. One set of parameters used in the study from Kelly *et al.* (2007) were also used in this study and the results can be compared for small and large variety numbers. Considering just the cases from the barley data set with perfect concurrence (as in Kelly *et al.* (2007)), for 10 varieties the uniform model was the most accurate model, while for 15 varieties the FA 1 model was the most accurate followed by the uniform model. However, for 25 and 50 varieties the FA model was the most accurate model, in agreement with the previous study by Kelly *et al.* (2007) with much larger variety numbers. However, it should be noted that in the Kelly *et al.* (2007) study the FA models outperformed both the diagonal and uniform models to a much greater extent.

Although the accuracy of all the models declined as the number of varieties decreased, it is important to recognise whether the level of accuracy reached was

practical. If even the most accurate model for a given variety number cannot adequately estimate the VxE effects, this implies that these trials do not contain enough varieties to accurately estimate genetic variance. However, if the least accurate models still reach a high enough level of accuracy this implies that these simpler models are sufficient to estimate genetic variance for a given variety number. For 10 varieties none of the models for any data generation model have a correlation between the true and estimated VxE effects greater than 0.89, and for 15 varieties the models rarely have a correlation above 0.89, aside from those in the mungbean and Desi chickpea data sets. It is interesting to note that for the mungbean and barley data sets, which have a similar (small) number of trials, all of the models for 25 and 50 varieties have correlations of 0.9 or greater. In contrast for the Kabuli chickpea and wheat data sets, and to a lesser extent the Desi chickpea data set, the uniform model has a poor level of accuracy regardless of the number of varieties, while the FA model frequently has a correlation of 0.9 or greater for all scenarios with 25 or 50 varieties per trial. This suggests that regardless of the VxE model used to analyse the data, the accuracy of the models are curtailed by the number of varieties. When MET data sets have very small numbers of varieties it is much more difficult to reach a practical level of accuracy, although, for all scenarios involving 15, 25 and 50 varieties there were models that were able to do this, suggesting the 10 varieties per trial is not enough in most cases if a correlation of greater than 0.85 is desired.

In addition to investigating the accuracy of FA models for small numbers of varieties per trial, the impact the underlying VxE variance structure has on the accuracy in such cases was also investigated. In the studies undertaken by Welham *et al*. (2010) and Kelly *et al*. (2007) only two data sets for the VxE variance-covariance patterns were used. Kelly *et al*. (2007) also used two versions of each variance structure, one derived from an FA model and one derived from an unstructured model. The differences between these two versions were not large enough to represent different data sets, but rather had mathematical properties appropriate to addressing the hypothesis of the study. Consequently their ability to investigate how the trends they noticed changed across scenarios was limited.

The five data sets from which the variance-covariance structures were drawn for this study were chosen as they reflect typical MET data sets with varying variance-covariance patterns and different numbers of trials. Given that the underlying patterns in the variance of the VxE effects are intrinsically linked to the number of trials in this study, it is impossible to determine which aspect of the data sets, the underlying VxE pattern or the number of trials, is driving the

results. Two of the data sets (mungbean and barley) have quite a low number of trials (≤ 10), while two of the remaining data sets (Kabuli chickpea and wheat) have a substantial number of trials (≥ 36). The underlying pattern in the variance of the VxE effects for the mungbean and barley data sets are quite consistent. Most of the trials have moderate positive genetic correlations with each other. In the barley data set there is one trial that has poor genetic correlations with the other trials, and in the mungbean data set there are two trials that have poor genetic correlations with the other trials. The VxE pattern for the Kabuli chickpea data set is also quite consistent with most trials having moderate to strong positive genetic correlations with each other. In contrast the patterns in the variance-covariance structure for the Desi chickpea and wheat data sets are much less consistent, with some trials having very strong positive genetic correlations and other pairs of trials having very strong negative genetic correlations. In both of these data sets the VxE patterns suggest there are distinct groups of trials in which varieties perform similarly, with the rankings of varieties in trials from different groups changing substantially. In this study, the data sets that have a greater range in genetic correlations also have a larger number of trials. Using this range of underlying patterns for the variance of the VxE effects allows the impact of small numbers of varieties to be investigated for a range of scenarios.

Comparing the individual FA models between data sets is difficult as the number of trials a data set contains heavily influences the order of FA model fitted and the order found to be most accurate. Despite this, the data sets have similar ranges in the correlations for the FA models considered for a given variety number. Higher order FA models are fitted for the data sets with more trials and when variety numbers are larger. The difference in the order of FA model that most frequently has the more accurate predictions for 10 and 50 varieties is greater when the MET data set has more trials. For example, an FA1 model most often results in the best predictions of the VxE effects for 10 varieties and frequently also for 15 varieties regardless of the data set, however, for 50 varieties per trial the model that most frequently has the smallest MSEP ranges from an FA1 to an FA6 model. The order of FA models fitted in this study are much higher than the case studies given in the literature. Smith *et al.* (2001) and Kelly *et al.* (2007) considered up to an FA3 model, while Welham *et al.* (2010) considered an FA2 model, and Smith *et al.* (2015) considered up to an FA5 model, though it should be noted this was a two-stage analysis with 129 trials.

The trend of accuracy decreasing as the number of varieties per trial decrease is consistent across the data sets representing different VxE patterns, however,

the ranking of the models differs between data sets, especially for very small numbers of varieties (10 - 15). The mungbean and barley data sets have similar characteristics. They have 9 and 10 trials respectively and in addition the range of genetic correlations for the mungbean data set is (-0.14, 0.78) and for the barley data set is (-0.09, 0.76) and the mean genetic correlation is 0.34 and 0.36 for the munbean and barley data sets respectively. Despite these similarities, the patterns in the results of the two data sets are very different. This suggests that it is the individual components in the variability in the VxE pattern rather than the broad characteristics that are driving the differences in the results, indicating that the underlying VxE variance structure can substantially impact how VxE models perform in comparison to each other.

In contrast to this, the Kabuli chickpea and wheat data sets have very similar patterns in their respective results, despite having quite different VxE patterns. The wheat data set is somewhat larger than the Kabuli chickpea data set, containing 13 more trials, although it should be noted that both data sets are large in terms of trial number. Despite this difference in trial numbers, the diagonal model more often results in more accurate predictions for the wheat data set than in the Kabuli chickpea data set, however, while the proportion of simulations in which a model is most accurate differs somewhat between the two data sets, the average correlations between the true VxE effects and the estimated VxE effects are very similar.

The accuracy of the uniform model seems to be most heavily influenced by the underlying characteristics of the data set, whether that be the VxE pattern or the number of trials. For two data sets it offered reasonably accurate predictions, however, in the other three it offered a poor fit across variety numbers. This is not overly surprising given that this model constrains the genetic variance to be equal across trials and constrains the genetic covariance between pairs of trials to be equal, providing only a measure of the magnitude of VxE interaction (Smith *et al.*, 2005). The two data sets in which the uniform model resulted in accurate VxE effects had nine and 10 trials, and in addition to this, had more consistent genetic correlations than the other data sets, meaning there was less loss of information due to averaging. However, for a data set with a similar number of trials but a much more variable VxE pattern these results may not hold.

Growing the same number of varieties at each trial, but differing these varieties between trials means that a larger number of varieties can be tested in a MET analysis, while still utilising the same resources. Collecting more data for

the same amount of work is a very appealing idea and consequently something that is regularly employed. Partial concurrence also frequently occurs within a breeding program due to selection over years. After a certain number of years a variety is either recommended for commercial release or dropped from the program and new varieties take their place, resulting in declining concurrence over years. If perfect concurrence was to be maintained over years new material could not enter the program, however, if too few varieties are in common between environments, the ability to accurately estimate genetic covariance is limited, not to mention unreliable and not representative. For this reason it was decided to investigate the impact of varietal concurrence on the accuracy of FA models when data sets contain small numbers of varieties per trial. Concurrence is not a factor that has been properly included in any of the existing studies, and has been highlighted as needing investigation (Smith *et al.*, 2015). Kelly *et al.* (2007) used perfect concurrence and the scenarios used by Welham *et al.* (2010) also contained high levels of varietal concurrence (86% and 92%).

Only two levels of varietal concurrence were investigated in this study, perfect concurrence in which all varieties were grown at every trial, and partial concurrence in which an average of 50% of varieties were in common between trials, meaning that the total unique varieties in the MET data set was double the number of varieties at any trial. Although only two levels of varietal concurrence were considered, there was a reasonable range in the numbers of varieties in common between two sites when partial concurrence was employed. When considering partial concurrence for a given variety number investigated, on average trials had half of this number of varieties in common. This resulted in trials having between two and nine varieties in common, between three and 12 varieties in common, between seven and 19 varieties in common, and between 16 and 35 varieties in common for the data generation models for 10, 15, 25, and 50 varieties per trial respectively.

Perfect concurrence resulted in a more accurate estimate of the VxE effects for the same variety number, however, it is important to note that these differences were generally small. It was expected that perfect concurrence would result in more accurate estimates of the VxE effects because fewer varieties (in total) were tested, hence more data was available on each variety. The results of the study suggest that the FA model is more heavily influenced by concurrence than the other models. For the uniform and diagonal model the difference in accuracy between perfect and partial concurrence is very small, if not non-existent. In contrast, the difference between concurrence levels for the FA model are more

pronounced. This is not surprising given it is the only model (aside from the unstructured model, which rarely converged) making a decent attempt at estimating the genetic covariance between environments. The range in the number of varieties in common means that for the scenarios with 10 or 15 varieties per trial in some cases the genetic covariance between trials was estimated based on the performance of only two common varieties. Despite this, the difference between the accuracy of the FA model with perfect concurrence and the FA model with partial concurrence for 10 and 15 varieties per trial tends to be similar to the change in accuracy when considering 25 or 50 varieties per trial. This suggests that even when trials have very low connectivity, provided on average they are reasonably connected, genetic covariance between trials can be adequately estimated. The concurrence patterns in this study were random, however, in practice a breeding program could be designed to test partially balanced sets of varieties across trials. This could potentially improve the accuracy of partial concurrence even more through a strategic allocation of concurrence.

For the different order of FA models the difference in accuracy between partial and perfect concurrence varies somewhat according to data set, as it does for the best FA model. For three data sets the difference between concurrence levels for the FA model was smaller for 10 and 15 varieties, and increased for 25 and 50 varieties, however, for the two remaining data sets the difference was greatest at 10 and 50 varieties. This suggests that concurrence level interacts with the underlying VxE patterns. However, the difference in accuracy between the concurrence levels for FA models is very small in practical terms, especially considering it allows for the testing of twice as many varieties. This suggests that any loss in accuracy is minimal and most plant breeders would most likely argue is worth it.

The results of this study indicate that the accuracy of the FA model is affected when there are small numbers of varieties per trial, however, these losses in accuracy for the most part have limited practical implications. The accuracy of the FA model compared to the other VxE models varied according to the underlying characteristics of the data set, such as the VxE pattern and the number of trials, however, in most cases for moderately small numbers of varieties (25 - 50) the FA model results in the most accurate estimate of the VxE effects. For very small numbers of varieties (10 - 15) the FA model is not always the most accurate model; however, in all cases for 15 varieties, and in nearly all for 10 varieties, its accuracy is of a practical level. It should also be noted that for very small numbers of varieties in nearly all cases the diagonal model for the variance of the VxE effects

was just as accurate as the FA model.

The MSEP was used in this study to determine which model resulted in the most accurate estimate of the VxE effects. However, when dealing with real data the MSEP cannot be used to compare models and the LLRT (or AIC) is instead used to compare model fit. There were many discrepancies between the models chosen using the two different methods. Despite the MSEP showing that in some cases the uniform and diagonal models can most accurately predict the VxE effects, the LLRT rarely ever selected either of these models for any data set. Among the FA models, the LLRT generally selected a model one or two orders higher than that selected by the MSEP. This suggests that models that are more complicated than necessary, and less accurate, are frequently being fitted in the analysis of MET data sets. For example the uniform model resulted in the best estimate of the VxE effects most often according to the MSEP for 10 varieties for the barley data set, however, using the LLRT an FA2 model was most often selected as offering the best fit to the data. These two models involve estimating two versus 29 variance parameters respectively. Smith *et al.* (2015) also noted the tendency of residual maximum likelihood ratio tests (and AIC) to lead to the selection of very high order models that are unnecessarily complicated. Both Smith *et al.* (2001) and Smith *et al.* (2015) suggest the ideal approach in the selection of the most appropriate order of FA model involves comparing a given model to the unstructured model, however, they note the difficulties in fitting this model, something that has been very evident in this study. Smith *et al.* (2015) suggest the need for an alternative test statistic, something that is the subject of their current research.

In order for variety number to be properly compared, within a given scenario all trials had the same number of varieties per trial. While this often occurs within a single year of a breeding program, it is frequently not the case across years. This may have some impact on the results, however, if the trials had similar numbers of varieties it is unlikely this would have a substantial effect, especially if larger numbers of varieties are being dealt with. Similarly the concurrence between trials was kept reasonably consistent to allow it to be best investigated, however, this may not reflect many data sets from the late stages of a crop breeding program.

4.1 Conclusions and future work

This study has provided important information regarding the analysis of MET data with small numbers of varieties. Although the aims of this study have been

successfully addressed, it has also highlighted areas for future work. Lower levels of concurrence could be investigated, along with a greater variability in concurrence patterns. As mentioned early, in this study concurrence was randomly allocated based on an overall average concurrence percentage. The impact of planning the breeding program to allow testing of partially balanced groups of varieties in a strategic way is worthy of further research. This study has highlighted the shortcomings of using the log-likelihood ratio test to select the most parsimonious model and future research could be conducted on a better model selection tool and the scenarios in which the MSEP and LLRT are in the greatest disagreement. In recent years the inclusion of pedigree information into the analysis of MET data has become more common (Oakey *et al.*, 2007). The impact the inclusion of such information would have on how the accuracy changes as variety numbers decrease is unknown and should be investigated in the future. In this study the variance-covariance structures used to represent different data sets were confounded with the different numbers of trials. This means that although differences between the data sets can be recognised, cause cannot be attributed solely to the underlying VxE pattern. A study could be structured with this aim in mind to allow this component to be further investigated.

This study has shown that 15 varieties per trial is sufficient in a MET data set to accurately estimate the VxE effects, and that in some cases MET data sets with even as few as 10 varieties could be considered (Aim 1). It was found that different data sets, characterised by the underlying pattern in the variance of the VxE effects and the number of trials, impacted on how the accuracy of the FA model compared to the accuracy of other models, especially for very small numbers of varieties (Aim 2). However, despite changes in the accuracy of the models according the the VxE pattern, the FA model was the superior model for moderately small numbers of varieties. While it was reasonably accurate for very small numbers of varieties, it was not necessarily the most accurate model. This study demonstrated that the FA model is affected by changes in concurrence more than the other models that were considered (Aim 3). Despite this, the impact of this change in accuracy is minimal, especially when considering the additional number of varieties for which data is collected when concurrence is lower.

References

Beeck, C. P., Cowling, W. A., Smith, A. B., & Cullis, B. R. (2010). Analysis of yield and oil from a series of canola breeding trials. Part I. Fitting factor analytic mixed models with pedigree information. *Genome* **53**, 992–1001.

Borgognone, M. G., Butler, D. G., Ogbonnaya, F. C., & Dreccer, M. F. (2016). Molecular marker information in the analysis of multi-environment trials helps differentiate superior genotypes from promising parents. *Crop Science* **56**, 2612–2628.

Butler, D., Cullis, B., Gilmour, A., & Gogel, B. (2009). Mixed models for S language environments, ASReml-R reference manual. Technical Report QE02001, Queensland Department of Agriculture, Fisheries and Forestry.

Byth, D. E., Eisemann, R. L., & De Lacy, I. H. (1976). Two-way pattern analysis of a large data set to evaluate genotypic adaptation. *Heredity* **37**, 215–230.

Christopher, J. T., Veyradier, M., Borrell, A. K., Harvey, G., Fletcher, S., & Chenu, K. (2014). Phenotyping novel stay-green traits to capture genetic variation in senescence dynamics. *Functional Plant Biology* **41**, 1035–1048.

Cullis, B., Gogel, B., Verbyla, A., & Thompson, R. (1998). Spatial Analysis of Multi-Environment Early Generation Variety Trials. *Biometrics* **54**, 1–18.

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

Cullis, B. R., Smith, A. B., Beek, C. P., & Cowling, W. A. (2010). Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaciont using factor analyis. *Genome* **53**, 1002–1016.

Finlay, K. W. & Wilkinson, G. N. (1963). The Analysis of Adaptation in a Plant-Breeding Programme. *Australian Journal of Scientific Research* **14**, 742–754.

Fox, G. P., Kelly, A., Poulsen, D., Inkerman, A., & Henry, R. (2006). Selecting for increased barley grain size. *Journal of Cereal Science* **43**, 198–208.

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

Gauch, H. J. (1992). *Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs*. Elsevier, Amsterdam.

Gilmour, A. R., Cullis, B. R., & Verbyla, A. P. (1997). Accounting for natural and extraneuous variation in the analysis of field experiments. *Journal of agricultural, biological, and environmental statistics* **2**, 269–273.

Gilmour, A. R., Thompson, R., & Cullis, B. R. (1995). Average Information REML: An Efficient Algorithm for Variance Parameter Estimation in Linear Mixed Models. *Biometrics* **51**, 1440–1450.

Gogel, B. J., Cullis, B. R., & Verbyla, A. P. (1995). REML Estimation of Multiplicative Effects in Multienvironment Variety Trails. *Biometrics* **51**, 744.

Kelly, A. M., Smith, A. B., Eccleston, J. A., & Cullis, B. R. (2007). The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Science* **47**, 1063–1070.

Kempton, R. A. (1984). The use of biplots in interpreting genotype by environment interactions. *Journal of Agricultural Science* **103**, 123–135.

Mardia, K., Kent, J., & Bibby, J. (1979). *Multivariate Analysis*. Academic Press, London.

Mungomery, V. E., Shorter, R., & Byth, D. E. (1974). Genotype x environment interactions and environmental adaptation. I. Pattern analysis - application to soya bean populations. *Australian Journal of Agricultural Research* **25**, 59–72.

Nabugoomu, F., Kempton, R. A., & Talbot, M. (1999). Analysis of series of trials where varieties differ in sensitivity to locations. *Journal of agricultural, biological, and environmental statistics* pages 310–325.

Oakey, H., Verbyla, A. P., Cullis, B. R., Wei, X., & Pitchford, W. S. (2007). Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. *Theoretical and Applied Genetics* **114**, 1319–1332.

Oakey, H., Verbyla, A. P., Pitchford, W. S., Cullis, B. R., & Kuchel, H. (2006). Joint modeling of additive and non-additive (genetic line) effects in single field trials. *Theoretical and Applied Genetics* **114**, 1319–1332.

Patterson, H. & Nabugoomu, F. (1992). REML and the analysis of series of crop variety trials. In *Proceedings from the 16th International Biometric Conference*, pages 77–93, Hamilton, New Zealand.

Patterson, H. & Silvey, V. (1980). Statutory and Recommended List Trials of Crop Varieties in the United Kingdom. *Journal of the Royal Statistical Society, A* **143**, 219–252.

Patterson, H. D., Silvey, V., Talbot, M., & Weatherup, S. T. C. (1977). Variability of yields of cereal varieties in U. K. trials. *The Journal of Agricultural Science* **89**, 239.

Patterson, H. D. & Thompson, R. (1971). Recovery of Inter-Block Information when Block Sizes are Unequal. *Biometrika* **58**, 545–554.

Piepho, H. P. (1997). Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* **53**, 761–766.

Piepho, H.-P., Denis, J.-B., & van Eeuwijk, F. a. (1998). Predicting Cultivar Differences Using Covariates. *Journal of Agricultural, Biological, and Environmental Statistics* **3**, 151.

Pike, E. W. & Silverberg, T. (1952). Designing meachanical computers. *Machine Design* **24**.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robinson, G. K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science* **6**, 15–32.

Rodda, M. S., Hobson, K. B., Forknall, C. R., Daniel, R. P., Fanning, J. P., Pounsett, D. D., Simpfendorfer, S., Moore, K. J., Owen, K. J., Sheedy, J. G., Thompson, J. P., Hollaway, G. J., & Slater, A. T. (2016). Highly heritable resistance to root-lesion nematode (Pratylenchus thornei) in Australian chickpea germplasm observed using an optimised glasshouse method and multi-environment trial analysis. *Australasian Plant Pathology* **45**, 309–319.

Searle, S. R. (1997). *Linear Models*. John Wiley & Sons.

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

Smith, A. B., Cullis, B. R., & Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science* **143**, 449.

Smith, A. B., Ganesalingam, A., Kuchel, H., & Cullis, B. R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **128**, 55–72.

Stefanova, K. T. & Buirchell, B. (2010). Multiplicative mixed models for genetic gain assessment in lupin breeding. *Crop Science* **50**, 880–891.

Thompson, J. P., Reen, R. A., Clewett, T. G., Sheedy, J. G., Kelly, A. M., Gogel, B. J., & Knights, E. J. (2011). Hybridisation of Australian chickpea cultivars with wild Cicer spp. increases resistance to root-lesion nematodes (Pratylenchus thornei and P. neglectus). *Australasian Plant Pathology* **40**, 601–611.

Thompson, R., Cullis, B., Smith, A., & Gilmour, A. (2003). A Sparse Implementation of the Average Information Algorithm for Factor Analytic and Reduced Rank Variance Models. *Australian & New Zealand Journal of Statistics* **45**, 445–459.

Verbyla, A. (1990). A Conditional Derivation of Residual Maximum Likelihood. *Australian Journal of Statistics* **32**, 227–230.

Welham, S. J., Gogel, B. J., Smith, A. B., Thompson, R., & Cullis, B. R. (2010). A comparison of analysis methods for late-stage variety evaluation triaLS. *Australian and New Zealand Journal of Statistics* **52**, 125–149.

Williams, E. J. (1952). The Interpretation of Interactions in Factorial Experiments. *Biometrika* **39**, 65–81.

Appendix A

Useful results

A.1 Joint normal distribution

If $y \sim \mathrm{N}\left(\mu, \Sigma\right)$ and $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, then

$$y_2|y_1 \sim \mathrm{N}\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(y_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$$

A.2 Orthogonal projection

If matrices X and L_2 are chosen such that $L_2'X = \mathbf{0}$ and H is positive definite then

$$H - HL_2\left(L_2'HL_2\right)^{-1}L_2'H = X\left(X'H^{-1}X\right)^{-1}X'$$

Proof (Verbyla, 1990)
If Equation 2.6 is transformed by the inverse square root of H, $H^{-1/2}$, orthogonal projections can be used. Then if $X^* = H^{-1/2}X$ a basis for the orthogonal complement of projections onto the column space of X^* is given by the columns of $L_2^* = H^{1/2}L_2$. Given that the orthogonal projection onto the orthogonal complement of the columns space of W is $I_n - W(W'W)^{-1}W'$, then

$$\begin{aligned}
L_2^*\left(L_2^{*\prime}L_2^*\right)^{-1}L_2^{*\prime} &= I_n - X^*\left(X^{*\prime}X^*\right)^{-1}X^{*\prime} \\
H^{1/2}L_2\left(L_2'H^{1/2}H^{1/2}L_2\right)^{-1}L_2'H^{1/2} &= I_n - H^{-1/2}X\left(X'H^{-1/2}H^{-1/2}X\right)^{-1}X'H^{-1/2} \\
H^{1/2}L_2\left(L_2'HL_2\right)^{-1}L_2'H^{1/2} &= I_n - H^{-1/2}X\left(X'H^{-1}X\right)^{-1}X'H^{-1/2} \\
H^{1/2}L_2\left(L_2'HL_2\right)^{-1}L_2'H^{1/2}H^{1/2} &= I_nH^{1/2} - H^{-1/2}X\left(X'H^{-1}X\right)^{-1}X'H^{-1/2}H^{1/2} \\
H^{1/2}L_2\left(L_2'HL_2\right)^{-1}L_2'H &= H^{1/2} - H^{-1/2}X\left(X'H^{-1}X\right)^{-1}X' \\
H^{1/2} - H^{1/2}L_2\left(L_2'HL_2\right)^{-1}L_2'H &= H^{-1/2}X\left(X'H^{-1}X\right)^{-1}X' \\
H^{1/2}H^{1/2} - H^{1/2}H^{1/2}L_2\left(L_2'HL_2\right)^{-1}L_2'H &= H^{1/2}H^{-1/2}X\left(X'H^{-1}X\right)^{-1}X' \\
H - HL_2\left(L_2'HL_2\right)^{-1}L_2'H &= X\left(X'H^{-1}X\right)^{-1}X'
\end{aligned}$$

From this it follows also that

$$\boldsymbol{I_n - HL_2\left(L_2'HL_2\right)^{-1}L_2' = X\left(X'H^{-1}X\right)^{-1}X'H^{-1}}$$

A.3 Derivative of $\boldsymbol{P}$

Given that

$$\boldsymbol{P = H^{-1} - H^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}}$$

the derivative of $\boldsymbol{P}$ is

$$\begin{aligned}
\frac{\partial \boldsymbol{P}}{\partial \kappa_i} =& \frac{\partial}{\partial \kappa_i}\left(\boldsymbol{H^{-1} - H^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}}\right) \\
=& \frac{\partial}{\partial \kappa_i}\left(\boldsymbol{H^{-1}}\right) - \frac{\partial}{\partial \kappa_i}\left(\boldsymbol{H^{-1}}\right)\boldsymbol{X\left(X'H^{-1}X\right)^{-1}X'H^{-1}} \\
& \boldsymbol{- H^{-1}X}\frac{\partial}{\partial \kappa_i}\left(\left(\boldsymbol{X'H^{-1}X}\right)^{-1}\right)\boldsymbol{X'H^{-1} - H^{-1}X\left(X'H^{-1}X\right)^{-1}X'}\frac{\partial}{\partial \kappa_i}\left(\boldsymbol{H^{-1}}\right) \\
=& \boldsymbol{- H^{-1}\dot{H}_iH^{-1} + H^{-1}\dot{H}_iH^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}} \\
& \boldsymbol{- H^{-1}X}\left(-\left(\boldsymbol{X'H^{-1}X}\right)^{-1}\frac{\partial}{\partial \kappa_i}\left(\boldsymbol{X'H^{-1}X}\right)\left(\boldsymbol{X'H^{-1}X}\right)^{-1}\right)\boldsymbol{X'H^{-1}} \\
& \boldsymbol{- H^{-1}X\left(X'H^{-1}X\right)X'\left(-H^{-1}\dot{H}_iH^{-1}\right)} \\
=& \boldsymbol{- H^{-1}\dot{H}_iH^{-1} + H^{-1}\dot{H}_iH^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}} \\
& \boldsymbol{- H^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}\dot{H}_iH^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}} \\
& \boldsymbol{+ H^{-1}X\left(X'H^{-1}X\right)X'H^{-1}\dot{H}_iH^{-1}} \\
=& \boldsymbol{- H^{-1}\dot{H}_i\left(H^{-1} - H^{-1}X\left(X'H^{-1}X\right)^{-1}X'H^{-1}\right)} \\
& \boldsymbol{+ H^{-1}X\left(X'H^{-1}X\right)X'H^{-1}\dot{H}_iH^{-1}\left(I_n - X\left(X'H^{-1}X\right)^{-1}X'H^{-1}\right)} \\
=& \boldsymbol{- H^{-1}\dot{H}_iP + H^{-1}X\left(X'H^{-1}X\right)X'H^{-1}\dot{H}_iP} \\
=& \boldsymbol{-\left(H^{-1} - H^{-1}X\left(X'H^{-1}X\right)X'H^{-1}\right)\dot{H}_iP} \\
=& \boldsymbol{- P\dot{H}_iP}
\end{aligned}$$

Appendix B

Matrix results

The following matrix results are drawn from Mardia *et al.* (1979).

B.1 Transpose

The transpose satisfies the properties

$$(\boldsymbol{A}')' = \boldsymbol{A} \qquad (\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}' \qquad (\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'.$$

If $\boldsymbol{A}$ is symmetric, then

$$\boldsymbol{A}' = \boldsymbol{A}.$$

B.2 Trace

The trace of $\boldsymbol{A}$ is given by

$$\operatorname{tr}(\boldsymbol{A}) = \sum a_{ij}$$

and satisfies the properties, when $\boldsymbol{A}(p \times p)$, $\boldsymbol{B}(p \times p)$, $\boldsymbol{C}(p \times n)$, and $\boldsymbol{D}(n \times p)$

$$\operatorname{tr}(\boldsymbol{A} \pm \boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A}) \pm \operatorname{tr}(\boldsymbol{B})$$
$$\operatorname{tr}(\boldsymbol{CD}) = \operatorname{tr}(\boldsymbol{DC}) = \sum_{i,j} c_{ij} d_{ij}$$

B.3 Determinants

A square matrix is non-singular if $|\boldsymbol{A}| \neq 0$, and the following result holds if a matrix is non-singular:

$$|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|.$$

B.4 Inverse

For non-singular matrices $\boldsymbol{A}(p \times p)$ and $\boldsymbol{C}(n \times n)$, and matrices $\boldsymbol{B}(p \times n)$ and $\boldsymbol{D}(n \times p)$,

$$(\boldsymbol{A} + \boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$$

B.5 Kronecker products

For matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$, the properties below hold:

(I) $\alpha(\boldsymbol{A} \otimes \boldsymbol{B}) = (\alpha\boldsymbol{A}) \otimes \boldsymbol{B} = \boldsymbol{A} \otimes (\alpha\boldsymbol{B})$ for scalar α

(II) $\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C}$

(III) $(\boldsymbol{A} \otimes \boldsymbol{B})' = \boldsymbol{A}' \otimes \boldsymbol{B}'$

(IV) $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{F} \otimes \boldsymbol{G}) = (\boldsymbol{A}\boldsymbol{F}) \otimes (\boldsymbol{B}\boldsymbol{G})$

(V) $(\boldsymbol{A} + \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{C} + \boldsymbol{B} \otimes \boldsymbol{C}$

(VI) $\boldsymbol{A} \otimes (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \otimes \boldsymbol{B} + \boldsymbol{A} \otimes \boldsymbol{C}$

B.6 Matrix differentiation

If the derivative of $f(\boldsymbol{X})$ with respect to $\boldsymbol{X}(n \times p)$ is defined as the matrix

$$\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{X}} = \left(\frac{\partial f(\boldsymbol{X})}{\partial x_{ij}}\right)$$

the following results hold

$$\frac{\partial \boldsymbol{a}'\boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{a}$$

$$\frac{\partial \boldsymbol{x}'\boldsymbol{x}}{\partial \boldsymbol{x}} = 2\boldsymbol{x} \qquad \frac{\partial \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{x} \qquad \frac{\partial \boldsymbol{x}'\boldsymbol{A}\boldsymbol{y}}{\partial \boldsymbol{x}} = \boldsymbol{A}\boldsymbol{y}$$

If $\boldsymbol{B}$ is a symmetric non-singular matrix whose elements are functions of $\boldsymbol{\kappa}$. Define $\dot{\boldsymbol{B}}_i = \partial \boldsymbol{B}/\partial \kappa_i$. Then the following results hold

$$\frac{\partial \log|\boldsymbol{B}|}{\partial \boldsymbol{\kappa}} = \operatorname{tr}\left(\boldsymbol{B}^{-1}\dot{\boldsymbol{B}}_i\right)$$

$$\frac{\partial \boldsymbol{B}^{-}1}{\partial \boldsymbol{\kappa}} = \operatorname{tr}\left(-\boldsymbol{B}^{-1}\dot{\boldsymbol{B}}_i\boldsymbol{B}^{-1}\right)$$

Appendix C

R code

C.1 Simulation code

```
fieldgen <- function(genodes,sitemn,evar,Ge){
    ########################################
    ## Code to generate data for a MET
    ########################################
    ## INPUT:
    ## genodes - list containing trial designs (one for each concurrence
    ##           level)
    ## sitemn - a vector of site means
    ## evar - a vector of error variances
    ## Ge - a genetic variance matrix

    ## OUTPUT:
    ## des - the inputted design with added simulated plot yields
    ## gxe - a matrix of sim gxe effects   m cols by p rows


    ## require necessary packages
    require(mvtnorm)
    require(MASS)

    g.check <- ldply(genodes,function(x){
        nm <- length(unique(x[["Trt"]]))        # no genos
        np <- length(unique(x[["Site"]]))       # no sites
        nn <- dim(x)[1]
        data.frame(nm,np,nn)
    })
    nm <- max(g.check$nm)

    ## generate genotype effects
    genetic <- mvrnorm(nm,mu=rep(0,unique(g.check$np)),Sigma=Ge)

    des.y <- list()
    for (i in names(genodes)){
        des <- genodes[[i]]
        ## y = Xt + Zu + e
```

```
        np <- length(unique(des[["Site"]]))     # no sites
        npi <- c(table(des[["Site"]])) # trial size for each site
        nn <- dim(des)[1]

        des <- des[order(des$Site,des$Column,des$Row),]
        ## plots within sites
        e <- matrix(rnorm(nn)*rep(sqrt(evar),npi),ncol=1)

        X <- matrix(0,ncol=np,nrow=nrow(des))
        for(j in 1:nrow(des)){
            X[j,des$Site[j]] <- 1
        }
        tau <- matrix(sitemn,ncol=1)

        ug <- matrix(genetic,ncol=1)
        Z <- matrix(0,ncol=nm*np,nrow=nrow(des))
        colnames(Z) <- paste(paste('S_',rep(1:np,each=nm),sep=''),
                             paste('G_',rep(1:nm,np),sep=''),sep='.')
        for(j in 1:nrow(des)){
            s <- des$Site[j]
            g <- des$Trt[j]
            ind <- paste(paste('S_',s,sep=''),paste('G_',g,sep=''),sep='.')
            Z[j,ind] <- 1
        }
        ## put it all together
        des$yvar <- X %*% tau + Z %*% ug + e

        des.y[[i]] <- des
    }
        list(des=des.y,gxe=genetic)
}
```

C.2 Code for analysis of simulated data

```
#########################################################
#########################################################
## code to compare models for different data sets
## and numbers of genotypes
#########################################################
#########################################################

## require packages
require(asreml)
require(myf)
options(width=185)
require(plyr)
require(reshape2)
require(ggplot2)
```

```
require(readxl)
source("simulate dataV3.R")
source("AIC.R")
source("sum_fa_func.R")

sitenumber <- data.frame(Set=c('Chickpea Kabuli','Chickpea Desi',
                               'Mungbean','Wheat','SA Barley'),
                         sites=c(39,18,9,52,10))
sitenumber$L <- paste(sitenumber$Set,sitenumber$sites,sep='_')

## import data
fielddes <- list()
site.mean <- list()
site.evar <- list()
site.Ge <- list()
con.sum <- list()

## sort out parameters
data.sets <- c("Chickpea Kabuli",'Chickpea Desi','Mungbean',
               'Wheat','SA Barley')
genos <- c(10,15,25,50)

## load up sets of parameters
for (k in data.sets){
  par <- read_excel(paste(k," data parameters.xlsx",sep=''),"Parameters")
  site.mean[[k]] <- par$mean
  site.evar[[k]] <- par$evar
  ge <- read_excel(paste(k," data parameters.xlsx",sep=''),
                   "Genetic variance")
  ge <- as.matrix(ge[,-1])
  dimnames(ge) <- list(NULL,NULL)
  site.Ge[[k]] <- ge

}

## load up designs
for (k in data.sets){
    wrkbook <- paste(k," designs.xlsx",sep='')
    s <- excel_sheets(wrkbook)
    fd <- list()
    for (i in s){
        fd[[i]] <- read_excel(wrkbook,i)
    }
    fielddes[[k]] <- fd
}

pred.alld <- list()
se.alld <- list()
```

```
gam.alld <- list()
con.alld <- list()
corel.alld <- list()
msep.alld <- list()
logl.alld <- list()
npar.alld <- list()
aic.alld <- list()
pvaf.alld <- list()
fa.level.alld <- list()
time.record.alld <- list()
model.time.alld <- list()
vxe.alld <- list()

for (k in data.sets){

    pred.allg <- list()
    se.allg <- list()
    gam.allg <- list()
    con.allg <- list()
    corel.allg <- list()
    msep.allg <- list()
    logl.allg <- list()
    npar.allg <- list()
    aic.allg <- list()
    pvaf.allg <- list()
    fa.level.allg <- list()
    time.record.allg <- list()
    model.time.allg <- list()
    vxe.allg <- list()

    for (h in genos){
        fd <- fielddes[[k]]
        conc <- names(fd)[grep(paste("G_",h,".",sep=''),names(fd))]
        fd <- fd[conc]

        g.check <- ldply(fd,function(x){
            nm <- length(unique(x[["Trt"]])) # no genos
            nm2 <- length(unique(x$Trt))

            nm.s <- unique(tapply(x$Trt,x$Site,function(d){
                length(unique(d))
            }))
            np <- length(unique(x[["Site"]])) # no sites
            nn <- dim(x)[1]
            data.frame(nm,nm.s,np,nn,nm2)
        })

        ns <- unique(g.check$np)
```

```
ng <- max(g.check$nm)

## maximum fa model that can be fitted
v <-  floor(0.5*(2*ns+1-sqrt(8*ns+1)))

## number of simulations
m <- 500
mod.names <- c('DIAG','CC',paste('FA',1:v,sep=''),'US')
n.mod <- length(mod.names)

vxe <- matrix(nrow=(ng*ns),ncol=m)
rownames(vxe) <- paste(rep(paste('Site',1:ns,sep='_'),each=ng),
                       rep(paste('Trt',1:ng,sep='_'),ns),sep=':')

pred <- list()
for(z in conc){
pred[[z]] <- list()
for (q in mod.names){
pred[[z]][[q]] <- matrix(nrow=(ng*ns),ncol=m)
rownames(pred[[z]][[q]]) <- paste(rep(paste('Site',1:ns,sep='_'),
                                      each=ng),
                                  rep(paste('Trt',1:ng,sep='_'),
                                    ns),sep=':')
    }
}

se <- list()
for(z in conc){
    se[[z]] <- list()
    for (q in mod.names){
    se[[z]][[q]] <- matrix(nrow=(ng*ns),ncol=m)
    rownames(se[[z]][[q]]) <- paste(rep(paste('Site',1:ns,sep='_'),
                                        each=ng),
                                    rep(paste('Trt',1:ng,sep='_'),
                                        ns),sep=':')
    }
}

gam <- list()
gam.s <- c(ns,2,((1:v+1)*ns),(ns*(ns+1)/2))
for(z in conc){
    gam[[z]] <- list()
    for (q in mod.names){
        gam[[z]][[q]] <- matrix(nrow=gam.s[
            which(q==mod.names)],ncol=m)
    }
}
```

```
con <- list()
for(z in conc){
    con[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(con[[z]]) <- mod.names
}

msep <- list()
for(z in conc){
    msep[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(msep[[z]]) <- mod.names
}

corel <- list()
for(z in conc){
    corel[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(corel[[z]]) <- mod.names
}

logl <- list()
for(z in conc){
    logl[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(logl[[z]]) <- mod.names
}

npar <- list()
for(z in conc){
    npar[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(npar[[z]]) <- mod.names
}

aic <- list()
for(z in conc){
    aic[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(aic[[z]]) <- mod.names
}

pvaf <- list()
for(z in conc){
    pvaf[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(pvaf[[z]]) <- mod.names
}

fa.level <- list()
time.record.start <- c()
time.record.finish <- c()
time.record <- c()

model.time <- list()
```

```
for(z in conc){
    model.time[[z]] <- matrix(nrow=n.mod,ncol=m)
    rownames(model.time[[z]]) <- mod.names
}

for (i in 1:500) {
    start.time <- Sys.time()

    fieldrun <- fieldgen2(fd,site.mean[[k]],site.evar[[k]],
                          site.Ge[[k]])
    des <- fieldrun$des
    gxe <- fieldrun$gxe
    vxe[,i] <- as.vector(gxe)

    for(o in conc){
        ## factorise
        fdes <- des[[o]]
        fdes$Trt <- factor(fdes$Trt)
        fdes$Site <- factor(fdes$Site)
        fdes$Column <- factor(fdes$Column)
        fdes$Row <- factor(fdes$Row)

        npi <- table(fdes$Site)
        np <- length(fdes$Trt)

        ## ################################
        ##   Start running models
        ## ################################
        ##==============##
        ## DIAG
        ##==============##
        model <- "DIAG"
        mt1 <- Sys.time()
        mod <- asreml(yvar ~ Site,
                      random = ~ diag(Site):Trt ,
                      rcov = ~ at(Site):Column:Row ,
                      data=fdes,workspace=80e6)
        mod <- update(mod)
        mt2 <- Sys.time()
        model.time[[o]][model,i] <- mt2 - mt1

        summary(mod)$varcomp

        con[[o]][model,i] <- mod$converge

        if(con[[o]][model,i]){
            prd <- predict(mod,class="Site:Trt",maxiter=1,
                           only='Site:Trt')
```

```
    prd <- prd$pred$pvals
    tt <- table(fdes$Trt,fdes$Site)
    prd$pres <- as.vector(tt)
    prd$predicted.value[prd$pres==0] <- NA
    prd$standard.error[prd$pres==0] <- NA
    prd$l <- paste('Site_',prd$Site,':Trt_',prd$Trt,
                   sep='')

    pred[[o]][[model]][
        match(prd$l,rownames(pred[[o]][[model]])),
        i] <- prd[order(prd$Site,prd$Trt),]$pred

    se[[o]][[model]][
        match(prd$l,rownames(pred[[o]][[model]])),
        i] <- prd[order(prd$Site,prd$Trt),]$stan

    gam[[o]][[model]][,i] <- mod$gammas[1:ns]
    if (is.null(rownames(gam[[o]][[model]]))){
        rownames(gam[[o]][[model]]) <- names(mod$gammas)[
            1:ns]}
    msep[[o]][model,i] <- sum((pred[[o]][[model]][,i]-
                                   as.vector(gxe))^2,
                              na.rm=TRUE)/
        (sum(!is.na(prd$predicted.value)))
    corel[[o]][model,i] <- cor(pred[[o]][[model]][,i],
                               as.vector(gxe),
                               use='pairwise.complete.obs')
    logl[[o]][model,i] <- mod$logl

    npar[[o]][model,i] <- sum(!(grepl('fa',rownames(
        summary(mod)$varcomp)) &
            summary(mod)$varcomp$constraint!='Fixed')
    aic[[o]][model,i] <- aic.fun(mod)
}else{
    model.time[[o]][model,i] <- NA
}

mod.d <- mod

##===============##
## CC
##===============##
model <- "CC"
mt1 <- Sys.time()
mod <- asreml(yvar ~ Site,
              random = ~ cor(Site):Trt ,
              rcov = ~ at(Site):Column:Row ,
              data=fdes,workspace=80e6)
```

```
mod <- update(mod)
mt2 <- Sys.time()
model.time[[o]][model,i] <- mt2 - mt1
con[[o]][model,i] <- mod$converge

## if didn't converge we don't want results
if(con[[o]][model,i]){
    prd <- predict(mod,class="Site:Trt",maxiter=1,
                   only='Site:Trt')
    prd <- prd$pred$pvals
    tt <- table(fdes$Trt,fdes$Site)
    prd$pres <- as.vector(tt)
    prd$predicted.value[prd$pres==0] <- NA
    prd$standard.error[prd$pres==0] <- NA
    prd$l <- paste('Site_',prd$Site,':Trt_',prd$Trt,sep='')

    pred[[o]][[model]][
        match(prd$l,rownames(pred[[o]][[model]])),
        i] <- prd[order(prd$Site,prd$Trt),]$pred
    se[[o]][[model]][
        match(prd$l,rownames(pred[[o]][[model]])),
        i] <- prd[order(prd$Site,prd$Trt),]$stan
    gam[[o]][[model]][,i] <- mod$gammas[1:2]
    if (is.null(rownames(gam[[o]][[model]]))){
        rownames(gam[[o]][[model]]) <- names(mod$gammas)[
            1:2]}
    msep[[o]][model,i] <- sum((pred[[o]][[model]][,i]-
                               as.vector(gxe))^2,
                              na.rm=TRUE)/
        (sum(!is.na(prd$predicted.value)))
    corel[[o]][model,i] <- cor(pred[[o]][[model]][,i],
                               as.vector(gxe),
                               use='pairwise.complete.obs')
    logl[[o]][model,i] <- mod$logl
    npar[[o]][model,i] <- sum(!(grepl('fa',rownames(
        summary(mod)$varcomp)) &
            summary(mod)$varcomp$constraint=='Boundary') &
            summary(mod)$varcomp$constraint!='Fixed')
    aic[[o]][model,i] <- aic.fun(mod)
}else{
    model.time[[o]][model,i] <- NA
}

##===============##
## FA
##===============##
model <- "FA1"
p <- 0
```

```
np <- c()
np.f <- c()
fa.logl <- c()
fa.con <- c()
g <- 1
mod.fa <- list()
## initial values from diag for fa1
mod.sv <- asreml(yvar ~ Site,
                 random = ~ fa(Site,g):Trt ,
                 rcov = ~ at(Site):Column:Row ,
                 data=fdes, start.values=TRUE)

ss <- c(1:ns)
diag.gam <- matrix(summary(mod.d,nice=T)$nice[["Site:Trt"]],
                   ncol=g)
dimnames(diag.gam) <- list(ss,c('psi'))
temp <- mod.sv$gammas.table
temp$Value[-grep('Trt',temp$Gamma)] <-  mod.d$gammas[
    -grep('Trt', names(mod.d$gammas))]
temp$Value[abs(temp$Value)<1e-5] <- .001
temp$Value[grep('*fa1',temp$Gamma)] <- c(rep(0.01,ns))
temp$Value[grep('Trt.*var',temp$Gamma)] <- diag.gam[,'psi']

mt1 <- Sys.time()
mod <- asreml(yvar ~ Site,
              random = ~ fa(Site,g):Trt ,
              rcov = ~ at(Site):Column:Row ,
              data=fdes,G.param=temp,R.param=temp,
              workspace=80e6)



for (l in 1:10){
    if(mod$converge) break
    mod <- update(mod)
}
mt2 <- Sys.time()
model.time[[o]][model,i] <- mt2 - mt1

summary(mod)$varcomp
np[g] <- length(summary(mod)$varcomp$constraint)
np.f[g] <- sum(summary(mod)$varcomp$constraint=="Fixed"
            |summary(mod)$varcomp$constraint=="Boundary")
fa.logl[g] <- mod$loglik

con[[o]][model,i] <- mod$converge
fa.con[g] <- mod$converge
```

```
if(con[[o]][model,i]){
    prd <- predict(mod,class="Site:Trt",maxiter=1,
                   only="fa(Site, g):Trt")
    prd <- prd$pred$pvals
    tt <- table(fdes$Trt,fdes$Site)
    prd$pres <- as.vector(tt)
    prd$predicted.value[prd$pres==0] <- NA
    prd$standard.error[prd$pres==0] <- NA
    prd$l <- paste('Site_',prd$Site,':Trt_',prd$Trt,
                   sep='')

    pred[[o]][[model]][
        match(prd$l,rownames(pred[[o]][[model]])),
        i] <- prd[order(prd$Site,prd$Trt),]$pred
    se[[o]][[model]][
        match(prd$l,rownames(pred[[o]][[model]])),
        i] <- prd[order(prd$Site,prd$Trt),]$stan
    gam[[o]][[model]][,i] <- mod$gammas[1:((g+1)*ns)]
    if (is.null(rownames(gam[[o]][[model]]))){
        rownames(gam[[o]][[model]]) <-
            names(mod$gammas)[1:((g+1)*ns)]
    }
    msep[[o]][model,i] <- sum((pred[[o]][[model]][,i]-
                                   as.vector(gxe))^2,
                              na.rm=TRUE)/
        (sum(!is.na(prd$predicted.value)))
    corel[[o]][model,i] <- cor(pred[[o]][[model]][,i],
                               as.vector(gxe),
                               use='pairwise.complete.obs')
    logl[[o]][model,i] <- mod$logl
    npar[[o]][model,i] <- sum(!(grepl('fa',rownames(
        summary(mod)$varcomp)) &
            summary(mod)$varcomp$constraint=='Boundary') &
            summary(mod)$varcomp$constraint!='Fixed')
    aic[[o]][model,i] <- aic.fun(mod)
    mod.sum <- sum.fa.func(mod)
    pvaf[[o]][model,i] <- mod.sum$'total %vaf'
}else{
    model.time[[o]][model,i] <- NA
}

mod.fa[[g]] <- mod
fac <- fa.con[g]
g <- 2


while (g<=v & p<0.05 & fac){
    model <- paste("FA",g,sep='')
```

```
mod.sv <- asreml(yvar ~ Site,
                 random = ~ fa(Site,g):Trt ,
                 rcov = ~ at(Site):Column:Row ,
                 data=fdes,start.values=TRUE)
## put gammas in matrix
fa.gam <- matrix(summary(mod,nice=T)$nice[[
    "fa(Site, g):Trt"]],ncol=(g))
dimnames(fa.gam) <- list(ss,c("psi",paste("lam",1:(g-1),
                                          sep='')))

temp <- mod.sv$gammas.table
temp$Value[-grep('Trt',temp$Gamma)] <- mod$gammas[
    -grep('Trt', names(mod$gammas))]
temp$Value[abs(temp$Value)<1e-5] <- .001

for (t in 1:(g-1)){
    temp$Value[grep(paste('*fa',t,sep=''),temp$Gamma)] <-
        fa.gam[,paste('lam',t,sep='')]
}
temp$Value[grep(paste('*fa',g,sep=''),temp$Gamma)] <-
    c(rep(0,g-1),rep(0.01,ns-(g-1)))
temp$Value[grep('Trt.*var',temp$Gamma)] <- fa.gam[,'psi']*0.8

mt1 <- Sys.time()
mod <- asreml(yvar ~ Site,
              random = ~ fa(Site,g):Trt ,
              rcov = ~ at(Site):Column:Row ,
              data=fdes,G.param=temp,R.param=temp,
              workspace=80e6)

for (t in 1:40){
    if(mod$converge) break
    mod <- update(mod)
}
mt2 <- Sys.time()
model.time[[o]][model,i] <- mt2 - mt1
summary(mod)$varcomp

## if didn't converge don't keep results
con[[o]][model,i] <- mod$converge
fa.con[g] <- mod$converge
if(con[[o]][model,i]){
     prd <- predict(mod,class="Site:Trt",maxiter=1,
                    only=paste("fa(Site, g):Trt",
                               sep=""))
     prd <- prd$pred$pvals
     tt <- table(fdes$Trt,fdes$Site)
     prd$pres <- as.vector(tt)
```

```
        prd$predicted.value[prd$pres==0] <- NA
        prd$standard.error[prd$pres==0] <- NA
        prd$l <- paste('Site_',prd$Site,':Trt_',
                       prd$Trt,sep='')

        pred[[o]][[model]][
            match(prd$l,rownames(pred[[o]][[model]])),
            i] <- prd[order(prd$Site,prd$Trt),]$pred
        se[[o]][[model]][
            match(prd$l,rownames(pred[[o]][[model]])),
            i] <- prd[order(prd$Site,prd$Trt),]$stan
        gam[[o]][[model]][,i] <- mod$gammas[1:((g+1)*ns)]
        if (is.null(rownames(gam[[o]][[model]]))){
            rownames(gam[[o]][[model]]) <-
                names(mod$gammas)[1:((g+1)*ns)]
        }
        msep[[o]][model,i] <- sum((pred[[o]][[model]][,i]-
                                       as.vector(gxe))^2,
                                  na.rm=TRUE)/
            (sum(!is.na(prd$predicted.value)))
        corel[[o]][model,i] <- cor(
            pred[[o]][[model]][,i],as.vector(gxe),
            use='pairwise.complete.obs')
        logl[[o]][model,i] <- mod$logl
        npar[[o]][model,i] <- sum(!(grepl('fa',rownames(
            summary(mod)$varcomp)) &
            summary(mod)$varcomp$constraint=='Boundary') &
            summary(mod)$varcomp$constraint!='Fixed')
        aic[[o]][model,i] <- aic.fun(mod)
        mod.sum <- sum.fa.func(mod)
        pvaf[[o]][model,i] <- mod.sum$'total %vaf'
    }else{
        model.time[[o]][model,i] <- NA
    }

    np[g] <- sum(!(grepl('fa',rownames(
        summary(mod)$varcomp)) &
            summary(mod)$varcomp$constraint=='Boundary') &
            summary(mod)$varcomp$constraint!='Fixed')
    dif <- (np[g]-np[g-1])
    fa.logl[g] <- mod$loglik
    logdiff <- 2*(fa.logl[g]-fa.logl[(g-1)])
    p <- 1- pchisq(logdiff, df=dif)
    mod.fa[[g]] <- mod
    fac <- fa.con[g]

    g <- g+1
    if (is.na(p)){p <- 1}
```

```
}

if (p>=0.05){
    fa.level[[o]][i] <- (g-2)
}else{
    fa.level[[o]][i] <- (g-1)
}

mod <- mod.fa[[fa.level[[o]][i]]]
mod.sum <- sum.fa.func(mod)
Gmat <- mod.sum$Gmat
us.init <- Gmat[!lower.tri(Gmat)]


##===============##
## US
##===============##
model <- 'US'
mod.sv <- asreml(yvar ~ Site,
                 random = ~ us(Site):Trt ,
                 rcov = ~ at(Site):Column:Row ,
                 data=fdes,start.values=TRUE)

temp <- mod.sv$gammas.table
temp$Value[1:(ns*(ns+1)/2)] <- us.init

mt1 <- Sys.time()
mod <- asreml(yvar ~ Site,
              random = ~ us(Site):Trt ,
              rcov = ~ at(Site):Column:Row ,
              data=fdes,G.param=temp,R.param=temp,
              workspace=80e6)

if(!mod$converge){
    mod <- update(mod)
    if(!mod$converge){
        for (t in 1:10){
            if (mod$converge) break
            mod <- update(mod)
        }

    }
}
mt2 <- Sys.time()
model.time[[o]][model,i] <- mt2 - mt1

## if didn't converge we don't want results
con[[o]][model,i] <- mod$converge
```

```
            if(con[[o]][model,i]){
                prd <- predict(mod,'Site:Trt',maxiter=1,
                               only='Site:Trt')
                prd <- prd$pred$pvals
                tt <- table(fdes$Trt,fdes$Site)
                prd$pres <- as.vector(tt)
                prd$predicted.value[prd$pres==0] <- NA
                prd$standard.error[prd$pres==0] <- NA
                prd$l <- paste('Site_',prd$Site,':Trt_',prd$Trt,sep='')

                pred[[o]][[model]][
                    match(prd$l,rownames(pred[[o]][[model]])),
                    i] <- prd[order(prd$Site,prd$Trt),]$pred
                se[[o]][[model]][
                    match(prd$l,rownames(pred[[o]][[model]])),
                    i] <- prd[order(prd$Site,prd$Trt),]$stan
                gam[[o]][[model]][,i] <- mod$gammas[1:(ns*(ns+1)/2)]
                if (is.null(rownames(gam[[o]][[model]]))){
                    rownames(gam[[o]][[model]]) <-
                        names(mod$gammas)[1:(ns*(ns+1)/2)]
                }
                msep[[o]][model,i] <- sum((pred[[o]][[model]][,i]-
                                           as.vector(gxe))^2,
                                          na.rm=TRUE)/
                    (sum(!is.na(prd$predicted.value)))
                corel[[o]][model,i] <- cor(pred[[o]][[model]][,i],
                                           as.vector(gxe),
                                           use='pairwise.complete.obs')
                logl[[o]][model,i] <- mod$logl
                npar[[o]][model,i] <- sum(!(grepl('fa',rownames(
                    summary(mod)$varcomp)) &
                        summary(mod)$varcomp$constraint=='Boundary') &
                        summary(mod)$varcomp$constraint!='Fixed')
                aic[[o]][model,i] <- aic.fun(mod)

            }else{
                model.time[[o]][model,i] <- NA
            }

        } ## end concurrence loop

        end.time <- Sys.time()
        time.record[i] <- end.time-start.time
        print(i)
        if((i %% 10)==0){
            save.image()
        }
    } ## end number simulations loop
```

```
time.record.allg[[paste('G_',h,sep='')]] <- time.record

pred <- ldply(pred,function(x){
    v <- ldply(x)
    colnames(v)[1] <- 'Model'
    v$Order <- rep(1:(ns*ng),length(mod.names))
    data.frame(v)
})
colnames(pred)[1] <- 'Conc'
colnames(pred)[c(-1,-2,-ncol(pred))] <- paste('Sim',1:m,sep='')
pred.allg[[paste('G_',h,sep='')]] <- pred

se <- ldply(se,function(x){
    v <- ldply(x)
    colnames(v)[1] <- 'Model'
    v$Order <- rep(1:(ns*ng),length(mod.names))
    data.frame(v)
})
colnames(se)[1] <- 'Conc'
colnames(se)[c(-1,-2,-ncol(se))] <- paste('Sim',1:m,sep='')
se.allg[[paste('G_',h,sep='')]] <- se

gam <- ldply(gam,function(x){
    v <- ldply(x,function(f){
        f <- data.frame(f);
        f$VC <- rownames(f);data.frame(f)
    })
    colnames(v)[1] <- 'Model'
    data.frame(v)
})
colnames(gam)[1] <- 'Conc'
colnames(gam)[c(-1,-2,-ncol(gam))] <- paste('Sim',1:m,sep='')
gam.allg[[paste('G_',h,sep='')]] <- gam

msep <- ldply(msep,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
colnames(msep)[1] <- 'Conc'
colnames(msep)[c(-1,-ncol(msep))] <- paste('Sim',1:m,sep='')
msep.allg[[paste('G_',h,sep='')]] <- msep

model.time <- ldply(model.time,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
```

```
colnames(model.time)[1] <- 'Conc'
colnames(model.time)[c(-1,-ncol(model.time))] <-
    paste('Sim',1:m,sep='')
model.time.allg[[paste('G_',h,sep='')]] <- model.time

corel <- ldply(corel,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
colnames(corel)[1] <- 'Conc'
colnames(corel)[c(-1,-ncol(corel))] <- paste('Sim',1:m,sep='')
corel.allg[[paste('G_',h,sep='')]] <- corel

logl <- ldply(logl,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
colnames(logl)[1] <- 'Conc'
colnames(logl)[c(-1,-ncol(logl))] <- paste('Sim',1:m,sep='')
logl.allg[[paste('G_',h,sep='')]] <- logl

npar <- ldply(npar,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
colnames(npar)[1] <- 'Conc'
colnames(npar)[c(-1,-ncol(npar))] <- paste('Sim',1:m,sep='')
npar.allg[[paste('G_',h,sep='')]] <- npar

aic <- ldply(aic,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
colnames(aic)[1] <- 'Conc'
colnames(aic)[c(-1,-ncol(aic))] <- paste('Sim',1:m,sep='')
aic.allg[[paste('G_',h,sep='')]] <- aic

pvaf <- ldply(pvaf,function(x){
    x <- data.frame(x)
    x$Model <- rownames(x)
    data.frame(x)
})
colnames(pvaf)[1] <- 'Conc'
colnames(pvaf)[c(-1,-ncol(pvaf))] <- paste('Sim',1:m,sep='')
```

```
    pvaf.allg[[paste('G_',h,sep='')]] <- pvaf

    con <- ldply(con,function(x){
        x <- data.frame(x)
        x$Model <- rownames(x)
        data.frame(x)
    })
    colnames(con)[1] <- 'Conc'
    colnames(con)[c(-1,-ncol(con))] <- paste('Sim',1:m,sep='')
    con.allg[[paste('G_',h,sep='')]] <- con

    fa.level <- ldply(fa.level)
    colnames(fa.level)[1] <- 'Conc'
    fa.level.allg[[paste('G_',h,sep='')]] <- fa.level

    vxe <- data.frame(vxe)
    vxe$Order <- 1:nrow(vxe)
    colnames(vxe)[c(-ncol(vxe))] <- paste('Sim',1:m,sep='')

    vxe.allg[[paste('G_',h,sep='')]] <- vxe

    save.image(paste(k,"_",h,".RData",sep=''))

} ## end geno loop

con.alld[[k]] <- ldply(con.allg)
colnames(con.alld[[k]])[1] <- 'genos'

pred.alld[[k]] <- ldply(pred.allg)
colnames(pred.alld[[k]])[1] <- 'genos'
se.alld[[k]] <- ldply(se.allg)
colnames(se.alld[[k]])[1] <- 'genos'
gam.alld[[k]] <- ldply(gam.allg)
colnames(gam.alld[[k]])[1] <- 'genos'
vxe.alld[[k]] <- ldply(vxe.allg)
colnames(vxe.alld[[k]])[1] <- 'genos'
msep.alld[[k]] <- ldply(msep.allg)
colnames(msep.alld[[k]])[1] <- 'genos'
model.time.alld[[k]] <- ldply(model.time.allg)
colnames(model.time.alld[[k]])[1] <- 'genos'
corel.alld[[k]] <- ldply(corel.allg)
colnames(corel.alld[[k]])[1] <- 'genos'
logl.alld[[k]] <- ldply(logl.allg)
colnames(logl.alld[[k]])[1] <- 'genos'
npar.alld[[k]] <- ldply(npar.allg)
colnames(npar.alld[[k]])[1] <- 'genos'
aic.alld[[k]] <- ldply(aic.allg)
colnames(aic.alld[[k]])[1] <- 'genos'
```

```
    pvaf.alld[[k]] <- ldply(pvaf.allg)
    colnames(pvaf.alld[[k]])[1] <- 'genos'
    fa.level.alld[[k]] <- ldply(fa.level.allg)
    colnames(fa.level.alld[[k]])[1] <- 'genos'
    time.record.alld[[k]] <- ldply(time.record.allg)
    colnames(time.record.alld[[k]])[1] <- 'genos'

} ## end data set loop
con.alld <- ldply(con.alld)
colnames(con.alld)[1] <- 'Set'
con.alld$Set <- sitenumber$L[match(con.alld$Set,sitenumber$Set)]

pred.alld <- ldply(pred.alld)
colnames(pred.alld)[1] <- 'Set'
pred.alld$Set <- sitenumber$L[match(pred.alld$Set,sitenumber$Set)]

vxe.alld <- ldply(vxe.alld)
colnames(vxe.alld)[1] <- 'Set'
vxe.alld$Set <- sitenumber$L[match(vxe.alld$Set,sitenumber$Set)]

se.alld <- ldply(se.alld)
colnames(se.alld)[1] <- 'Set'
se.alld$Set <- sitenumber$L[match(se.alld$Set,sitenumber$Set)]

gam.alld <- ldply(gam.alld)
colnames(gam.alld)[1] <- 'Set'
gam.alld$Set <- sitenumber$L[match(gam.alld$Set,sitenumber$Set)]

corel.alld <- ldply(corel.alld)
colnames(corel.alld)[1] <- 'Set'
corel.alld$Set <- sitenumber$L[match(corel.alld$Set,sitenumber$Set)]

msep.alld <- ldply(msep.alld)
colnames(msep.alld)[1] <- 'Set'
msep.alld$Set <- sitenumber$L[match(msep.alld$Set,sitenumber$Set)]

model.time.alld <- ldply(model.time.alld)
colnames(model.time.alld)[1] <- 'Set'
model.time.alld$Set <- sitenumber$L[match(model.time.alld$Set,
                                          sitenumber$Set)]

logl.alld <- ldply(logl.alld)
colnames(logl.alld)[1] <- 'Set'
logl.alld$Set <- sitenumber$L[match(logl.alld$Set,sitenumber$Set)]

npar.alld <- ldply(npar.alld)
colnames(npar.alld)[1] <- 'Set'
npar.alld$Set <- sitenumber$L[match(npar.alld$Set,sitenumber$Set)]
```

```
aic.alld <- ldply(aic.alld)
colnames(aic.alld)[1] <- 'Set'
aic.alld$Set <- sitenumber$L[match(aic.alld$Set,sitenumber$Set)]

pvaf.alld <- ldply(pvaf.alld)
colnames(pvaf.alld)[1] <- 'Set'
pvaf.alld$Set <- sitenumber$L[match(pvaf.alld$Set,sitenumber$Set)]

fa.level.alld <- ldply(fa.level.alld)
colnames(fa.level.alld)[1] <- 'Set'
colnames(fa.level.alld)[c(-1,-2,-3)] <- paste('Sim',1:m,sep='')
fa.level.alld$Set <- sitenumber$L[match(fa.level.alld$Set,sitenumber$Set)]

time.record.alld <- ldply(time.record.alld)
colnames(time.record.alld)[1] <- 'Set'
colnames(time.record.alld)[c(-1,-2)] <- paste('Sim',1:m,sep='')
time.record.alld$Set <- sitenumber$L[match(time.record.alld$Set,sitenumber$Set)]
```

Appendix D

Results

D.1 MSEP

Table D.1: Average mean square error of prediction for 500 simulations for the data generation models from the mungbean variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.0080	0.0080	0.0070	0.0071	0.0062	0.0063	0.0057	0.0058
DIAG	0.0078	0.0078	0.0068	0.0068	0.0060	0.0059	0.0053	0.0053
FA*	0.0083	0.0079	0.0068	0.0067	0.0058	0.0056	0.0050	0.0047
US				0.0049		0.0056	0.0045	0.0046
FA1	0.0077	0.0077	0.0065	0.0066	0.0056	0.0056	0.0050	0.0049
FA2	0.0083	0.0079	0.0068	0.0067	0.0058	0.0057	0.0050	0.0048
FA3	0.0083	0.0079	0.0069	0.0068	0.0059	0.0056	0.0050	0.0047
FA4	0.0086	0.0076	0.0067	0.0068	0.0059	0.0055	0.0050	0.0046
FA5		0.0071	0.0061	0.0069	0.0060	0.0056	0.0050	0.0047

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.2: Average mean square error of prediction for 500 simulations for the data generation models from the barley variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.0152	0.0155	0.0136	0.0136	0.0123	0.0123	0.0113	0.0112
DIAG	0.0158	0.0158	0.0140	0.0138	0.0125	0.0125	0.0113	0.0113
FA*	0.0176	0.0161	0.0147	0.0138	0.0123	0.0120	0.0106	0.0104
US				0.0141		0.0123	0.0112	0.0107
FA1	0.0160	0.0157	0.0137	0.0134	0.0118	0.0118	0.0105	0.0104
FA2	0.0175	0.0162	0.0145	0.0137	0.0122	0.0119	0.0106	0.0103
FA3	0.0182	0.0166	0.0151	0.0143	0.0125	0.0122	0.0108	0.0105
FA4	0.0198	0.0162	0.0151	0.0149	0.0126	0.0126	0.0110	0.0107
FA5		0.0195	0.0171	0.0142	0.0122	0.0126	0.0114	0.0112
FA6				0.0197	0.0119	0.0147	0.0122	0.0115

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.3: Average mean square error of prediction for 500 simulations for the data generation models from the Desi chickpea variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.0097	0.0097	0.0085	0.0087	0.0079	0.0079	0.0074	0.0075
DIAG	0.0088	0.0089	0.0076	0.0077	0.0067	0.0066	0.0061	0.0060
FA*	0.0077	0.0074	0.0059	0.0057	0.0046	0.0042	0.0033	0.0029
FA1	0.0070	0.0074	0.0055	0.0059	0.0047	0.0047	0.0039	0.0039
FA2	0.0071	0.0073	0.0053	0.0055	0.0042	0.0041	0.0033	0.0031
FA3	0.0078	0.0074	0.0058	0.0056	0.0044	0.0040	0.0032	0.0028
FA4	0.0084	0.0075	0.0063	0.0059	0.0048	0.0044	0.0034	0.0030
FA5	0.0097	0.0075	0.0068	0.0069	0.0051	0.0048	0.0036	0.0033
FA6			0.0078	0.0062	0.0053	0.0051	0.0039	0.0037
FA7						0.0061		0.0053

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.4: Average mean square error of prediction for 500 simulations for the data generation models from the Kabuli chickpea variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.0224	0.0225	0.0216	0.0219	0.0210	0.0211	0.0205	0.0205
DIAG	0.0146	0.0144	0.0131	0.0132	0.0119	0.0118	0.0109	0.0108
FA*	0.0152	0.0137	0.0130	0.0120	0.0111	0.0097	0.0086	0.0073
FA1	0.0138	0.0137	0.0121	0.0123	0.0106	0.0107	0.0095	0.0094
FA2	0.0150	0.0145	0.0124	0.0123	0.0104	0.0103	0.0089	0.0087
FA3	0.0155	0.0142	0.0129	0.0120	0.0103	0.0099	0.0084	0.0079
FA4	0.0162	0.0157	0.0132	0.0118	0.0105	0.0096	0.0083	0.0075
FA5	0.0181	0.0146	0.0138	0.0120	0.0108	0.0095	0.0082	0.0072
FA6	0.0197	0.0214	0.0147	0.0133	0.0111	0.0093	0.0082	0.0069
FA7		0.0147	0.0137	0.0142	0.0113	0.0096	0.0085	0.0073
FA8				0.0196	0.0116	0.0114	0.0088	0.0076
FA9							0.0089	0.0079
FA10							0.0088	0.0084
FA11								0.0088

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.5: Average mean square error of prediction for 500 simulations for the data generation models from the wheat variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.0424	0.0425	0.0409	0.0412	0.0405	0.0400	0.0395	0.0396
DIAG	0.0235	0.0233	0.0206	0.0211	0.0187	0.0186	0.0169	0.0170
FA*	0.0239	0.0229	0.0201	0.0195	0.0168	0.0150	0.0130	0.0106
FA1	0.0229	0.0229	0.0194	0.0200	0.0170	0.0171	0.0150	0.0151
FA2	0.0242	0.0236	0.0195	0.0198	0.0164	0.0163	0.0139	0.0137
FA3	0.0246	0.0228	0.0199	0.0196	0.0162	0.0156	0.0132	0.0128
FA4	0.0264	0.0248	0.0209	0.0194	0.0161	0.0150	0.0126	0.0119
FA5		0.0237	0.0219	0.0196	0.0164	0.0146	0.0120	0.0109
FA6		0.0283	0.0227	0.0207	0.0168	0.0144	0.0117	0.0101
FA7			0.0251	0.0214	0.0174	0.0148	0.0125	0.0106
FA8					0.0182	0.0162	0.0131	0.0112
FA9					0.0192		0.0136	0.0115
FA10							0.0141	0.0115
FA11							0.0147	0.0129

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

D.2 Correlation

Table D.6: Average correlation for 500 simulations for the data generation models from the mungbean variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.8778	0.8804	0.8919	0.8935	0.9033	0.9035	0.9103	0.9106
DIAG	0.8784	0.8815	0.8925	0.8965	0.9059	0.9068	0.9148	0.9152
FA*	0.8723	0.8812	0.8928	0.8982	0.9091	0.9132	0.9206	0.9270
US				0.9219		0.9150	0.9288	0.9274
FA1	0.8795	0.8836	0.8982	0.8997	0.9118	0.9122	0.9210	0.9223
FA2	0.8721	0.8805	0.8929	0.8978	0.9094	0.9116	0.9206	0.9239
FA3	0.8719	0.8799	0.8925	0.8975	0.9082	0.9131	0.9208	0.9255
FA4	0.8719	0.8866	0.8958	0.8992	0.9100	0.9151	0.9217	0.9272
FA5		0.8820	0.8984	0.9050	0.9074	0.9158	0.9220	0.9277

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.7: Average correlation for 500 simulations for the data generation models from the barley variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.8739	0.8761	0.8886	0.8913	0.9000	0.9013	0.9069	0.9082
DIAG	0.8670	0.8719	0.8841	0.8883	0.8973	0.8987	0.9057	0.9065
FA*	0.8549	0.8696	0.8788	0.8884	0.8991	0.9032	0.9114	0.9145
US				0.8932		0.9014	0.9002	0.9125
FA1	0.8657	0.8732	0.8863	0.8920	0.9028	0.9048	0.9125	0.9147
FA2	0.8547	0.8687	0.8799	0.8883	0.8996	0.9037	0.9116	0.9150
FA3	0.8510	0.8686	0.8756	0.8846	0.8976	0.9017	0.9105	0.9142
FA4	0.8474	0.8713	0.8784	0.8856	0.8984	0.9018	0.9084	0.9134
FA5		0.8180	0.8772	0.8887	0.9045	0.9015	0.9064	0.9120
FA6				0.8934	0.8925	0.8984	0.8954	0.9143

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.8: Average correlation for 500 simulations for the data generation models from the Desi chickpea variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.8363	0.8386	0.8571	0.8528	0.8660	0.8639	0.8748	0.8727
DIAG	0.8530	0.8526	0.8742	0.8715	0.8865	0.8875	0.8980	0.8995
FA*	0.8792	0.8845	0.9087	0.9112	0.9266	0.9337	0.9463	0.9539
FA1	0.8858	0.8804	0.9119	0.9044	0.9235	0.9233	0.9362	0.9374
FA2	0.8859	0.8855	0.9171	0.9129	0.9318	0.9341	0.9464	0.9507
FA3	0.8774	0.8835	0.9099	0.9117	0.9297	0.9366	0.9489	0.9562
FA4	0.8742	0.8850	0.9039	0.9086	0.9234	0.9306	0.9447	0.9521
FA5	0.8563	0.8931	0.8985	0.8957	0.9197	0.9244	0.9418	0.9485
FA6			0.8923	0.8962	0.9200	0.9186	0.9376	0.9402
FA7						0.9086		0.9285

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.9: Average correlation for 500 simulations for the data generation models from the Kabuli chickpea variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.7280	0.7252	0.7406	0.7381	0.7487	0.7473	0.7547	0.7546
DIAG	0.8261	0.8302	0.8469	0.8478	0.8604	0.8621	0.8722	0.8735
FA*	0.8295	0.8432	0.8547	0.8666	0.8779	0.8919	0.9036	0.9189
FA1	0.8398	0.8406	0.8613	0.8611	0.8774	0.8775	0.8900	0.8911
FA2	0.8301	0.8317	0.8594	0.8628	0.8807	0.8833	0.8968	0.9005
FA3	0.8268	0.8401	0.8557	0.8694	0.8826	0.8894	0.9033	0.9104
FA4	0.8198	0.8480	0.8537	0.8736	0.8817	0.8933	0.9054	0.9156
FA5	0.8156	0.8157	0.8482	0.8670	0.8795	0.8950	0.9068	0.9198
FA6	0.7864	0.8513	0.8480	0.8516	0.8775	0.8957	0.9073	0.9232
FA7		0.7826	0.8581	0.8279	0.8763	0.8913	0.9041	0.9193
FA8				0.8344	0.8756	0.8865	0.9016	0.9161
FA9							0.9016	0.9141
FA10							0.9038	0.9048
FA11								0.8949

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

Table D.10: Average correlation for 500 simulations for the data generation models from the wheat variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	0.7164	0.7113	0.7262	0.7254	0.7395	0.7419	0.7475	0.7471
DIAG	0.8428	0.8425	0.8611	0.8600	0.8776	0.8787	0.8895	0.8896
FA*	0.8452	0.8500	0.8698	0.8750	0.8957	0.9059	0.9187	0.9343
FA1	0.8493	0.8470	0.8707	0.8687	0.8897	0.8897	0.9027	0.9026
FA2	0.8434	0.8421	0.8716	0.8714	0.8947	0.8957	0.9105	0.9123
FA3	0.8427	0.8493	0.8709	0.8739	0.8966	0.9008	0.9153	0.9188
FA4	0.8363	0.8354	0.8666	0.8772	0.8981	0.9059	0.9197	0.9250
FA5		0.8437	0.8634	0.8767	0.8975	0.9088	0.9236	0.9313
FA6		0.8016	0.8648	0.8832	0.8956	0.9109	0.9260	0.9373
FA7			0.8756	0.8551	0.8937	0.9065	0.9218	0.9345
FA8					0.8904	0.8928	0.9182	0.9306
FA9					0.8884		0.9153	0.9285
FA10							0.9126	0.9237
FA11							0.9105	0.9279

* Average of FA models which offered the best fit to the data according to a log-likelihood ratio test.

D.3 Sample size of FA models

Table D.11: Number of times model was used in 500 simulations for the data generation models from the mungbean data set.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
FA1	500	500	500	500	500	500	500	500
FA2	500	500	500	500	500	500	500	500
FA3	459	384	428	419	386	449	446	495
FA4	105	110	183	247	208	342	305	470
FA5	0	17	4	79	40	177	132	394

Table D.12: Number of times model was used in 500 simulations for the data generation models from the barley data set.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
FA1	500	500	500	500	500	500	500	500
FA2	500	500	500	500	500	500	500	500
FA3	436	306	384	325	376	371	390	457
FA4	108	59	164	96	148	113	129	191
FA5	0	4	13	20	39	23	25	39
FA6	0	0	0	3	1	2	2	6

Table D.13: Number of times model was used in 500 simulations for the data generation models from the Desi chickpea data set.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
FA1	500	500	500	500	500	500	400	400
FA2	500	500	500	500	500	500	400	400
FA3	497	484	500	499	500	500	400	400
FA4	296	198	422	332	477	481	400	400
FA5	7	13	131	55	226	174	202	159
FA6	0	0	1	5	25	26	33	24
FA7	0	0	0	0	0	4	0	1

Table D.14: Number of times model was used in 500 simulations for the data generation models from the Kabuli chickpea data set.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
FA1	500	500	500	500	500	500	500	500
FA2	500	499	500	500	500	500	500	500
FA3	440	303	490	383	500	472	500	500
FA4	140	115	375	192	494	357	500	494
FA5	12	37	159	81	489	222	500	493
FA6	2	10	33	27	469	117	500	493
FA7	0	1	2	7	388	36	499	487
FA8	0	0	0	1	128	3	413	347
FA9	0	0	0	0	0	0	161	133
FA10	0	0	0	0	0	0	10	21
FA11	0	0	0	0	0	0	0	1

Table D.15: Number of times model was used in 500 simulations for the data generation models from the wheat data set.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
FA1	500	500	500	500	500	500	500	500
FA2	493	488	500	493	500	497	500	498
FA3	338	299	481	382	500	482	500	497
FA4	62	125	266	232	487	406	495	494
FA5	1	48	84	139	458	330	491	490
FA6	0	13	15	62	390	220	484	481
FA7	0	2	1	13	243	113	483	469
FA8	0	0	0	0	66	11	465	260
FA9	0	0	0	0	7	0	319	82
FA10	0	0	0	0	0	0	91	11
FA11	0	0	0	0	0	0	5	1

D.4 Convergence

Table D.16: Percentage of time model converged in 500 simulations for the data generation models from the mungbean variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	100	100	100	100	100	100	100	100
DIAG	100	100	100	100	100	100	100	100
FA1	99	99	99	100	100	99	100	100
FA2	100	100	100	100	100	100	100	100
FA3	98	100	100	100	100	100	100	100
FA4	91	98	89	100	100	100	100	100
FA5		76	75	97	80	99	100	100
US	0	0	0	1	0	10	1	53

Table D.17: Percentage of time model converged in 500 simulations for the data generation models from the barley variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	100	100	100	100	100	100	100	100
DIAG	100	100	100	100	100	100	100	100
FA1	100	100	100	100	100	100	100	100
FA2	100	100	100	100	100	100	100	100
FA3	99	100	100	100	100	100	100	100
FA4	97	100	99	100	100	100	100	100
FA5		100	92	100	92	100	100	100
FA6				100	100	100	100	100
US	0	0	0	40	0	78	1	90

Table D.18: Percentage of time model converged in 500 simulations for the data generation models from the Desi chickpea variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	100	100	100	100	100	100	100	100
DIAG	100	100	100	100	100	100	100	100
FA1	100	100	100	100	100	100	100	100
FA2	100	100	100	100	100	100	100	100
FA3	100	100	100	100	100	100	100	100
FA4	98	99	100	100	100	100	100	100
FA5	100	100	92	100	100	100	100	100
FA6			100	100	72	100	94	100
FA7						100		100
US	0	0	0	0	0	0	0	0

Table D.19: Percentage of time model converged in 500 simulations for the data generation models from the Kabuli chickpea variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	100	100	100	100	100	100	100	100
DIAG	100	100	100	100	100	100	100	100
FA1	100	100	100	100	100	100	100	100
FA2	100	96	100	98	100	100	100	100
FA3	100	65	100	96	100	100	100	100
FA4	96	62	100	82	100	100	100	100
FA5	100	53	100	65	100	100	100	100
FA6	100	50	100	70	100	96	100	100
FA7		100	100	86	100	94	100	100
FA8				100	99	100	100	100
FA9							100	99
FA10							100	100
FA11								100
US	0	0	0	0	0	0	0	0

Table D.20: Percentage of time model converged in 500 simulations for the data generation models from the wheat variance-covariance structure.

Model	10 Varieties		15 Varieties		25 Varieties		50 Varieties	
	Partial	Perfect	Partial	Perfect	Partial	Perfect	Partial	Perfect
UNIF	100	100	100	100	100	100	100	100
DIAG	100	100	100	100	100	100	100	100
FA1	99	98	100	99	100	99	100	100
FA2	98	75	100	98	100	100	100	100
FA3	88	53	99	81	99	98	99	100
FA4	45	54	89	69	98	94	99	99
FA5	0	54	85	50	97	85	99	98
FA6		23	87	29	95	75	100	99
FA7		0	100	23	84	54	99	93
FA8					88	55	98	88
FA9					100		99	85
FA10							99	73
FA11							100	100
US	0	0	0	0	0	0	0	0