Plant Breeding

doi:10.1111/pbr.12251

Trial heterogeneity and variance models in the genetic evaluation of potato tuber yield

Mark F. Paget[1,5], Peter A. Alspach[2], John A.D. Anderson[3], Russell A. Genet[1] and Luis A. Apiolaza[4]

[1]The New Zealand Institute for Plant & Food Research Limited, Lincoln, Christchurch, 8140, New Zealand; [2]The New Zealand Institute for Plant & Food Research Limited, Motueka, 7198, New Zealand; [3]The New Zealand Institute for Plant & Food Research Limited, Pukekohe, 2676, New Zealand; [4]School of Forestry, University of Canterbury, Christchurch, New Zealand; [5]Corresponding author, E-mail: mark.paget@plantandfood.co.nz

With 4 figures and 3 tables

Received July 16, 2014 / Accepted December 16, 2014
Communicated by H.-P. Piepho

Abstract

Genetic evaluation aims to identify genotypes with high empirical breeding values (EBVs) for selection as parents. In this study, 2157 potato genotypes were evaluated for tuber yield using 8 years of early-stage trial data collected from a potato breeding programme. Using linear mixed models, spatial parameters to target greater control of localised spatial heterogeneity within trials were estimated and variance models to account for across-trial genetic heterogeneity were tested. When spatial components improved model fit, correlations of errors were mostly small and negative for marketable tuber yield (MTY) and total tuber yield (TTY), suggesting the presence of interplot competition in some years. For the analysis of multi-environment trials, a variance model with a simple correlation structure (with heterogeneous variances) was the most favourable variance structure fitted for TTY and PTY (per cent marketable yield). There was very little difference in model fit when comparing a factor analytic structure of order 2 (FA2) with either FA1 or simple correlation structures for MTY, indicating that simple variance models may be preferable for early-stage genetic evaluation of potato yield.

Key words: empirical breeding values — genotype-by-environment — multi-environment trials — potato selection — spatial models

Potato tuber yield is an important selection criterion and contributes to maximising the (implicit) economic objective in potato breeding programmes. It is typical that information on tuber yield under either a phenotypic or 'genotypic' (progeny test) selection strategy is not available until at least the second clonal stage of a breeding scheme, when there are enough tubers available for establishing formal replicated trials. Under a traditional phenotypic recurrent selection strategy, the use of a promising parental candidate will often be delayed until the breeder has enough confidence in its individual 'production worth' after further years of extensive field trials. Furthermore, the candidate's 'breeding worth' is not necessarily formally evaluated from the performance of its progeny when this information eventually comes to hand. Evaluation of potato genotypes may therefore fail to take advantage of all available information in multi-environment trial (MET) data, which might improve the precision of breeding value estimation in the early stages of testing; trial analyses often assume the independence of genotypes both within and across trials, but these assumptions are not realistic (Smith et al. 2005). Genetic evaluation of yield and other traits using trial data and ancestry information can provide predictions of breeding and genotype values from informal mating designs (e.g. Crossa et al. 2006, Oakey et al. 2006, 2007, Kelly et al. 2009). Furthermore, observations that the residuals of neighbouring plots in field trials are often more alike than those of non-neighbours have led to the development of a number of statistical approaches to deal with this localised trial heterogeneity that augment standard blocking in trial designs (e.g. Gleeson 1997, Edmondson 2004).

For genetic evaluation, different variance structures can be set up within a mixed model to accommodate the genetic (co)variances that exist among trials or 'environments' in MET data (e.g. Smith et al. 2001, Crossa et al. 2006, Kelly et al. 2007, Meyer 2009), allowing varying degrees of complexity to be modelled. Trial evaluation may be enhanced, for example, by fitting a homogeneous covariance structure that models different within-trial variances and the same genetic correlation between trials. This structure may be relatively simple to fit but may not be reasonable when trials are performed over diverse environments, and the assumed homogeneous genetic correlation structure does not adequately deal with the genetic heterogeneity that may exist. The most general form is an unstructured (US) covariance matrix, which models both heterogeneity of trial variance and different covariances for each pairwise combination of trials, but is recognised as computationally difficult to fit. An alternative is the factor analytic (FA) model (Piepho 1998, Smith et al. 2001). To simplify the calculations, the FA approach attempts to confine the genotype-by-environment (G×E) interaction effects into a small number of components (unobserved latent variables) that aim to explain most of the interaction and in this respect is analogous to ordination methods previously developed to study G×E, such as the additive main effects and multiplicative interaction (AMMI) model (Gauch and Zobel 1988, Crossa et al. 1991).

The general aim of this study was to identify an appropriate genetic evaluation model for analysing MET potato yield data for early-stage selection in a potato breeding programme. The study used 21 trials from 8 years of field data, which included 2157 genotypes, from the early-stage selection trials of The New Zealand Institute for Plant & Food Research (PFR) potato breeding programme. Spatial models were investigated for greater control of local-scale heterogeneity within potato field trials, and different variance structures were modelled to account for across-trial heterogeneity of MET field data.

Materials and Methods

Data: Trials were performed over a number of years (1999–2003, 2006–2007 and 2012) mostly at Pukekohe, South Auckland (37°.12′S 174°.52′E, 141 m asl) but with two trials in Palmerston North, Manawatu (40°.20′S 175°.28′E, 30 m asl). The 21 trials were designed as early-clonal stage two, three and four (C2, C3 and C4, respectively)

'early-main' (EM) crop and 'main' (MN) crop yield trials. Early-main crop trials were planted in mid to late September and harvested in late February, approximately 150 days after planting. Main crop trials were planted in early November and harvested in mid May, with weather conditions sometimes delaying harvest into June. C2 trials were treated as early-main crop trials. Selected genotypes from the C2 stage were entered into main crop (Manawatu: MW) and early-main and main crop (Pukekohe: PK) trials at the C3 and C4 stages. Each trial comprised a rectangular array of rows by columns, typically of 60–90 genotypes replicated twice, designed as a resolvable latinised row-column with CycDesigN v4.0 (CycSoftware 2009) and previous versions of the trial design software. Each plot was made up of 12 tubers in total, planted in a six by two arrangement, with a width of 1.55 m and a length of 2.0 m. Spacing between neighbouring plots on the shorter plot side was 0.58 m and on the longer plot side was 0.77 m. Plot yield was recorded at harvest as both a total tuber yield (TTY) and a marketable tuber yield (MTY) and converted to t ha^{-1} (metric tonnes per hectare) for analyses. MTY, the trait of most interest, was the saleable (graded) yield after undersized (<80 g), and defective tubers had been removed. Yield was also expressed as the percentage marketable fraction of the total yield (PTY), and logit transformed so that the response used for analysis was $ln(p/(1\text{-}p))$, where p is the proportion of MTY to TTY.

TTY, MTY and PTY were analysed in 21 early-stage potato breeding trials for the estimation of variance components and spatial parameters. Fifteen of these trials (1999–2003) showed reasonable concurrence of genotypes across both trials years (Table 1). This representative series of early-stage trials were therefore used to test different variance structures to account for trial heterogeneity and to estimate breeding values for potato yield for both (1050) tested genotypes and all genotypes in the pedigree.

Single-trial analysis: Single trials were analysed to estimate variance components for each trial, as represented by a general form of the linear mixed model:

$$\mathbf{y} = \mathbf{1}m + \mathbf{Z_1 b} + \mathbf{Z_2 g} + \mathbf{e},$$

where $\mathbf{y}$ is the $n \times 1$ vector of yield observations, m is the overall trial mean as a fixed effect, $\boldsymbol{b} \sim N(0, \mathbf{I}\sigma^2_{\mathbf{b}})$ and $\boldsymbol{g} \sim N(0, \mathbf{I}\sigma^2_{\mathbf{g}})$ are $q \times 1$ and $w \times 1$ vectors that represent random (non-genetic) design factors, for example replicate/row/column and genotypic effects, respectively, and $\boldsymbol{e} \sim N(0, \mathbf{I}\sigma^2_{\mathbf{e}})$ is the $n \times 1$ vector of random error terms. $\mathbf{Z_1}$ ($n \times q$) and $\mathbf{Z_2}$ ($n \times w$) are known incidence matrices of the random effects (trial design and genetic effects), and $\mathbf{I}$ are the relevant $q \times q$, $w \times w$ or $n \times n$ identity matrices.

A randomisation-based approach (the base model) was first used in analyses to reflect the experimental design. This included the independent row and column effects, respecting the latinisation of the trials, and the complete replicate effects if necessary. The *base* model was then compared with an *extended* row × column model that included random row and column effects, but was augmented by row and/or column spatial correlation parameters as an attempt to better describe localised heterogeneity. A separable autoregressive process of order one (AR1) has previously been shown to provide a suitable variance structure for local spatial trend in annual crop evaluation trials (Gilmour et al. 1997, Smith et al. 2001) and, in general, compare well with alternative linear variance models (Müller et al. 2010). Following Gilmour et al. (1997) and Smith et al. (2001), spatial analysis, **e,** was decomposed into spatially dependent and spatially independent errors. The following matrix shows the pattern of spatially dependent errors, modelled as the AR1 correlation coefficients (ρ) between plots in the same ordered column or row (q) of size n:

$$\mathbf{AR1}\,(\rho_{\mathbf{q}}) = \begin{bmatrix} \mathbf{1} & \rho_{\mathbf{q}} & \rho^2_{\mathbf{q}} & \cdots & \rho^{\mathbf{n-1}}_{\mathbf{q}} \\ \rho_{\mathbf{q}} & \mathbf{1} & \rho_{\mathbf{q}} & \cdots & \vdots \\ \rho^2_{\mathbf{q}} & \rho_{\mathbf{q}} & \mathbf{1} & \mathbf{1} & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{\mathbf{n-1}}_{\mathbf{q}} & \cdots & \cdots & \cdots & \mathbf{1} \end{bmatrix}$$

This is generalised to give the correlation coefficient between plots not located in the same row or in the same column as $\rho_r^{|i-i'|}\rho_c^{|j-j'|}$ for plots separated by $|i\text{-}i'|$ rows and $|j\text{-}j'|$ columns from the direct product **AR1** ⊗ **AR1**. The best fitting model was selected as the preferred model on the basis of the Akaike information criterion (AIC) goodness-of-fit test: AIC = $-2(logl - N_p)$, where *logl* is the REML estimate of the log-likelihood and N_p is a penalty term representing the number of variance parameters fitted. Smaller values of AIC represented a better fitting model. The model was then revised with any large-scale field trend (large-scale dependence or global trend in the mean of errors) present across rows and/or columns accounted for by either fitting fixed linear or polynomial regressions (Gilmour et al. 1997), or fixed linear regressions and cubic smoothing splines (Verbyla et al. 1999) to the spatial coordinates. From the inspection of residual variograms, global trends were detected by non-stationarity and fixed linear and polynomial global terms were tested using approximate incremental F-tests based on Wald statistics, with non-significant fixed regression terms sequentially dropped from the model. Splines were tested with the AIC. An *extended* model was therefore considered as a model that included localised spatial error components and/or global field trends, as well as row and column block effects.

The modelling procedures then incorporated the pedigree. Models were tested using an 'individual plant' model, with $\mathbf{I}\sigma^2_{\mathbf{g}}$ replaced by $\mathbf{A}\sigma^2_{\mathbf{a}}$, the variance–covariance matrix of the additive genetic effects (breeding values), where $\mathbf{A}$ as the numerator relationship matrix that provides the between-genotype relationship as two times the coefficient of coancestry. Empirical breeding values (EBVs) were obtained from the BLUPs of genotype effects, (e.g. Smith et al. 2005, p.458). As variance components were unknown, empirical breeding values resulted from applying variance components in the mixed model equations that were estimated from the data, thus giving empirical BLUEs (best linear unbiased estimators) of fixed effects and empirical BLUPs of random effects. The changes in ranking using Spearman's rank correlations for all tested genotypes and the percentage concurrence of genotypes selected between the *extended* and *base* models (when the top 10 per cent of genotypes ranked on EBV were selected from each analysis) were also measured.

Table 1: Concurrence of genotypes across 5 years of potato tuber yield trials; diagonal entries are the number of genotypes tested in individual years

	1999	2000	2001	2002	2003
1999	462	114	33	19	8
2000		577	158	64	26
2001			233	101	55
2002				131	68
2003					89

Variance models for MET analysis: The single-trial models were then expanded into a multivariate MET analysis by analysing 15 trials of yield observations $(\mathbf{y_1}, \mathbf{y_2}, \ldots \mathbf{y_{15}})$ between 1999 and 2003. These trials were used to represent data obtained from early-stage selection trials, providing an initial opportunity for a breeder to obtain breeding values for yield of newly tested parents. Therefore, all trials after 2003 were excluded from analysis (trial PK-EM-00 was excluded, as only a small number of genotypes in this trial were shared with other trials), with data vectors and design matrices constructed as follows:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{15} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_1 m_2 \\ \mathbf{1}_2 m_2 \\ \vdots \\ \mathbf{1}_{15} m_{15} \end{bmatrix} + \begin{bmatrix} \mathbf{Z_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{\mathbf{1}_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z}_{\mathbf{1}_{15}} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_{15} \end{bmatrix}$$
$$+ \begin{bmatrix} \mathbf{Z}_{2_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{2_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z}_{2_{15}} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{15} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_{15} \end{bmatrix}$$

Random effects were assumed to follow a multivariate normal distribution with means and variances defined by:

$$\begin{pmatrix}\mathbf{b}\\ \mathbf{a}\\ \mathbf{e}\end{pmatrix} \sim \mathbf{N}\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{bmatrix}\mathbf{B_0}\otimes\mathbf{I_b} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G_0}\otimes\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{R}\otimes\mathbf{I_n}\end{bmatrix}\right]$$

where $\mathbf{0}$ are null matrices. $\mathbf{B_0}$, $\mathbf{G_0}$ and $\mathbf{R}$ are covariance matrices for design factors, genetic (additive) and residual effects, respectively and $\otimes$ is the direct (Kronecker) product. The matrix $\mathbf{B_0}$ is a diagonal matrix of (non-genetic) scaled identity matrices, and the variance–covariance structure of plot error effects $\mathbf{R}$ is assumed to be block diagonal. As each trial (t) analysed comprised a rectangular array of r_t rows by c_t columns ($n_t = r_t c_t$), local spatial trend, as described for the single-trial analyses outlined previously, was specified through $\mathbf{R}$ as a separable AR1 process (Gilmour et al. 1997, Smith et al. 2001) and included the $c_t \times c_t$ and r_t x r_t correlation matrices associated with the coordinates of the column and row layout of the trials, respectively, and the independent measurement error variance for trial t.

The assumption was that the variance matrix for the additive genotype effects has the separable form $\mathbf{G_a} = \mathbf{G_0}\otimes\mathbf{A}$ (e.g. Kelly et al. 2009), where $\mathbf{G_0}$ is the matrix of additive variances and covariances between environments and $\mathbf{A}$ is the covariance matrix between genotypes – the numerator relationship matrix. Non-pedigree-based models were also tested, so that the independent genotype effects were of the form $\mathbf{G_g} = \mathbf{G_0}\otimes\mathbf{I}$, where $\mathbf{I}$ in this particular case is an identity matrix of order g (number of genotypes). Using the important non-genetic terms identified from each single-trial analysis, four forms of the genetic variance matrix were then compared with each other. $\mathbf{G_0}$ is the genetic variance matrix, with the diagonal elements representing genetic variances for each trial and the off-diagonal elements representing genetic covariances between pairs of trials. Definitions of the forms of $\mathbf{G_0}$ tested are as follows:

SIMPLE: all variances within trials are assumed to be equal, and all pairwise covariances between trials are assumed to be independent and therefore zero:

$$\mathbf{G_0} = \begin{bmatrix}\sigma^2 & 0 & \cdots & 0\\ 0 & \sigma^2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma^2\end{bmatrix}$$

DIAG: variances within trials are assumed to be different, and all pairwise covariances between trials are assumed to be zero:

$$\mathbf{G_0} = \begin{bmatrix}\sigma_1^2 & 0 & \cdots & 0\\ 0 & \sigma_2^2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma_t^2\end{bmatrix}$$

CORH: variances within trials are assumed to be different, and a constant non-zero correlation is assumed between all pairwise combinations of trials:

$$\mathbf{G_0} = \begin{bmatrix}\sigma_1^2 & \rho & \cdots & \rho\\ \rho & \sigma_2^2 & \cdots & \rho\\ \vdots & \vdots & \ddots & \vdots\\ \rho & \rho & \cdots & \sigma_t^2\end{bmatrix}$$

where ρ represents a constant correlation of additive genetic effects between trials.

FAk: factor analytic, a parsimonious approximation to the US genetic (co)variance matrix (Piepho 1997, 1998, Smith et al. 2001), which identifies common factors (as the leading principal components) and residuals, or 'specific variances', and is given by: $\mathbf{G_0} = \mathbf{\Lambda}\mathbf{\Lambda}' + \boldsymbol{\psi}$; where $\mathbf{\Lambda}$ is a (t x k) matrix of environmental loadings (or common factors):

$$\mathbf{\Lambda} = \begin{bmatrix}\Phi_{11} & \Phi_{12} & \cdots & \Phi_{1k}\\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2k}\\ \vdots & \vdots & \ddots & \vdots\\ \Phi_{t1} & \Phi_{t2} & \cdots & \Phi_{15\mathbf{k}}\end{bmatrix}$$

and $\boldsymbol{\psi}$ is a (t x t) diagonal matrix of specific variances, with constraints placed on the FA loadings for identifiability when $k > 1$.

Pedigree information was built from historic PFR field books, and a publicly available pedigree database (van Berloo et al. 2007). R (R Development Core Team 2012) was used for data analyses, with the mixed models fitted using ASReml-R (Butler et al. 2009). AIC was used as the test criterion for the various forms of $\mathbf{G_0}$. Variance–covariance models were also compared by simulating the response to selection (Piepho and Möhring 2007). The best linear unbiased predictors (BLUPs) of breeding values were obtained as the arithmetic average across environments of the predictions for each environment. With the assumption that breeding values estimated from the data were the 'true' values, residuals at each simulation round (1000 rounds) were resampled with replacement and added to the fitted values. The simulated data were then re-analysed to provide the BLUPs of breeding values, and at selection fraction s, the top (ranked) s100% genotypes based on the simulated BLUPs, were identified. The simulated BLUPs were then replaced with the true BLUPs for the selected group of genotypes. The difference between the true breeding value mean of the selected genotypes and the true mean of the breeding population was considered to be the response to selection.

Results

Single-trial analysis

The distributions for MTY are illustrated in Fig. 1 and showed a high degree of yield variability across trials.

Table 2 shows the fitted fixed effects, variance components and spatial parameter estimates for the preferred (*base* or *extended*) model from single-trial analyses for MTY (results not tabulated for TTY and PTY). The proportion of additive (V_A) to phenotypic (total) variance (V_P), a measure of narrow-sense heritabilities, for MTY and TTY (0.57–0.88) were high for all trials. Heritability values, in general, were lower for PTY (ranging from 0.27 to 0.69). Including an independent measurement error component frequently resulted in convergence problems and was therefore not included in analyses. Fixed linear regressions, in either rows or columns or both, were included for over half of all trials. A second-order polynomial regression (rows) was chosen for trial PK-MN-01B for both TTY and MTY, which appeared to account for curvature present across the trial, as observed from the variogram of residuals. Random spline (row) effects were found to be important in only one trial (MW-MN-02) for TTY and MTY, but there was very little effect on the re-ranking of genotype breeding values in this case. For TTY, the percentage of common genotypes selected from both models was greater than 80 for all trials with the exception of PK-C2-99A (74%). The range of percentage concurrence of genotypes selected across all 21 trials was greater for both MTY and PTY than for TTY and included ranges of between 67 and 100 and 66 and 100 for MTY and PTY, respectively. AR1 correlation coefficients, when considered to improve model fit, were generally small overall, that is mostly between −0.35 and +0.30, but featured more for rows (the shorter plot to plot distance) than for columns, in general. These spatial correlation estimates were all

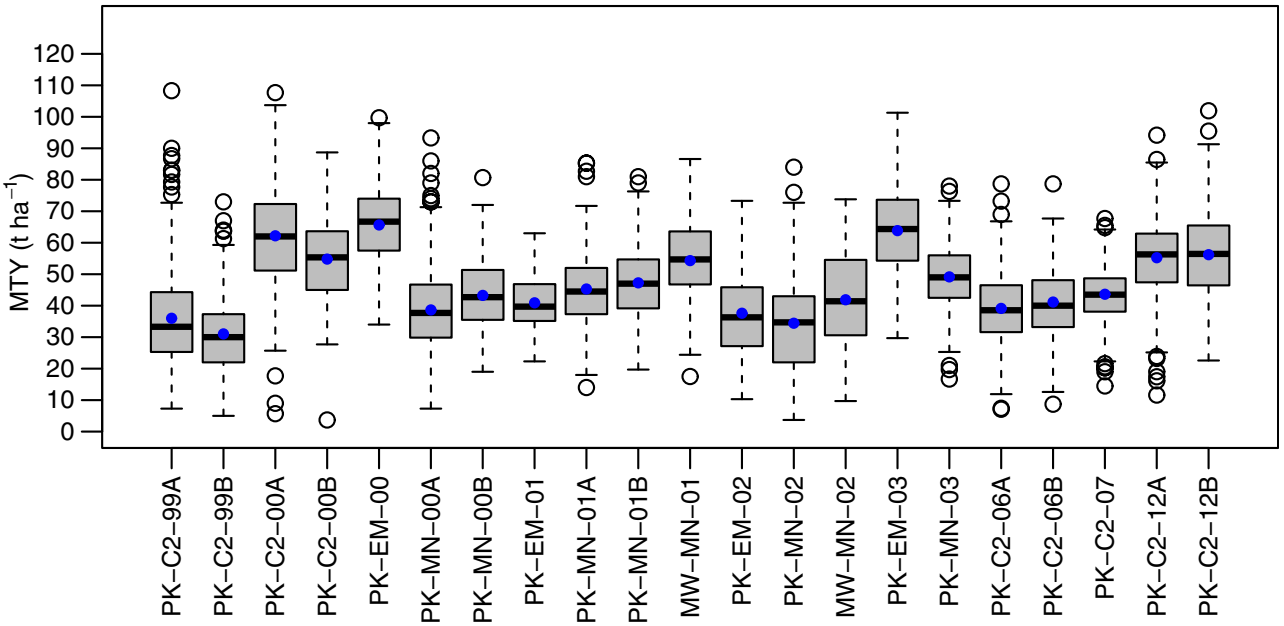

Fig. 1: Box plots of marketable tuber yield (MTY) (t/ha) from 21 early-stage potato breeding trials over 8 years. PK and MW trial prefixes refer to Pukekohe and Manawatu locations, respectively. Mean yields are indicated by the filled circles

Table 2: Trial summary and REML estimates of variance components from models for marketable tuber yield (MTY, t/ha); fixed and random effects and autoregressive (AR1) parameters, Spearman's rank correlations (*rho*) and % concurrence of the top 10% of genotypes (ranked on empirical breeding values (EBVs) of MTY) between *base* and *extended* models

					Error model				Spatial correlation					
Trial code	Dimension	Trial mean	Fixed regressions	σ_a^2	σ_b^2	σ_r^2	σ_c^2	spl_{row}	ρ_r	ρ_c	σ_e^2	*rho*	% concurrence	V_A/V_P
PK-C2-99A	40 × 16	36	l_{row}	201.0		9.1	12.0				34.7	0.95	81	0.78
PK-C2-99B	38 × 8	31	l_{row} l_{col}	121.3			2.9			0.25	43.7	0.93	73	0.72
PK-C2-00A	60 × 8	62	l_{row}	241.9		4.5					66.1	0.98	92	0.77
PK-C2-00B	14 × 8	55		280.8		2.2	2.5		−0.41		44.4	0.99	83	0.85
PK-EM-00	33 × 8	66	l_{col}	108.5	20.9	5.4	1.9		−0.20		37.8	0.97	67	0.62
PK-MN-00A	20 × 24	39	l_{row}	172.4	25.9	0.9			0.12	0.10	59.2	0.92	79	0.67
PK-MN-00B	14 × 8	43		143.3		5.2	2.3				44.4	‡	‡	0.73
PK-EM-01	24 × 8	41	l_{row}	34.2		0.8	1.8			−0.26	23.2	0.97	83	0.57
PK-MN-01A	12 × 8	45	l_{row} l_{col}	182.9							46.5	0.98	80	0.80
PK-MN-01B	34 × 8	47	pl_{row}	94.8		1.3	2.6		−0.14	−0.21	31.5	0.97	77	0.73
MW-MN-01	12 × 8	54		150.5		0.9	13.4				58.8	‡	‡	0.67
PK-EM-02	24 × 12	38	l_{row}	181.5		2.3	1.4				37.2	0.99	100	0.82
PK-MN-02	14 × 8	34	l_{row}	171.7					−0.30		65.5	0.98	67	0.72
MW-MN-02	12 × 10	42	l_{row}	151.9		11.7	8.3	5.2			26.4	0.99	100	0.75
PK-EM-03	30 × 8	64	l_{row} l_{col}	160.6		1.6	1.4		−0.21	−0.31	53.1	0.98	88	0.74
PK-MN-03	18 × 8	49		94.4			12.1		−0.21		39.4	0.99	80	0.65
PK-C2-06A	54 × 10	39	l_{row}	109.2		0.5	14.3		−0.30		50.3	0.95	85	0.63
PK-C2-06B	26 × 10	41	l_{col}	122.2		6.3	3.1		−0.27	−0.29	34.1	0.98	83	0.74
PK-C2-07	34 × 20	44	l_{row}	58.0		5.2	2.9				21.3	0.99	88	0.66
PK-C2-12A	56 × 7	55	l_{col}	149.9		4.9	4.4		−0.37	−0.14	48.2	0.98	95	0.72
PK-C2-12B	54 × 7	56		154.8		0.7	2.2				87.5	‡	‡	0.63

Trial mean is the observed mean tuber yield (t/ha), l_{row} and l_{col} represent a linear regression of the marketable yield on column or row number, respectively; pl_{row} represents a polynomial regression (of order 2); σ_a^2 is the additive genetic variance; σ_b^2, σ_r^2, σ_c^2 and spl_{row} are the replicate, row and column variances and random row splines, respectively, and represent the error model; ρ_r and ρ_c are the spatial correlation parameters; σ_e^2 is the residual error variance; V_A/V_P is the proportion of additive genetic variance to the phenotypic variance – a measure of the narrow-sense heritability. ‡indicates that the *base* model (no spatial error component or trend term) was the best fitting model.

negative for TTY and also for MTY with the exception of two trials (PK-C2-99B and PK-MN-00A). For PTY, there were approximately equal numbers of positive and negative spatial correlations. In exploratory analyses of TTY and MTY, examination of a null model (*genotype* + *spatial*) sometimes yielded positive AR1 spatial correlation estimates but these were often effectively reduced to zero when random row and column terms were added. For example, for MTY in trial PK-C2-99A, positive AR1 estimates for rows and columns of 0.46 and 0.42, respectively, were reduced to 0.11 and 0 when random row and column terms were added. Spatial effects, when fitted, often had an impact on the concurrence between *base* and *extended* models

Table 3: Summary of genetic variance models (G_0), number of genetic and total variance parameters (N_p), Log-Likelihood (Log-L) and Akaike information criterion (AIC) goodness-of-fit for total tuber yield (TTY), marketable tuber yield (MTY) and the fraction of marketable yield (PTY)

			TTY				MTY				PTY			
	N_p		Log-L	AIC	Log-L	AIC	Log-L	AIC	Log-L	AIC	Log-L	AIC	Log-L	AIC
G_0 structure	G_0	total	$\mathbf{I}\sigma_g^2$		$\mathbf{A}\sigma_a^2$		$\mathbf{I}\sigma_g^2$		$\mathbf{A}\sigma_a^2$		$\mathbf{I}\sigma_g^2$		$\mathbf{A}\sigma_a^2$	
SIMPLE	1	54	−10535	405	−10488	311	−10266	345	−10195	283	703	98	692	121
DIAG	15	68	−10499	362	−10411	246	−10189	298	−10159	239	710	113	699	135
CORH	16	69	−10375	115	−10317	0	−10063	49	−10041	4	767	0	764	6
FA1	30	83	−10366	126	−10307	8	−10056	81	−10027	4	778	7	773	16
FA2	44	97	−10349	121	−10289	2	−10051	82	−10010	0	785	22	779	34

AIC expressed as the difference from the best fitting model. $\mathbf{I}\sigma_g^2$ represents the independent genotypic variance (no pedigree fitted) and $\mathbf{A}\sigma_a^2$ represents the pedigree-based genotypic variance (pedigree included).

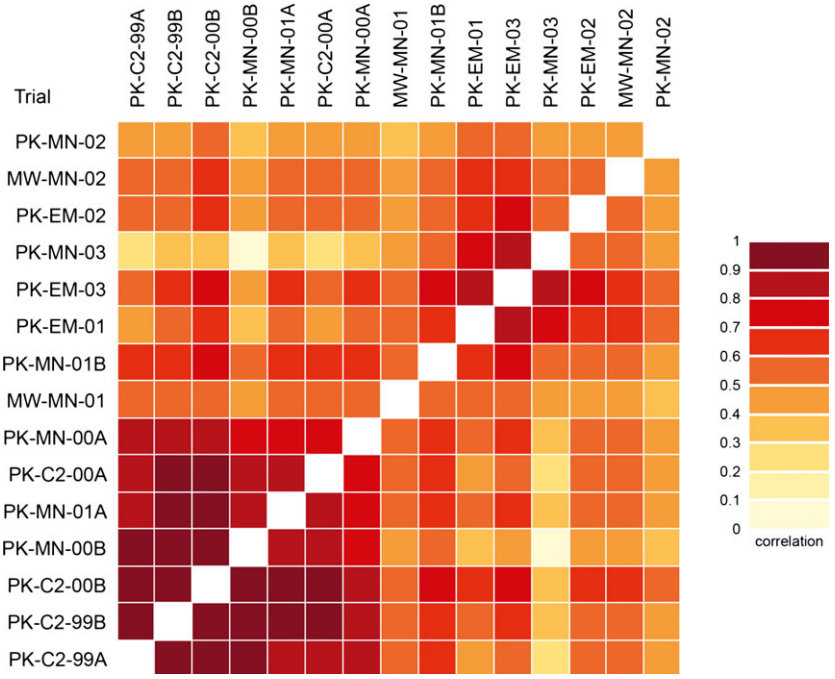

Fig. 2: Genetic (additive) correlation estimates of MTY (marketable yield) between early-stage potato selection trials from a factor analytic structure of order 2 (FA2)

for the top 10 per cent of genotypes selected, but agreement was usually greater than 80%.

MET analysis and variance models

A summary of genetic variance models tested is presented in Table 3. Based on AIC, there was a large improvement of CORH and FA*k* variance structures over DIAG for all three traits. For TTY and PTY, CORH was a small improvement over both FA1 and FA2, with and without the pedigree fitted. The trial-to-trial genetic correlation estimate from the CORH model was 0.69 and 0.72 for TTY and PTY, respectively. For MTY (with pedigree included), there was no difference in model fit between FA1 and CORH, and for FA2, there was very little improvement over both FA1 and CORH. For MTY, the trial-to-trial genetic correlation estimate from the CORH model was 0.69. A heatmap plot of REML estimates of the additive genetic correlations from the FA model for MTY is shown in Fig. 2. There was a pattern of decreasing genetic correlations over time, but there were no negative genetic correlations estimated between any trials. For MTY, the lowest genetic correlations were found between PK-MN-03 and the 1999–2000 trials. These ranged from 0.06 to 0.39. For MTY, fitting a FA1 model accounted for 71% of the variance, which increased to 77% with a FA2 fit. There were problems with convergence for the FA3 models, as there were with attempts to fit an unstructured (US) model to the data. Including a relationship matrix (pedigree) in analyses improved model fit for TTY and MTY but not for PTY (Table 3).

Empirical breeding values for tested genotypes predicted from CORH, FA1 and FA2 were all highly correlated with each other, with product-moment correlation coefficients between 0.98 and 0.99. There was a 93% concurrence of the top-ranked 10% of selected genotypes between FA2 and FA1 and 91% between FA2 and CORH. Figure 3 illustrates a shrinkage effect of MTY when data were fitted to a FA2 model, with plots of EBVs slightly departing from the line of unity between: (a) CORH and FA2, and (b) FA1 and FA2. Simulations of the response to selection over all levels of p (the proportion of the top-ranked genotypes selected) were similar for CORH and FA1 but were reduced for FA2 (Fig. 4), which also reflected the shrinkage of empirical breeding values.

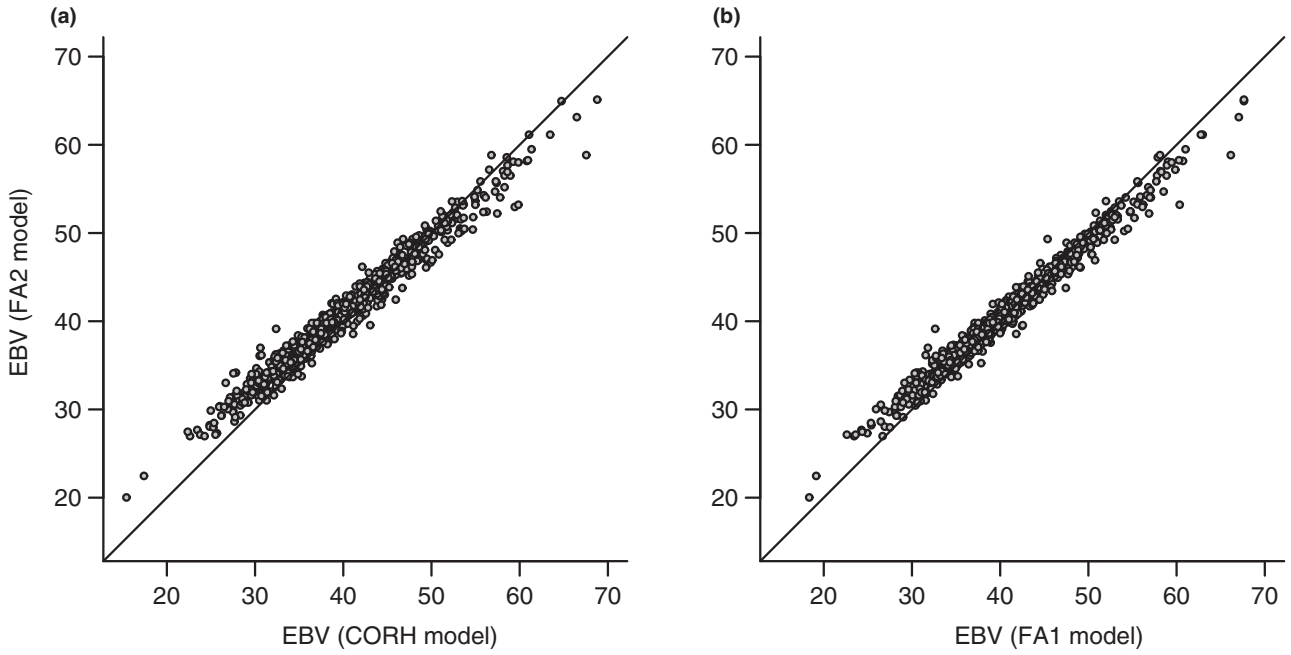

Fig. 3: Scatter plots of empirical breeding values (EBV) for marketable tuber yield (MTY) for: (a) EBVs predicted from heterogeneous variance–homogeneous correlation model (CORH) and factor analytic structure FA2 models, and (b) EBVs predicted from FA1 and FA2 models

Discussion

Spatial parameters and interplot competition

Local spatial trends were not a consistent feature of the potato trials tested; extending models to include spatial effects did not always improve model fit. By contrast, a number of studies in other field crops, particularly cereals, have demonstrated considerable advantages when including spatial terms (e.g. Gilmour et al. 1997, Qiao et al. 2000) and are routinely included in analyses (e.g. Oakey et al. 2006, Beeck et al. 2010). Spatial analysis has been considered as an alternative model to the traditional analysis of complete or incomplete block designs, but based on the comprehensive re-analysis of 53 lentil variety trials, Sarker et al. (2001) recommended that block design methods could often be enhanced but not replaced with spatial methods. Müller et al. (2010) also emphasised that prudence was probably the best approach and advised that over-complication should be avoided if possible. They found that a standard block model outperformed a spatial model in most cases when analysing 293 sugar beet and 64 barley trials. In the present study, the blocking designed for trials often appeared to deal adequately with localised heterogeneity. However, the effort expended in checking for spatial effects is small, compared with the effort and costs involved in setting up and managing field trials. Spatial modelling should therefore be a consideration in potato evaluation to account for possible field heterogeneity that may be caused by localised factors within a trial site, such as soil chemical and physical properties (e.g. Redulla et al. 2002, Po et al. 2010), and which may be trait dependent (e.g. Dutkowski et al. 2006).

In trials, where spatial effects did appear to be important, spatial correlation estimates were often small and mostly negative (for TTY and MTY). Similarly, small negative spatial correlations were also found by Stringer and Cullis (2002) in sugarcane trials and were attributed to interplot (or intergenotype) competition. The use of larger interplot distances has been suggested as a means to eliminate intergenotype competition in breeders' trials, but this implies a reduction in selection intensity when the total trial area is fixed. Furthermore, bias due to competitive ability may be replaced by bias due to the occurrence of a genotype × plant density interaction, whereby the performance of genotypes is density dependent (Bos and Caligari 2008). Work by Connolly et al. (1993) identified competitive effects of yield in single-row plots of potatoes, although these effects were not ubiquitous over all trials tested. They found little reranking of genotypes but there was shrinkage in the range of yield estimates from high- and low-yielding plots and closer agreement with pure-stand yields after accounting for competitive effects.

In the presence of interplot competition, the AR1 model, as a 'power' model, implies a negative correlation between a plot and its immediate neighbour as well as its third, fifth, etc.

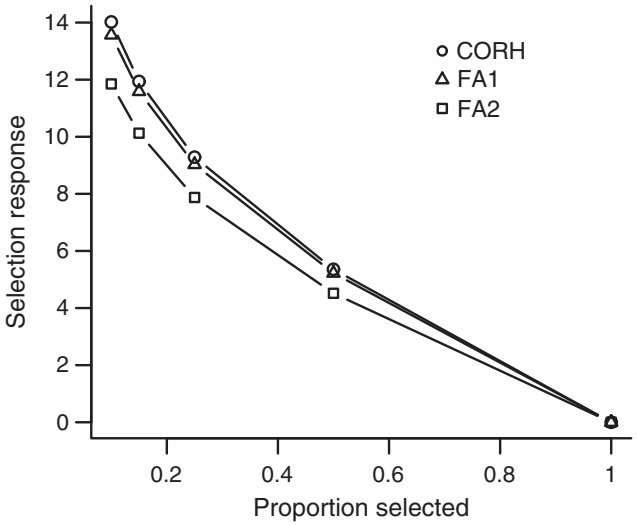

Fig. 4: Simulated response to selection for marketable tuber yield (MTY, t/ha) for variance models heterogeneous variance–homogeneous correlation model (CORH), factor analytic structure FA1 and FA2

neighbours. Fitting a negative correlation beyond the immediate neighbour (along with positive correlations for the second, fourth, etc. neighbours) does not seem a sensible model and appears to have no biological justification. The weak negative correlations found in the present study suggested that the correlations beyond immediate neighbours were negligible and could be reasonably overlooked. However, when interplot competition is found to be strong, the fitting of an AR1 model may not be appropriate and suggests that competition effects should be fitted explicitly.

Genetic variance models

A heterogeneous variance and homogeneous correlation structure (CORH) was found to be adequate for modelling the G×E effects for TTY and PTY in the early-stage potato trials tested. In a previous study, a simple model was also found to be suitable for the evaluation of resistance to tuber powdery scab disease from 12 years of potato field trials (Paget et al. 2014). For MTY, there was only a small difference between the AIC of the factor analytic model of order 2 (FA2) and both the FA1 and CORH models. This suggested that the three models performed equally well for the analysis of MTY. EBV plots between CORH and FA2, and FA1 and FA2 (Fig. 3) indicated that the FA2 model was possibly over fitting the data, and some degree of bias was introduced into the process when moving from a FA1 to a FA2 model. This was also reflected in the simulations of the responses to selection in which both CORH and FA1 were similar over all levels of p. In comparison, responses were reduced when selections on EBVs were based on simulations from FA2 (Fig. 4). These results also suggested that environments, which were mainly temporal in the current study, that is different years or growing seasons in the same location, were relatively homogeneous for these data. The results are, of course, presented in the context of early-stage potato selection trials when the extent of testing over multiple locations is restricted because of limited planting material. From model comparisons, previous studies have found that FA variance structures were suitable for both early- and late-stage evaluation trials (Smith et al. 2001, Crossa et al. 2006, Kelly et al. 2007, Burgueño et al. 2011) in other crops. These studies were generally based on MET information from more extensive trial data and, most probably, more diverse environments, such as the extensive international wheat trials of CIMMYT (Crossa et al. 2006). From cross-validation studies, Burgueño et al. (2011) found that when G×E was not complex, that is a rescaling of performance rather than a reranking of genotypes, both FA and simple non-FA models gave good predictive ability. So and Edwards (2011) found that because of poor genetic links of maize hybrids across environments, modelling heterogeneous genotype covariances did not improve predictions. Where access to suitable software and experience is limited, evaluation by fitting a homogeneous correlation structure may provide a more approachable method to fit a genetic variance structure to MET data if appropriate.

The present study ignored the fitting of genotype main effects, nesting genotypes within environments and obtaining the BLUPs of breeding values as the arithmetic average across environments of the predictions for each environment. This was considered to be appropriate in the current context as genotypes (with limited plant material available in the early-selection stages) were generally tested over multiple years rather than locations and that the correlations between trials were positive and generally high. Further, when genotype main effects were included in models, there was little difference to the predictions, that is no reranking of genotypes compared with those obtained when excluded (results not shown). However, it is likely that the standard errors of the predictions will be underestimated within a restricted inference space. It is therefore emphasised that under different circumstances, the inclusion of genotype effects may be more appropriate to broaden inference beyond the test environments. When genotype main effects and G×E interaction effects are not separated, it is difficult to monitor the behaviour of particular genotypes in particular environments. Separation of these effects is therefore useful for studying the specific adaptation of genotypes to environments (Crossa et al. 2006).

Use of ancestry information

For pedigree-based BLUP of breeding values in plants and animals, it is recommended that all data that have been used in selection decisions should be included in the evaluation for the estimate of breeding values (Piepho and Möhring 2006). An assumption of the present analysis was that selection for yield had been absent in the initial generations (seedling and C1). This seemed reasonable, as selection for yield in the seedling or C1 generations is generally avoided. Previous work has considered the low efficiency of potato selection in the early generations, showing a poor association between selection for performance as seedlings and performance in the C1 and later generations (e.g. Anderson and Howard 1981, Brown et al. 1984, Brown and Caligari 1986). PTY, however, was highly correlated with a general impression score, which is a categorical preference score of tubers given on a 1–9 scale by breeders, in the present study (results not shown). General impression is a trait for which there has been selection in the initial generations but the basis on which these decisions were made was not recorded. Using a pedigree-based genotypic variance,V_G (i.e. $V_G = \mathbf{A}\sigma_A^2$), as in the present study, Piepho and Möhring (2007) reported a better model fit in some analyses when assuming independence between genotypes (i.e.$V_G = \mathbf{I}\sigma_G^2$). It was suggested that selection has possibly taken place, and the information on which selection had been based was not included in the analysis.

The additive relationship matrix **A** was based on disomic inheritance. Under the assumptions of no past selection, double reduction or inbreeding, the expected additive genetic covariances both of diploid and tetraploid relatives are equivalent (Lynch and Walsh 1998), which may or may not be appropriate when dealing with autotetraploid potato. From our unpublished genetic analysis of potato (tuber) starch data, there was very little difference in the BLUPs when a diploid relationship matrix was replaced with a tetrasomic-based relationship matrix that was derived using the method of Kerr et al. (2012). This used a double reduction coefficient of 0.1 which seemed reasonable, based on the estimates of Haynes and Douches (1993) and Slater et al. (2014). Similar results were also reported by Slater et al. (2014). Fitting a non-additive component may reduce bias in breeding value estimation, and variance estimates could be exploited by selecting favourable parental combinations (Mrode 2005, Oakey et al. 2007). It may also be more appropriate for clonal selection in cultivar development, that is selection of individuals with a high total genetic value – a high 'potential production ability' or 'production worth', which demands further investigation in potato.

In conclusion, the use of historic field data provides an opportunity to explore statistical models that improve the methods and precision of identifying new high-yielding genotypes for use as

parents, as well as potential and worthy cultivars. In the analyses of potato field trials, spatial effects were not important in all years but there was evidence of interplot competition in some years. For the genetic evaluation of potato yield, a simple (homogeneous) correlation structure to model G×E effects (allowing for heterogeneity of trial genetic variance) was suitable for the series of early-stage MET trials tested. Simple models are easier to fit than unstructured or factor analytic models, particularly when a pedigree is included, and therefore offer advantages for routine genetic evaluation of potato.

Acknowledgements

We gratefully acknowledge the funding of this study by the Potatoes NZ Charitable Trust. We thank Fred Braam and Moe Jeram for management of the potato field trials. We also thank Alasdair Noble of AgResearch NZ Ltd, Steve Lewthwaite of The New Zealand Institute for Plant and Food Research Ltd and the anonymous reviewers for their comments and helpful suggestions.

References

Anderson, J.A.D and H.W. Howard 1981: Effectiveness of selection in the early stages of potato breeding programmes. Potato Res. **24** (3), 289—299.

Beeck, C. P., W. A. Cowling, A. B. Smith, and B. R. Cullis, 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.

van Berloo, R., R. C. B. Hutten, H. J. van Eck, and R. G. F. Visser, 2007: An online potato pedigree database resource. Potato Res. **50**, 45—57.

Bos, I., and P. D. S. Caligari, 2008: Selection Methods in Plant Breeding, 2nd edn. Springer, Dordrecht.

Brown, J., and P. D. S. Caligari, 1986: The efficiency of seedling selection for yield and yield components in a potato breeding programme. Zeitschrift fuer Pflanzenzuechtung **96**, 53—62.

Brown, J., P. D. S. Caligari, G. R. Mackay, and G. E. L. Swan, 1984: The efficiency of seedling selection by visual preference in a potato breeding programme. J. Agric. Sci. **103**, 339—346.

Burgueño, J., J. Crossa, J. M. Cotes, F. San Vicente, and B. Das, 2011: Prediction assessment of linear mixed models for multi-environment trials. Crop Sci. **51**, 944—954.

Butler, D. G., B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2009: ASReml-R Reference Manual. Version 3, Queensland Department of Primary Industries & Fisheries, Australia.

Connolly, T., I. Currie, J. Bradshaw, and J. McNicol, 1993: Inter-plot competition in yield trials of potatoes *(Solanum tuberosum* L.) with single-drill plots. Ann. Appl. Biol. **123**, 367—377.

Crossa, J., P. N. Fox, W. H. Pfeiffer, S. Rajaram, and H. G. J. Gauch, 1991: AMMI adjustment for statistical analysis of an international wheat yield trial. Theor. Appl. Genet. **81**, 27—37.

Crossa, J., J. Burgueño, P. L. Cornelius, G. McLaren, R. Trethowan, and A. Krishnamachari, 2006: Modeling genotype x environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. Crop Sci. **46**, 1722—1733.

CycSoftware, 2009: CycDesigN 4.0 A Package for the Computer Generation of Experimental Designs. Version 4.0, CycSoftware Ltd, Hamilton, New Zealand.

Dutkowski, G. W., J. C. E. Silva, A. R. Gilmour, H. Wellendorf, and A. Aguiar, 2006: Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. Can. J. For. Res. **36**, 1851—1870.

Edmondson, R. N., 2004: Past developments and future opportunities in the design and analysis of crop experiments. J. Agric. Sci. **143**, 27—33.

Gauch, H. G., and R. W. Zobel, 1988: Predictive and postdictive success of statistical analyses of field trials. Theor. Appl. Genet. **76**, 1—10.

Gilmour, A. R., B. R. Cullis, A. P. Verbyla, and A. C. Gleeson, 1997: Accounting for natural and extraneous variation in the analysis of field experiments. J. Agric. Biol. Environ. Stat **2**, 269—293.

Gleeson, A. C., 1997: Spatial analysis, In: R. A. Kempton, and P. N. Fox, (eds), Statistical Methods for Plant Variety Evaluation, 69—85. Chapman & Hall, London.

Haynes, K. G., and D. S. Douches, 1993: Estimation of the coefficient of double reduction in the cultivated tetraploid potato Theor. Appl. Genet. **85**, 857—862.

Kelly, A. M., A. B. Smith, J. A. Eccleston, and B. R. Cullis, 2007: The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. Crop Sci. **47**, 1063—1070.

Kelly, A. M., B. R. Cullis, A. R. Gilmour, J. A. Eccleston, and R. Thompson, 2009: Estimation in a multiplicative mixed model involving a genetic relationship matrix. Genet. Sel. Evol. **41**:33 doi:10.1186/1297-9686-41-33.

Kerr, R. J., L. Li, B. Tier, G. W. Dutkowski, and T. A. McRae, 2012: Use of the numerator relationship matrix in genetic analysis of autopolyploid species. Theor. Appl. Genet. **124**, 1271—1282.

Lynch, M., and B. Walsh, 1998: Genetics and Analysis of Quantitative Traits. Sinauer Associates Inc., Sunderland.

Meyer, K., 2009: Factor-analytic models for genotype x environment type problems and structured covariance matrices. Genet. Sel. Evol. **41**:21 doi:10.1186/1297-9686-41-21.

Mrode, R. A., 2005: Linear Models for the Prediction of Animal Breeding Values, 2nd edn. CABI Publishing, Wallingford.

Müller, B. U., K. Kleinknecht, J. Möhring, and H. P. Piepho, 2010: Comparison of spatial models for sugar beet and barley trials. Crop Sci. **50**, 794—802.

Oakey, H., A. Verbyla, W. Pitchford, B. Cullis, and H. Kuchel, 2006: Joint modeling of additive and non-additive genetic line effects in single field trials. Theor. Appl. Genet. **113**, 809—819.

Oakey, H., A. P. Verbyla, B. R. Cullis, X. M. Wei, and W. S. Pitchford, 2007: Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. Theor. Appl. Genet. **114**, 1319—1332.

Paget, M. F., P. A. Alspach, R. A. Genet, and L. A. Apiolaza, 2014: Genetic variance models for the evaluation of resistance to powdery scab (*Spongospora subterranea* f. sp. *subterranea*) from long-term potato breeding trials. Euphytica **197**, 369—385.

Piepho, H. P., 1997: Analyzing genotype-environment data by mixed models with multiplicative terms. Biometrics **53**, 761—767.

Piepho, H. P., 1998: Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Theor. Appl. Genet. **97**, 195—201.

Piepho, H. P., and J. Möhring, 2006: Selection in cultivar trials – is it ignorable? Crop Sci. **146**, 193—202.

Piepho, H. P., and J. Möhring, 2007: Computing heritability and selection response from unbalanced plant breeding trials. Genetics **177**, 1881—1888.

Po, E. A., S. S. Snapp, and A. Kravchenko, 2010: Potato yield variability across the landscape. Agron. J. **102**, 885—894.

Qiao, C. G., K. E. Basford, I. H. DeLacy, and M. Cooper, 2000: Evaluation of experimental designs and spatial analyses in wheat breeding trials. Theor. Appl. Genet. **100**, 9—16.

R Development Core Team, 2012: R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing. Austria, Vienna.

Redulla, C. A., J. R. Davenport, R. G. Evans, M. J. Hattendorf, A. K. Alva, and R. A. Boydston, 2002: Relating potato yield and quality to field scale variability in soil characteristics. Am. J. Potato Res. **79**, 317—323.

Sarker, A., M. Singh, and W. Erskine, 2001: Efficiency of spatial methods in yield trials in lentil (*Lens culinaris* ssp *culinaris*). J. Agric. Sci. **137**, 427—438.

Slater, A. T., G. M. Wilson, N. O. I. Cogan, J. W. Forster, and B. J. Hayes, 2014: Improving the analysis of low heritability complex traits for enhanced genetic gain in potato. Theor. Appl. Genet. **127**, 809—820.

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

Smith, A. B., B. R. Cullis, and R. Thompson, 2005: The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. J. Agric. Sci. **143**, 449—462.

So, Y. S., and J. Edwards, 2011: Predictive ability assessment of linear mixed models in multi-environment trials in corn. Crop Sci. **51**, 542—552.

Stringer, J. K., and B. R. Cullis, 2002: Application of spatial analysis techniques to adjust for fertility trends and identify interplot competition in early stage sugarcane selection trials. Aust. J. Agric. Res. **53**, 911—918.

Verbyla, A. P., B. R. Cullis, M. G. Kenward, and S. J. Welham, 1999: The analysis of designed experiments and longitudinal data by using smoothing splines. J. R. Stat. Soc. Ser. C-Appl. Stat. **48**, 269—300.