Theor Appl Genet (2004) 108:1256–1264
DOI 10.1007/s00122-003-1541-4

ORIGINAL PAPER

C. G. Qiao · K. E. Basford · I. H. DeLacy · M. Cooper

Advantage of single-trial models for response to selection in wheat breeding multi-environment trials

Received: 27 June 2003 / Accepted: 14 November 2003 / Published online: 20 December 2003

Abstract An investigation was conducted to evaluate the impact of experimental designs and spatial analyses (single-trial models) of the response to selection for grain yield in the northern grains region of Australia (Queensland and northern New South Wales). Two sets of multi-environment experiments were considered. One set, based on 33 trials conducted from 1994 to 1996, was used to represent the testing system of the wheat breeding program and is referred to as the multi-environment trial (MET). The second set, based on 47 trials conducted from 1986 to 1993, sampled a more diverse set of years and management regimes and was used to represent the target population of environments (TPE). There were 18 genotypes in common between the MET and TPE sets of trials. From indirect selection theory, the phenotypic correlation coefficient between the MET and TPE single-trial adjusted genotype means [$r_{p(MT)}$] was used to determine the effect of the single-trial model on the expected indirect response to selection for grain yield in the TPE based on selection in the MET. Five single-trial models were considered: randomised complete block (RCB), incomplete block (IB), spatial analysis (SS), spatial analysis with a measurement error (SSM) and a combination of spatial analysis and experimental design information to identify the preferred (PF) model. Bootstrap-resampling methodology was used to construct multiple MET data sets, ranging in size from 2 to 20 environments per MET sample. The size and environmental composition of the MET and the single-trial model influenced the $r_{p(MT)}$. On average, the PF model resulted in a higher $r_{p(MT)}$ than the IB, SS and SSM models, which were in turn superior to the RCB model for MET sizes based on fewer than ten environments. For METs based on ten or more environments, the $r_{p(MT)}$ was similar for all single-trial models.

Communicated by H.C. Becker

C. G. Qiao · K. E. Basford (✉) · I. H. DeLacy · M. Cooper
School of Land and Food Sciences,
The University of Queensland,
4072 Brisbane, Queensland, Australia
e-mail: k.e.basford@uq.edu.au

Present address:
C. G. Qiao, Statistics New Zealand,
P.O. Box 2722, Wellington, New Zealand

Present address:
M. Cooper, Pioneer Hi-Bred International Inc.,
7300 N.W. 62nd Ave., P.O. Box 1004, Johnston, IA 50131, USA

Introduction

An investigation was conducted to evaluate the impact of experimental designs and spatial analyses (referred to as single-trial models) on the response to selection for grain yield of wheat. Multi-environment trials (METs) are conducted within plant breeding programs to evaluate test genotypes and predict their expected responses to selection in a target population of environments (TPE) (Comstock 1977). The presence of genotype-by-environment (G×E) interactions and experimental error complicates the selection of genotypes with superior performance in the TPE (Comstock and Moll 1963; Basford and Cooper 1998). Systematic heterogeneity among the experimental units within replicates and/or appreciable spatial correlation between neighbouring units has been widely recognised as a problem for the commonly used randomised complete block design and analysis model for such trials (Zimmerman and Harville 1991). Advanced experimental designs (Yates 1939; Cochran and Cox 1957; Williams 1986; John and Whitaker 1993; Williams and Matheson 1994) and spatial analysis methodologies (Gleeson and Cullis 1987; Cullis and Gleeson 1989, 1991; Grondona et al. 1996; Gilmour et al. 1997) have been developed to cope with these problems, and have been demonstrated to improve statistical efficiency or precision for genotype comparisons. Qiao et al. (2000) applied both advanced experimental designs and spatial analyses in the analysis of the single trials in a wheat MET in the northern grains region of Australia. They found that these alternative single-trial models improved trial efficiency and also affected the

selection of superior genotypes in comparison to the randomised complete block (RCB) design and analysis. Since the overall objective of a breeding program is to maximise realised genetic gain for the traits of interest within a TPE (Comstock 1977; Magnussen 1993), the benefits from these analytical methods can be assessed in terms of their impact on indirect response to selection for grain yield from the METs to the TPE, using an independent set of experiments representative of the TPE for the northern grains region of Australia. While there has been considerable investigation of the impact of single-trial analysis models at the level of efficiency of individual trials, there is limited information on the effects of these experimental designs and spatial analysis methods on the selection of superior genotypes, as judged by their realised superior performance in independent experiments (Kirk et al. 1980; Matassa et al. 1998).

Falconer (1952, 1989) extended correlated genetic advance theory to study G×E interactions. Cooper and DeLacy (1994) used this framework to study G×E interactions for grain yield of wheat in the northern grains region. Within this framework, METs can be considered as samples of environments from the TPE. MET samples of environments that deviate from expectations of the TPE are likely to result in suboptimal selection response in the TPE (Cooper and Podlich 1999). Response in the TPE can be considered as a case of indirect response to selection $\Delta G_{\mathrm{T|M}}$, where response in the TPE is evaluated as an indirect response to selection on the results of the MET. This relationship can be expressed in the form of a prediction equation:

$$\Delta G_{\mathrm{T|M}} = i_{\mathrm{M}} h_{\mathrm{M}} h_{\mathrm{T}} r_{\mathrm{g(MT)}} \sigma_{\mathrm{p(T)}}, \quad (1)$$

where i_{M} is the standardised selection differential applied in the MET; h_{M} and h_{T} are the square roots of the heritabilities in the MET and the TPE, respectively; $r_{\mathrm{g(MT)}}$ is the genetic correlation between genotype mean performance in the MET and the TPE; and $\sigma_{\mathrm{p(T)}}$ is the phenotypic standard deviation of the selection units in the TPE. DeLacy et al. (1996) expressed this indirect selection response equation between the two sets of environments in a simplified form:

$$\Delta G_{\mathrm{T|M}} = i_{\mathrm{M}} r_{\mathrm{p(MT)}} \sigma_{\mathrm{p(T)}}, \quad (2)$$

where it is assumed that there is no environmental or error correlation among environments and $r_{\mathrm{p(MT)}}$ is the phenotypic correlation between genotype performance in the MET and TPE. Equations 1 and 2 can be used to examine expected response to selection in the TPE, as it is influenced by adjustments to selection based on different treatments of data from a MET. For Eq. 2, $r_{\mathrm{p(MT)}}$ dominates the indirect response to selection, particularly when it is reasonable to assume that the selection pressure applied in the MET and the phenotypic variation in the TPE are constant. Improvements in predicted indirect response to selection in the TPE should be detectable in terms of improvements in $r_{\mathrm{p(MT)}}$.

In this paper, we use $r_{\mathrm{p(MT)}}$ to measure changes in predicted indirect response to selection in the TPE based on adjustment to genotype mean data from analysis of the MET. Here, we focus on the impact of different single-trial models in combination with the number of environments used in a MET. The aims of this study were to examine: (1) the expected effect of single-trial models on prediction of indirect response to selection in the TPE, based on adjusted genotype mean performance measured in a MET; and (2) interactions between the effects of single-trial models and both number of environments and combinations of environments sampled in a MET.

Materials and methods

Definition of environment types

Eighty wheat breeding trials conducted over a period spanning 11 years, 1986–1996, with a common set of 18 genotypes (advanced lines and cultivars), were considered here (Appendix). These trials were broken into two independent sets for the purposes of this study using a combination of experimental design and practical considerations (e.g. similar to the approach used by Lin and Morrison 1992). The set of 33 trials covering the period 1994–1996 was used to represent the selection set of environments sampled by the breeding program and is referred to as the MET. The other set of 47 trials conducted during the period 1986–1993 was considered to provide a measure of genotype performance in the target population of environments and is referred to as the TPE. The 33 MET trials all involved advanced designs with three replicates, and this set of trials is regarded as representative of the MET selection regime that is currently used in the wheat pedigree-breeding programs that target the northern grains region of Australia. The TPE trials, on the other hand, were diverse in design and layout, had different levels of replication, and most were not designed as incomplete block experimental designs. They covered a larger number of years than the MET and also a wider range of management and abiotic stress regimes and were conducted to be representative of the range of target production environments in the Australian northern grains region.

For the 33 wheat trials representing the MET, the total number of entries was 72. Triple-lattice designs were used for all trials, with nine laid out in a one-column-by-many-rows array (called 'long-face trials') and the rest laid out in a row-column array (called 'row-column trials') (Appendix). For the TPE trials, the number of entries and experimental layout were variable. The 18 genotypes used in this study were part of the full set of lines included in each trial.

Analytical methods

Indirect response to selection from the MET to the TPE

Two sets of trials were used as the basis for evaluating the effects of the MET design and analysis options on response to selection in the TPE, using a generalisation of Falconer's (1952, 1989) correlated genetic advance theoretical framework. The impact of experimental designs and spatial analysis methods on components of the prediction equation for correlated response to selection was examined within the MET and TPE. From Eq. 2, the correlation between average genotype performance in the MET and TPE [$r_{\mathrm{p(MT)}}$], both adjusted for the results of the defined single-trial model, was used as a measure of the predicted correlated genetic advance in the TPE in response to the single-trial analytical methodologies. An increase in the MET-TPE correlation relative to the basic randomised complete block model was taken as evidence of an improvement in the predicted correlated response to selection in the TPE.

Following the same procedures as in Qiao et al. (2000), five single-trial models (analysis of variance models for the single trials) were used (wherever possible) to analyse the grain yield (*t/ha*) data for each of the 33 trials in the MET:

1. Randomised complete block (RCB)

$$y_{ij} = \mu + g_i + r_j + \varepsilon_{ij}$$

2. Incomplete block (IB)

$$y_{ijk} = \mu + g_i + r_j + (b|r)_{jk} + \varepsilon_{ijk}$$

3. Standard spatial (SS)

$$y_{ikl} = \mu + g_i + (AR1 \times AR1)_{kl} + \varepsilon_{ikl}$$

4. Standard spatial fitted with measurement error (SSM)

$$y_{ikl} = \mu + g_i + (AR1 \times AR1)_{kl} + \eta_{ikl} + \varepsilon_{ikl}$$

5. Combination of spatial analysis and experimental design information to identify the preferred (PF) model (Appendix)

For each of the 47 trials in the TPE, the PF model was applied for estimation of the adjusted genotypic means. In the above models, μ is the grand mean, g_i is the effect of the ith genotype, r_j is the effect of the jth replicate, and the error terms (ε_{ij}, ε_{ijk} and ε_{jkl}) are assumed to be independent and normally distributed ~N(0, σ^2_ε). In the RCB model, y_{ij} is the observation of the ith line in the jth replicate (block), and ε_{ij} is the corresponding error effect. In the IB model, y_{ijk} is the observation of the ith line in the kth incomplete block within the jth replicate, $(b|r)_{jk}$ is the effect of the kth incomplete block within the jth replicate, and ε_{ijk} is the corresponding error effect. In the two spatial models, SS and SSM, y_{ikl} is the observation of the ith line at the kth row and the lth column, $(AR1 \times AR1)_{kl}$ is the separable, first-order autoregressive process along the kth row and the lth column and is the key term in the two-dimensional spatial model [for the long-face trials, this term simplifies to $(AR1)_l$, the first-order autoregressive process along the lth column and is the key term in the one-dimensional spatial model], and ε_{ikl} is the corresponding error effect. The term η_{ikl}, known as measurement error or nugget effect, is an intrinsic factor with one level for each experimental unit. It is used as a random factor to fit a second independent residual term when a covariance structure is applied to the usual residuals (Gilmour et al. 1999). The difference between the SS and SSM models is that the latter was fitted with a measurement error, while the former was not. The reason for adopting the SSM model was given in Gilmour et al. (1997), who found further improvement from fitting the measurement error in SSM in comparison to the SS model. In the above models, all effects were treated as random, unless otherwise stated.

RCB and IB models were fitted as described by Cochran and Cox (1957). The one- and two-dimensional spatial analyses of Gleeson and Cullis (1987) and of Cullis and Gleeson (1991) were used for the long-face and row-column trials, respectively, to produce the SS and SSM models called 'spatial models' (Grondona et al. 1996; Gilmour et al. 1997). The approach of Gilmour et al. (1997) was used for the PF model search, starting with the SS model. Then, the variograms (graphical displays of the semi-variance of the difference between plots h rows or columns apart) and the related diagnostics, such as log-likelihood ratio, error mean square, standard error, auto-correlation between plots within a row, auto-correlation between plots within a column and residual plots, were used to facilitate the search for the PF model. The PF model has a target of a flat variogram, largest log-likelihood ratio, and lowest error mean square and standard error. In the PF model, the spatial adjustment was not constrained by the block structure, based on recommendations from Gilmour et al. (1997).

The correlation between the best linear unbiased predictions (BLUPs) based on the mean across all 33 trials in the MET and the mean performance across all 47 trials in the TPE, $r_{p(MT)}$, was estimated, as defined in Eq. 2. This was based on the genotype BLUPs from the different single-trial models applied in the MET in combination with the genotype BLUPs from the preferred model for the TPE.

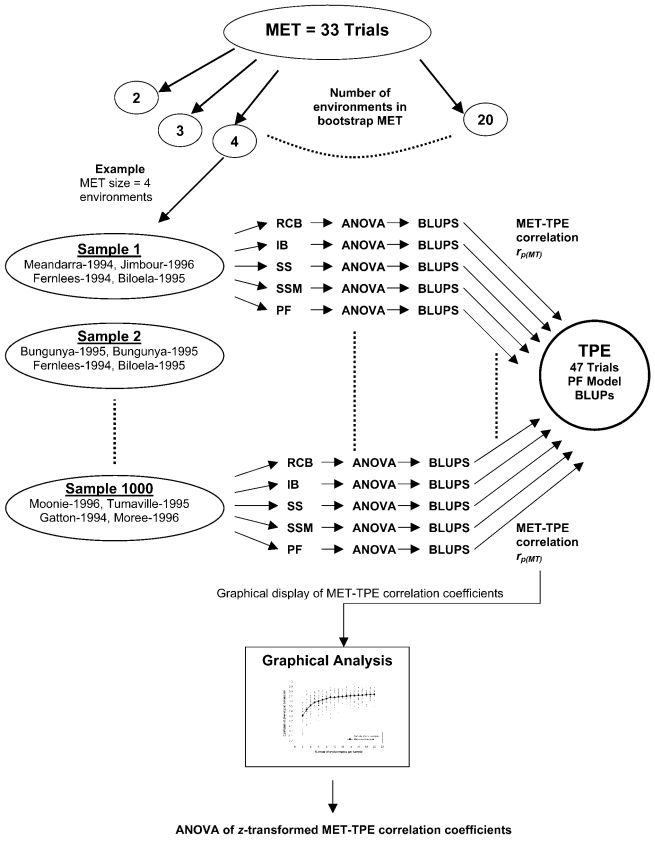

Fig. 1 Schematic representation of the bootstrap sampling of the 33 multi-environment trial (*MET*) regime, from which 1,000 bootstrap samples were taken for MET sizes of 2–20 environments and compared to the genotype best linear unbiased predictions (BLUPs) from the target population of environments (TPE) trials to estimate and analyse the MET-TPE correlation coefficient [$r_{p(MT)}$]

Simulation experiment (bootstrap resampling the MET)

In the current study, the MET data set of the 33 trials can be regarded as the base data set of a sample that comes from the TPE. There are numerous ways that the 33 trials can be combined to represent an MET. A bootstrap simulation was conducted to sample many possible MET structures that could be used to predict relative performance and investigate selection responses measured in the TPE. After initial investigations based on 100 bootstrap samples, we extended this to consider 1,000 bootstrap samples for each of 2, 3, . . . , 20 environments taken from the 33 trials in the full MET data set using the approach of random sampling with replacement (Fig. 1).

For each MET bootstrap sample, a two-stage analysis was conducted for the five genotype-environment data sets based on the adjusted data from application of the five single-trial models: RCB, IB, SS, SSM and PF (Qiao et al. 2000; Fig. 1). In the first stage, the best linear unbiased estimators (BLUEs) of the 18 genotypes were computed for individual trials included in the MET bootstrap sample. These BLUEs were used as the input for the second stage, the combined analysis across environments following the proce-

dures given by Cullis et al. (1996). The model for combined analysis of variance was:

$$y_{ij} = m + g_i + e_j + (ge)_{ij} + [\varepsilon_{ijk}], \quad (3)$$

where y_{ij} is the phenotypic performance of genotype i (i =1, . . . , g) for environment j (j =1, . . . , e), m is the mean of all observations, g_i is the fixed effect of genotype i~N(0, σ^2_g), e_j is the fixed effect of environment j with sum to zero constraint, $(ge)_{ij}$ is the effect of the interaction between genotype i and environment j~N(0, σ^2_{ge}), and $[\varepsilon_{ijk}]$ is the residual plot error effect for genotype i in plot k in environment j~N(0, σ^2_ε), which was estimated as a pooled residual term across all environments. Since only one BLUE and one error estimate were available for each genotype in an environment for the combined analysis, Eq. 3 is a partition of the BLUEs into the relevant genotypic, environmental and G × E interaction effects, plus a residual or error term. The effects of replicates and other within-trial factors were accommodated in the single-trial models and are therefore not included in the combined trial models. For the combined analysis, weights were adopted for each trial according to its replication and the inverse of its residual variance estimate (Cullis et al. 1996). In total, 95,000 combined analyses of the METs were conducted for sample sizes covering 2–20 environments (Fig. 1). This provided a measure of the range and distribution of discrimination among the genotypes for various numbers and combinations of environments sampled from the MET data set. Note that this investigation was not an exhaustive treatment of all possible analyses. The focus was on single-trial model selection, and further improvements may be possible by considering alternative combined analysis models (e.g. Smith et al. 2002).

For each of the five single-trial models used in a sample of environments in the MET, $r_{p(MT)}$ was computed for adjusted genotype mean yield, measured as the BLUPs of the 18 genotypes from the combined analysis of variance between the MET bootstrap sample and the TPE (Fig. 1). The bootstrap-sampling process, as described by Efron and Tibshirani (1993), and estimation of $r_{p(MT)}$ were conducted using a specifically designed Fortran program 'BOOTSSE'. The ASREML software (Gilmour et al. 1999) was used to obtain the genotype BLUPs from the combined analysis for each bootstrap MET sample of environments. The analysis of variance step was automated as part of the BOOTSSE software.

Effect of sample size and analytical methods

The impact of the single-trial models and MET sample size (number of environments) on the estimates of the MET-TPE correlation was portrayed graphically. The effects of the single-trial models on the correlation were examined at all MET sample sizes. The average $r_{p(MT)}$ between MET bootstrap samples and the TPE genotype performance over samples and/or models was calculated using the method of z-transformation described by Sokal and Rohlf (1995), as this transformation was used to analyse the $r_{p(MT)}$ values.

Results

Correlations between the MET and the TPE

The MET-TPE correlation coefficients between adjusted genotype grain yield performance across all 33 MET environments and the grain yield BLUPs from the PF model for the combined analysis of the 47 trials in the TPE were similar for the five single-trial models. They were all significant and positive: $r_{p(MT)}$=0.81, 0.82, 0.80, 0.80 and 0.81 for the RCB, IB, SS, SSM and PF models, respectively.

For single environments, the range of the MET-TPE correlation coefficients varied from no correlation for Gilgandra 1995 and Myall Vale 1995 to as high as 0.73 for Biloela 1995. An effect of single-trial models on the correlation with the TPE genotype mean yield was observed for some of the 33 environments, but the results were not consistent. For example, for locations Billa Billa, Gatton and Warren, the correlation with the TPE was higher for the PF model than for the RCB model, while for locations Fernlees and Moree, the reverse was found for the two single-trial models (data not shown). In general, the Queensland trials were more representative of the TPE than the New South Wales trials. This was likely in part a consequence of the greater use of Queensland locations to generate the TPE data set in this study (Appendix). Future work will focus on broadening the regional representation in the TPE data set. Since no consistent trends were observed for single-environment analyses, these results are not considered further. Greater attention was given to the effects of the single-trial models in combination with multiple-environment analysis.

Effect of MET sample size, environmental composition and single-trial analysis

As the number of environments in the bootstrap MET sample increased, the estimates of $r_{p(MT)}$ increased rapidly for a MET sample size of 2–9 environments for all single-trial models (Fig. 2). The rate of increase was reduced beyond nine environments per MET sample. For each MET size of 2–20 environments, different MET-TPE correlation coefficients were observed for the different combinations of environments represented by the 1,000 bootstrap samples. Therefore, this indicates that for any MET size, the combination of environments sampled can have a large influence on the MET-TPE correlation. However, the variability of these estimates decreased as the MET sample size increased (Fig. 2). Therefore, the influence of differences in MET environmental composition on the MET-TPE correlation decreased as MET size increased.

For MET sample sizes of 2–9 environments, the application of experimental designs (IB model), spatial analysis (SS and SSM models), and their combined use in the PF model displayed advantages over the RCB model in the estimates of the MET-TPE correlation (Fig. 3). The PF model gave the highest estimates of $r_{p(MT)}$, followed by the IB, SS and SSM models. The advantages of the alternative single-trial models became smaller as the environment sample size increased to around ten environments. A close examination of the variance-covariance structure showed that the increase in correlation between samples of the MET and the TPE by the alternative single-trial models over the RCB model was due to the increase in the covariance between the two sets of environments, while the variances of the two sets remained relatively unchanged. This suggests that advanced experimental designs and spatial analyses could have an impact on the correlated genetic advance from the

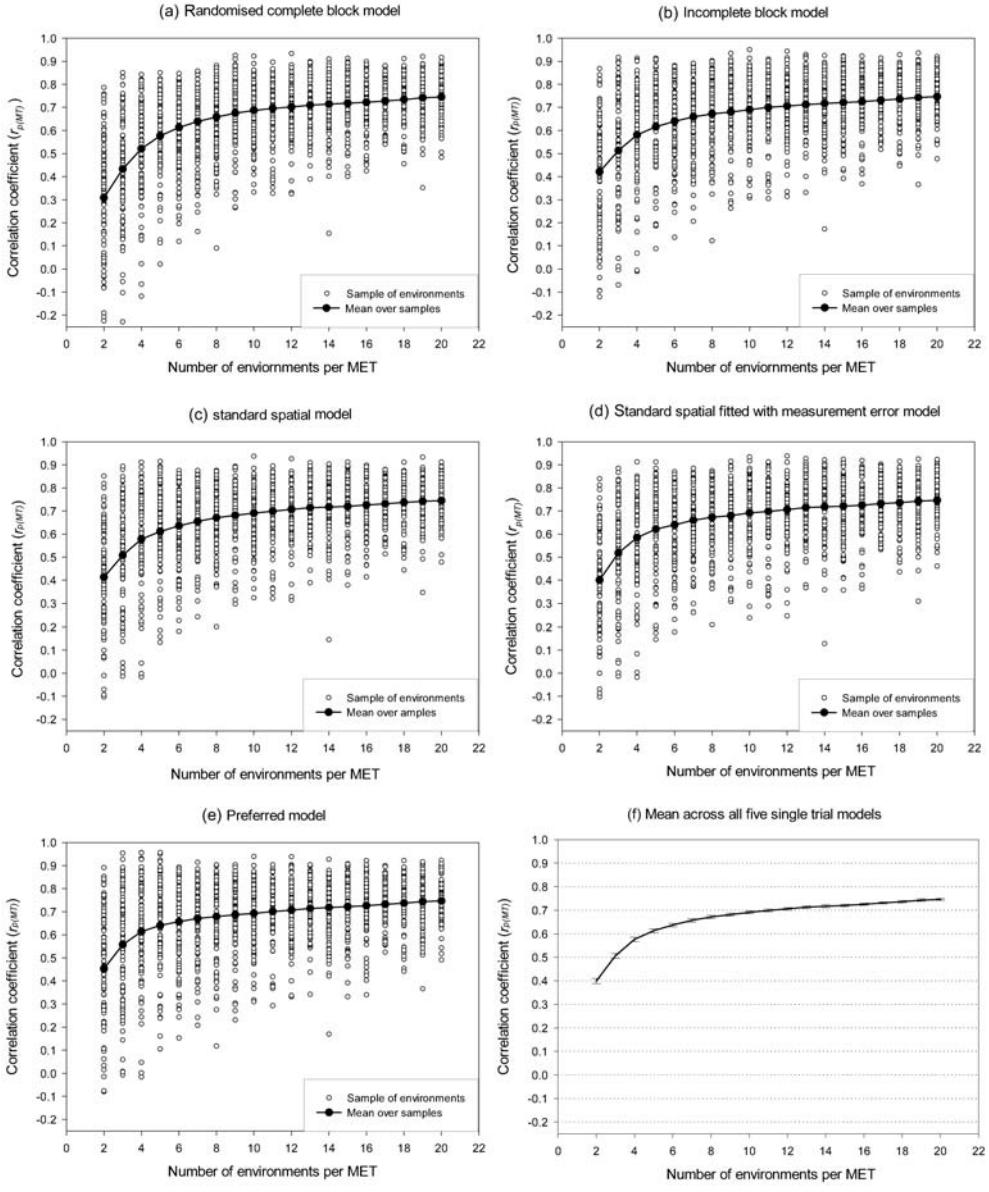

Fig. 2a–f Bootstrap sample estimates of $r_{p(MT)}$ from the combined analysis of bootstrap MET sample sizes of 2–20 environments for five single-trial models: **a** randomised complete block (RCB), **b** incomplete block (IB), **c** spatial analysis (SS), **d** spatial analysis with measurement error (SSM), **e** preferred model (PF), and **f** mean (and standard error) across single-trial models and samples. The *solid lines* connect the bootstrap means across number of environments. Only the first 100 bootstrap MET-TPE correlations are displayed in **a**–**e** to illustrate the variability

MET to the TPE by influencing the covariance between the two sets of environments.

To get some measure of confidence in the differences within the MET sizes (number of environments per sample) appearing in Fig. 3, an analysis of variance of the five correlation values (from the five models) for each of the 1,000 samples was conducted for each MET size. Whilst it is accepted that these 1,000 bootstrap samples were not independent in that they were drawn from the same original 33 METs, the sampling process was independent from one draw to another. It was found that there were significant differences among the five single-trial models at low MET sample sizes (2–9 environments), while there was little difference at higher sample sizes (10–20 environments) (Table 1; Fig. 3). This indicates that the single-trial model analyses, previously examined in terms of reduction in experimental errors and genotype selection for this data set (Qiao et al. 2000), have contributed positively to estimates of correlated response to selection in the TPE for the combined analysis of these MET environments. The same interpretation was obtained when the RCB model was removed from the analyses. In this study, the practical implications of single-trial model analysis for improving rate of genetic gain in a breeding program were measured empirically by evaluating the improvement in the relationship between the MET testing and an estimate of performance in the TPE for a set of genotypes.

Discussion

This study examined the impact of five single-trial models on the correlation coefficient between the esti-

Table 1 Summary analysis of variance of the correlation coefficient, $r_{p(MT)}$, for the five single trial models applied to 1,000 samples for each multi-environment trial (*MET*) size of 2–20 environments (df_{total}=4,999, df_{models}=4, df_{error}=3,996)

MET size	MS_{models}	MS_{error}	*F*	*P*
2	3.623	0.006	651.35	<0.01
3	1.939	0.005	394.56	<0.01
4	1.308	0.005	253.41	<0.01
5	0.952	0.009	104.60	<0.01
6	0.484	0.007	73.30	<0.01
7	0.290	0.008	34.41	<0.01
8	0.135	0.011	12.52	<0.01
9	0.057	0.015	3.71	0.01
10	0.036	0.019	1.93	0.10
11	0.029	0.017	1.69	0.15
12	0.022	0.016	1.36	0.25
13	0.019	0.015	1.27	0.28
14	0.016	0.014	1.11	0.35
15	0.012	0.012	0.95	0.43
16	0.010	0.011	0.89	0.47
17	0.009	0.011	0.80	0.52
18	0.005	0.008	0.57	0.68
19	0.002	0.006	0.39	0.82
20	0.001	0.004	0.21	0.93

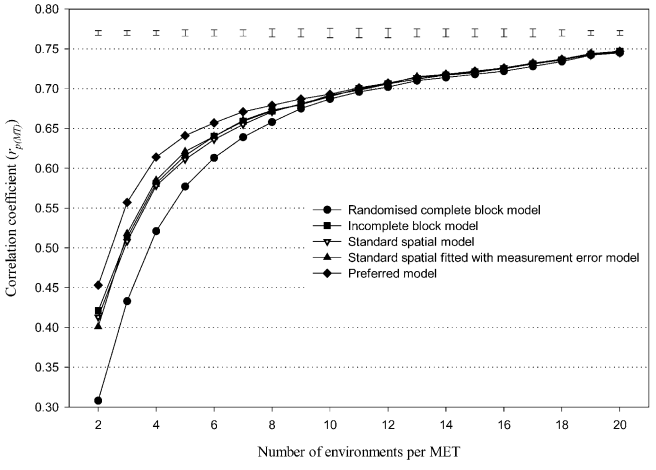

Fig. 3 Bootstrap sample estimates of the mean $r_{p(MT)}$ for five single-trial models averaged over all samples at MET sample sizes of 2–20 environments. Least significant differences among the means for each number of environments per MET are shown as *bars* across the *top* of the figure

mates of genotype mean yield for 18 common genotypes in two series of experiments, one representative of the current breeding program MET and the other representative of the TPE for the breeding program in the northern grains region of Australia. The purpose of this investigation was to obtain an empirical estimate of the expected impact of the alternative single-trial models on response to selection in the TPE when they are applied within the wheat-breeding MET system of the northern grains region of Australia.

For the study by Qiao et al. (2000), using the MET data sets considered in this paper, improvements in efficiency detected at the single-trial level were associated with changes in the ranks and selection of genotypes. On average, these changes in selection of genotypes are expected to improve the correlated response to selection in the TPE. In the present study, for the analysis of the 33 MET trials taken individually, none except one (Biloela 1994) of the single-trial models showed a clear advantage over the RCB model when measured in terms of the correlation between genotype performance in the individual trial and in the TPE. However, advantages were detected when the single-trial models were used in combination with multi-environment testing.

The estimates of the MET-TPE correlation coefficient increased with the number of environments in the MET. Therefore, with an increasing number of environments sampled in the METs, on average, there was an improvement in the representation of genotype performance in the TPE by the MET. Increasing the MET size beyond ten environments resulted in a phase of diminishing returns for additional environments. In addition to the number of environments sampled in the METs, the combination or composition of environments in the sample was important in determining the response to selection in the TPE, and there was a large variation in the MET-TPE correlation for all sizes of MET. Both MET sample size and single-trial models influenced the MET-TPE correlation, with MET sample size having a larger effect in this study. There was an interaction between MET sample size and the single-trial model. Therefore, the influence of the single-trial model was dependent on the MET sample size. This study has demonstrated a consistent advantage of the PF model over the RCB model for the MET-TPE correlation for MET sizes from two to nine environments.

While, theoretically, one would have expected that the importance of the single-trial model would diminish as the number of environments goes up, this study is, as far as we know, the first comprehensive empirical demonstration of the theoretical expectation. Smith et al. (2002) showed a similar result using simulation.

These findings indicate advantages from improving the structure of the wheat-breeding METs in the northern grains region of Australia by using experimental design and spatial analysis methods in combination with METs based on 2–9 environments that sample the diverse conditions of the TPE. No advantage was detected for METs based on greater than nine environments. Therefore, this finding is likely to be especially useful in situations where the wheat breeders are confronted with testing a large number of genotypes in a few environments. The early generation trials of the wheat breeding programs at the Leslie Research Centre and the recurrent selection strategy of the Wheat Germplasm Enhancement Program both encounter this situation. The following areas are identified for further investigation:

– Broadening the environmental representation of the current TPE data set to include environments from the two Australian states represented in the northern grains region, Queensland and particularly New South Wales.

– Evaluating these findings for different sets of genotypes that represent different reference populations and stages of selection (i.e. early generation vs advanced line testing) relevant to wheat breeding in the northern grains region.
– Evaluating the types of environmental conditions that contribute to the variable MET-TPE correlation with a view to identifying a MET regime that contributes to a consistently higher MET-TPE correlation.

Acknowledgements The contributions of the field trial staff at the Leslie Research Centre, Toowoomba, and the Plant Breeding Institute, Narrabri, are gratefully acknowledged. We greatly appreciate the assistance of Dr. Dean Podlich with the BOOTSSE software and automation of the bootstrap analysis. Comments by the referees were helpful and resulted in an improved version of this manuscript. The Grain Research Development Corporation of Australia funded this research.

Appendix

Description of the 80 wheat trials. The first 33 were used to represent the multi-environment trial (MET) regime, while the last 47 were used to represent the target population of environments (TPE) in the northern grains region.

Trial	Year	Location	Trial type	Design	Field layout	No. of lines	No. of Replicates	Preferred model (PF)[a]
1	1994	Biloela, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+s(row)+AR1
2	1994	Jimbour, Qld.	Long-face	Triple lattice	1×149	49	3	rep+blk/rep
3	1994	Meandarra, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+blk/rep+AR1
4	1994	Oakeleigh, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+blk/rep+AR1
5	1994	The Gum, Qld.	Long-face	Triple lattice	1×149	49	3	blk/rep+AR1
6	1995	Billa Billa, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+s(row)+rep+AR1
7	1995	Biloela, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+s(row)+blk/rep+AR1
8	1995	Bungunya, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+s(row)+blk/rep+AR1
9	1995	Tumaville, Qld.	Long-face	Triple lattice	1×149	49	3	l(row)+s(row)+blk/rep+AR1
10	1994	Fernlees, Qld.	Row-column	Triple lattice	7×21	49	3	fcol+frow+AR1×AR1
11	1994	Gatton, Qld.	Row-column	Triple lattice	5×15	25	3	fcol+frow+row/rep+AR1×AR1
12	1995	Gatton, Qld.	Row-column	Triple lattice	5×15	25	3	frow+col+col/rep+AR1×AR1
13	1996	Billa Billa, Qld.	Row-column	Triple lattice	5×15	25	3	fcol+frow+rep+AR1×AR1
14	1996	Bungunya, Qld.	Row-column	Triple lattice	5×15	25	3	fcol+frow+l(row)+l(col)+AR1×AR1
15	1996	Fernlees, Qld.	Row-column	Triple lattice	5×15	25	3	fcol+frow+col/rep+AR1×AR1
16	1996	Gatton, Qld.	Row-column	Triple lattice	5×15	25	3	frow+col+AR1×AR1
17	1996	Jimbour, Qld.	Row-column	Triple lattice	5×15	25	3	fcol+frow+rep+AR1×AR1
18	1996	Moonie, Qld.	Row-column	Triple lattice	5×15	25	3	fcol+frow+AR1×AR1
19	1994	Moree, NSW	Row-column	Triple lattice	5×15	25	3	l(row)+l(col)+s(row)+s(col)+row/rep +AR1×AR1
20	1994	Myall Vale, NSW	Row-column	Triple lattice	5×15	25	3	fcol+frow+AR1×AR1
21	1994	Narrabri, NSW	Row-column	Triple lattice	5×15	25	3	fcol+frow+AR1×AR1
22	1994	Narrabri Late, NSW	Row-column	Triple lattice	5×15	25	3	fcol+frow+AR1×AR1
23	1994	North Star, NSW	Row-column	Triple lattice	5×15	25	3	fcol+frow+AR1×AR1
24	1994	Warren, NSW	Row-column	Triple lattice	5×15	25	3	fcol+frow+blk/rep+AR1×AR1
25	1995	Gilgandra, NSW	Row-column	Triple lattice	6×18	36	3	col+col/rep+row/rep+AR1×AR1
26	1995	Myall Vale, NSW	Row-column	Triple lattice	6×18	36	3	row+col+AR1×AR1
27	1995	North Star, NSW	Row-column	Triple lattice	6×18	36	3	fcol+frow+AR1×AR1
28	1996	Gilgandra, NSW	Row-column	Triple lattice	18×6	36	3	fcol+frow+rep+row/rep+AR1×AR1
29	1996	Moree, NSW	Row-column	Triple lattice	18×6	36	3	fcol+row+rep+AR1×AR1
30	1996	Myall Vale, NSW	Row-column	Triple lattice	18×6	36	3	fcol+frow+ AR1×AR1
31	1996	Narrabri, NSW	Row-column	Triple lattice	18×6	36	3	fcol+frow+rep+AR1×AR1
32	1996	North Star, NSW	Row-column	Triple lattice	18×6	36	3	fcol+row/rep+AR1×AR1
33	1996	Spring Ridge, NSW	Row-column	Triple lattice	18×6	36	3	fcol+rep+AR1×AR1
34	1988	Emerald, Qld.	Row-column	Augmented	73×2	55	-	l(col)+row+s(col)+AR1×AR1+ nugget
35	1988	Emerald, Qld.	Row-column	Augmented	73×2	55	-	l(col)+row+s(col)/row+AR1×AR1
36	1988	Emerald, Qld.	Row-column	Augmented	73×2	21	-	l(col)+row+s(col)+AR1
37	1988	Emerald, Qld.	Row-column	Augmented	79×2	55	-	l(col)+row+s(col)/row+AR1
38	1988	Emerald, Qld.	Row-column	Augmented	73×2	55	-	l(col)+row+s(col)/row+AR1×AR1
39	1988	Emerald, Qld.	Row-column	Augmented	37×2	21	-	l(col)+s(col)+AR1×AR1
40	1989	Emerald, Qld.	Row-column	Augmented	160×2	251	-	row+AR1×AR1
41	1989	Emerald, Qld.	Row-column	Augmented	160×2	251	-	l(col)+row+s(col)+AR1×AR1
42	1989	Emerald, Qld.	Row-column	Augmented	160×2	251	-	l(col)+row+s(col)+AR1×AR1
43	1988	Kingthorpe, Qld.	Row-column	Augmented	73×2	55	-	l(col)+row+s(col)+AR1×AR1
44	1988	Kingthorpe, Qld.	Row-column	Augmented	79×2	56	-	l(col)+row+s(col)/row+AR1×AR1
45	1988	Kingthorpe, Qld.	Long-face	Augmented	64×1	21	-	l(col)+rep+s(col)/rep+AR1
46	1988	Kingthorpe, Qld.	Row-column	Augmented	73×2	54	-	l(col)+row +s(col)/row+AR1×AR1
47	1988	Kingthorpe, Qld.	Row-column	Augmented	73×2	53	-	l(col)+s(col)+AR1×AR1
48	1988	Kingthorpe, Qld.	Long-face	Augmented	64×1	21	-	l(col)+AR1
49	1989	Kingthorpe, Qld.	Row-column	Augmented	160×2	251	-	l(col)+s(col)/row+AR1×AR1
50	1989	Kingthorpe, Qld.	Row-column	Augmented	160×2	251	-	l(col)+row+s(col)/row+AR1×AR1

Trial	Year	Location	Trial type	Design	Field layout	No. of lines	No. of Replicates	Preferred model (PF)[a]
51	1989	Kingthorpe, Qld.	Row-column	Augmented	160×2	251	-	l(col)+row+s(col)+AR1×AR1
52	1989	Kingthorpe, Qld.	Row-column	Augmented	160×2	251	-	l(col)+row+s(col)+AR1×AR1
53	1988	Gatton, Qld.	Row-column	Augmented	74×2	55	-	l(col)+s(col)+AR1×AR1
54	1988	Gatton, Qld.	Row-column	Augmented	74×2	55	-	l(col)+row+s(col)+AR1×AR1
55	1988	Gatton, Qld.	Row-column	Augmented	15×3	15	-	l(row)+l(col)+s(col)/row+s(row)/col+AR1×AR1
56	1988	Gatton, Qld.	Row-column	Augmented	15×3	15	-	l(col)+l(row)+s(col)/row+AR1×AR1
57	1989	Gatton, Qld.	Row-column	Augmented	54×6	246	-	l(col)+col/row+AR1×AR1
58	1989	Gatton, Qld.	Row-column	Augmented	54×6	251	-	l(col)+l(row)+col+s(col)+s(row)+AR1×AR1
59	1988	Biloela, Qld.	Long-face	Incomplete block	48×1	15	3	l(col)+blk/rep+s(col)/rep+AR1
60	1988	Brookstead, Qld.	Long-face	Incomplete block	48×1	15	3	l(col)+s(col)/rep+AR1
61	1988	Fernlees, Qld.	Long-face	Incomplete block	48×1	15	3	blk/rep+AR1
62	1988	Jimbour, Qld.	Long-face	Incomplete block	48×1	15	3	l(col)+blk/rep+s(col)/rep+AR1
63	1988	The Gum, Qld.	Long-face	Incomplete block	48×1	15	3	l(col)+blk/rep+s(col)/rep+AR1
64	1988	Toobeah, Qld.	Long-face	Incomplete block	48×1	15	3	l(col)+blk/rep+s(col)+AR1
65	1988	Gatton, Qld.	Row-column	Randomised complete block	15×2	15	2	l(col)+s(col)+AR1×AR1
66	1988	Gatton, Qld.	Row-column	Randomised complete block	15×2	15	2	l(col)+rep+s(col)/rep+AR1×AR1
67	1988	Gatton, Qld.	Row-column	Randomised complete block	15×2	15	2	rep+AR1×AR1
68	1986	Brookstead, Qld.	Row-column	Triple lattice	7×21	49	3	l(col)+l(row)+rep+col/rep+s(row)/col+AR1×AR1
69	1986	Brookstead, Qld.	Row-column	Triple lattice	7×21	49	3	l(col)+l(row)+rep+row/rep+col/rep+s(row)/col+AR1×AR1
70	1986	Cecil Plains, Qld.	Row-column	Triple lattice	7×21	49	3	l(col)+row/rep+col/rep+AR1×AR1
71	1986	Cecil Plains, Qld.	Row-column	Triple lattice	7×21	49	3	l(col)+l(row)+s(col)/row+AR1×AR1
72	1987	Gatton, Qld.	Row-column	Triple lattice	21×7	49	3	l(row)+row+col/rep+s(row)/rep+AR1×AR1
73	1987	Gatton, Qld.	Row-column	Triple lattice	21×7	49	3	l(col)+l(row)+row/rep+col/rep+s(row)+s(col)/rep+AR1×AR1
74	1987	Tummaville, Qld.	Row-column	Triple lattice	21×7	49	3	l(col)+l(row)+row/rep+col/rep+s(row)+s(col)/rep+AR1×AR1
75	1987	Tummaville, Qld.	Row-column	Triple lattice	21×7	49	3	l(col)+l(row)+row+col/rep+s(col)/row+s(row)+AR1×AR1
76	1993	Gatton, Qld.	Long-face	Triple lattice	147×1	49	3	l(col)+blk/rep+s(col)/rep+AR1
77	1993	Inglestone, Qld.	Long-face	Triple lattice	147×1	48	3	blk/rep+AR1
78	1993	Brookstead, Qld.	Long-face	Triple lattice	147×1	49	3	l(col)+blk/rep+s(col)+AR1
79	1993	Narrabri, NSW	Long-face	Triple lattice	147×1	47	3	l(col)+blk/rep+s(col)+AR1
80	1993	Biloela, Qld.	Long-face	Triple lattice	147×1	49	3	l(col)+blk/rep+s(col)/rep+AR1

[a] row, random row effect; row/rep, row within replicates effect; col/rep, column within replicates effect; blk/rep, block within replicates effect; col, random column effect; frow, fixed row effect; fcol, fixed column effect; rep, replicate effect; l(col), fixed linear column effect; l(row), fixed linear row effect; ARI, one-dimensional auto-regressive process; ARI×ARI, two-dimensional auto-regressive process; s(row), a random curvature component for rows, spline row effect; s(col), a random curvature component for columns, spline column effect; s(col)/rep, spline column within replicates effect; s(row)/rep, spline row within replicates effect; s(col)/row, spline column within rows effect; s(row)/col, spline row within columns effect; col/row, column within rows effect; row/col, row within columns effect; nugget, measurement error

Note that a spline effect is fitted in two components—a fixed linear trend and a random curvature component

References

Basford KE, Cooper M (1998) Genotype × environment interactions and some considerations of their implications for wheat breeding in Australia. Austr J Agric Res 49:153–174

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

Comstock RE (1977) Quantitative genetics and the design of breeding programmes. In: Pollack E, Kempthorne O, Bailey TBJ (eds) Proceedings of the International Conference on Quantitative Genetics. Iowa State University Press, Ames, Iowa, pp 705–718

Comstock RE, Moll RH (1963) Genotype-environment interactions. In: Hanson WD, Robinson HF (eds) Statistical genetics and plant breeding. Natl Acad Sci Natl Res Pub 982:164–196

Cooper M, DeLacy IH (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. Theor Appl Genet 88:561–572

Cooper M, Podlich DW (1999) Genotype × environment interactions, selection response and heterosis. In: Coors JG, Pandey S

(eds) The genetics and exploitation of heterosis in crops. ASA-CSSA-SSSA, Madison, Wis., pp 81–92
Cullis BR, Gleeson AC (1989) Efficiency of neighbour analysis for replicated variety trials in Australia. J Agric Sci (Camb) 113:233–239
Cullis BR, Gleeson AC (1991) Spatial analysis of field experiments—an extension to two dimensions. Biometrics 47:1449–1460
Cullis BR, Thomson FM, Fisher JA, Gilmour AR, Thompson R (1996) The analysis of the NSW wheat variety database. II. Variance component estimation. Theor Appl Genet 92:28–39
DeLacy IH, Basford KE, Cooper M, Bull JK, McLaren CG (1996) Analysis of multi-environment trials—an historical perspective. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, IRRI and ICRISAT, Wallingford, UK, pp 39–124
Efron B, Tibshirani R (1993) An introduction to the bootstrap. Chapman and Hall, New York
Falconer DS (1952) The problem of environment and selection. Am Nat 86:293–298
Falconer DS (1989) Introduction to quantitative genetics, 3rd edn. Longman, London
Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiments. J Agric Biol Environ Stat 2:269–293
Gilmour AR, Cullis BR, Welham SJ, Thompson R (1999) ASREML Program User Manual. New South Wales Agriculture, Orange, New South Wales
Gleeson AC, Cullis BR (1987) Residual maximum likelihood (REML) estimation of a neighbour model for field experimentation. Biometrics 43:277–288
Grondona MO, Crossa J, Fox PN, Pfeiffer WH (1996) Analysis of variety yield trials using two-dimensional separable ARIMA processes. Biometrics 52:763–770
John JA, Whitaker D (1993) Construction of resolvable row-column designs using simulated annealing. Austr J Stat 35:237–245
Kirk HJ, Haynes FL, Monroe RJ (1980) Application of trend analysis to horticultural field trials. J Am Soc Hort Sci 105:189–193
Lin CS, Morrison MJ (1992) Selection of test locations for regional trials of barley. Theor Appl Genet 83:968–972
Magnussen S (1993) Design and analysis of tree genetic trials. Can J For Res 23:1114–1149
Matassa VJ, Basford KE, Stringer JK, Hogarth DM (1998) The application of spatial analysis to sugar cane variety trials. Proc Austr Soc Sugar Cane Technol 20:1–7
Qiao CG, Basford K E, DeLacy IH, Cooper M (2000) Evaluation of experimental designs and spatial analyses in wheat breeding trials. Theor Appl Genet 100:9–16
Smith A, Cullis B, Luckett D, Hollamby G, Thompson R (2002) Exploring variety-environment data using random effects AMMI models with adjustments for spatial field trend. Part 2: Applications. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding. CAB International, Wallingford, UK, pp 337–351
Sokal RR, Rohlf FJ (1995) Biometry, 3rd edn. Freeman, New York
Williams ER (1986) Row and column designs with contiguous replicates. Austr J Stat 28:154–163
Williams ER, Matheson AC (1994) Experimental design and analysis for use in tree improvement. CSIRO, Melbourne, Australia
Yates F (1939) The recovery of the inter-block information in variety trials arranged in three-dimensional lattices. Ann Eugen 9:136–156
Zimmermann DL, Harville DA (1991) A random field approach to the analysis of field-plot experiments and other spatial experiments. Biometrics 47:223–239