Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials

Scott C. Chapman

Abstract Crop simulation models of plant processes capture the biological interactions between the sensing of signals at an organ level (e.g. drought affecting roots), the response of the plant at a biochemical level (e.g. change in development rate) and the result at the organ (or crop) level (e.g. reduced growth). In dissecting the complex control of phenotypes like yield, simulation models have several roles. Models have been used to generate an index of the climatic environment (e.g. of drought stress) for breeding programme trials. In wheat and sorghum grown in northern Australia, this has shown that mid-season drought generates large genotype by environment interaction. By defining gene action to calculate the value of input trait parameters to crop models, simulated multi-environment trials estimate the yield of 'synthetic' sorghum cultivars grown in historical or artificial climates with current or potential management regimes. In this way, the biological interactions among traits constrain the crop yields to only those that are biologically possible in the given set of environments. This allows the construction of datasets that are more 'realistic' representations of gene by trait by environment interaction than is possible using only the statistical attributes (e.g. means, variances and correlations) of real-world trait datasets. This approach has an additional advantage in that 'biological and experimental noise' can be manipulated separately. These 'testbeds' for statistical techniques can be extended to the interpretation of a crossing and selection programme where the processes of chromosomal recombination are simulated using a quantitative genetics model and applied to the trait parameters. Statisticians are challenged to develop improved methods for the resulting simulated phenotype datasets, with the objective of revealing the (known) underlying genetic and environment structure that was input to the simulations. These improved methods can then be applied to existing plant breeding programmes.

Keywords Crop physiology · Crop simulation models · Genotype by environment interaction · GxE · Principal component analysis · PCA

Introduction

The principal objective of a plant-breeding programme is the generation and selection of new gene combinations to create genotypes with trait performance that is superior to current genotypes, within the target population of environments (TPE) (Comstock 1977). This objective applies equally to conventional, molecular and combined approaches. Traditional methods of crop improvement have met with continued success over

S. C. Chapman (✉)
CSIRO Plant Industry, Queensland Bioscience Precinct,
305 Carmody Rd, St. Lucia, QLD, Australia
e-mail: scott.chapman@csiro.au

the last century. In recent years, biotechnology has rev-olutionised development of new varieties with genes being identified for engineering or marker-assisted selection for a range of 'simple' trait phenotypes in several crop species (see Chapman et al. (2002b) for a short review and some examples). While there is cur-rently substantial investment in the study of more com-plex trait phenotypes, such as response to drought, there are not yet success stories based on the identifica-tion of specific genes and their utilisation for this chal-lenge.

Plant breeders have continued to improve crops for performance under drought using both conventional testing and selection regimes, as well as trait-based selection involving physiological indicators of perfor-mance such as anthesis-silking interval (ASI) in maize (Edmeades et al. 2004); transpiration efficiency (TE) in wheat (Condon et al. 2004) and modifications of phenology and growth habit in many species to suit seasonal water supply. The major successes have occurred where a single trait has been identified, and its response to the environment has either been mini-mal (e.g. TE in wheat); or has been exploited through careful management of phenotype screening (e.g. greater genotypic expression of ASI under managed drought conditions (Edmeades et al. 1999)). Explic-itly, by identifying and selecting an adaptive trait, the breeders and/or physiologists have been able to understand what drives at least a part of the genotype by environment interaction in such environments. However, in many cases of plant adaptation to the abiotic environment, these drivers are difficult to identify as their effects are masked by different types of interactions: between the genes controlling 'com-ponent' traits, physiological interactions among com-ponent traits and target selection traits and interactions between genes, traits and environments (seasons and locations). In field experiments, these different types of interactions (sources of variation) cannot be separated out from components of variance that are typically assigned to genotype by environ-ment interaction and experimental error.

Both phenotypic and 'gene-based' selection have been successful in assembling defence-gene net-works. These are similar to the gene networks that apparently respond to extremes of abiotic stress and are a major subject of research in molecular biology, e.g. Bartels et al. (1996). Traits like solute accumula-tion are some of last processes to occur in stressed

plant tissues prior to cell death (Hsiao 1973; Ludlow and Muchow 1990) and the plant has already under-taken many adaptations in growth and development processes prior to response. In this respect, major genes controlling normal growth and development are also 'stress' genes and are fundamental to 'drought adaptation'. For example, cell expansion, rather than photosynthesis or solute accumulation is the first pro-cess to be affected by decreased tissue water potential (Hsiao 1973). Cell expansion (driving leaf elonga-tion) has been shown to be under strong genetic con-trol (Reymond et al. 2003) and is related to how the canopy develops and the balance of water use through the season. Accordingly, coping with the seasonal variation in radiation, temperature and water supply is largely a matter of a favourable coordination of the genes that are already expressed under 'optimal' growing conditions. It is this aspect of coordinated growth and development to enable the plant to con-tinue to efficiently capture resources (i.e. continuing canopy and root expansion) for improved crop yield that will generate economic benefits in the many envi-ronments where plants experience substantial, but not catastrophic, levels of stress (Bidinger et al. 1996).

As outlined above and discussed in detail else-where (for examples see Chapman et al. 2002b; Edmeades et al. 2004; Campos et al. 2004; Cooper et al. 2006), substantial research effort in drought adaptation is being invested in the two approaches: phenotypic selection for candidate traits (including selection for yield) and identification and selection or engineering for potential adaptive genes. Candidates for the latter are a point of dispute, as some are 'magic bullets' from desert plants, while others are regulators to control the growth and development traits dis-cussed above. There is a need for a comprehensive framework to allow the identification of efficient eval-uation and selection strategies in both of these approaches. The combination of crop simulation and genetic simulation models together with extensive databases of input and output information provides a testable implementation of this framework (Chapman et al. 2003).

Crop simulation models vary greatly in their utility for the applications discussed. These models are described as dynamic (taking daily weather data as input) and deterministic (comprised of interacting modules that describe major plant processes) in the terminology of (Thornley and France 2004). The

models discussed here have been developed by groups of scientists using over 20 years of detailed experiments on germplasm that exhibit a range of genetic variation for various adaptive traits similar to that seen in the breeding programmes of the region. The model development has the objective of accommodating physiological relationships between traits to allow expression of biological interactions among traits and the responses of growth to both soil and aerial environments (Hammer et al. 1996a, 2006; Chapman et al. 2002b). It is essential that the model development research applies this focus on capturing the trait dependencies when attempting to replicate observable genotype by environment interaction, rather than simply adopting conventional models that have been used for agronomic prediction in farming systems. Chenn et al. (2008) demonstrate this principle by integrating an hourly time-step model of the genetic and environmental control of leaf growth (Reymond et al. 2003) into a crop simulation model for maize.

This paper outlines some of the challenges to which this framework has been applied in breeding programmes within our research group to dryland crops (wheat and sorghum) in northern Australia and much of the detail has been previously reported. This paper is an overview of the existing and potential statistical methodologies that might be tested using the simulation framework to investigate GEI. Three example areas of the application are considered: (i) models as 'environment integrators' to interpret genotype by environment interaction observed in field trials; (ii) model simulation of GEI in multi-environment trials to understand how yield is driven by different combinations of traits; (iii) interpretation of changes in GEI with selection in a simulated breeding programme.

Models as environment integrators of drought patterns

There have been numerous (100's) papers published to study the effects of drought on GEI in a range of crops. In some cases, the authors have focused on explaining GEI in terms of component traits and their interactions with the environment (e.g. (Chapman et al. 1997; Campos et al. 2004 ; Rizza et al. 2004) for maize; Cooper et al. 1995, 2006) while others have

used environmental indices to interpret GEI (e.g. Voltas et al. 1999). In wheat, Saulescu and Kronstad (1995) proposed the use of a 'simulated entry' to be used as an unbiased 'check' for the environment effects, and that type of approach was also considered in sorghum work here with Hammer and Jordan (2007) giving a more detailed framework for its use. In past work, we have used models to describe the environment in terms of drought stress patterns and here revisit examples for wheat and sorghum in Australia. Loffler et al. (2005) applied the approach of a crop model index to explain a substantial proportion of GEI in US maize breeding trials.

GEI in wheat is related to seasonal timing of drought stress

In the northern wheat-growing region of Australia, spring wheat is planted in autumn and grows largely on stored soil water through the winter, resulting in mean farm yields of 1.5–2.5 t/ha. Rainfall during winter and spring may contribute to increased yields, but is unreliable, as is the summer rainfall (average ca. 500 mm in September to March with a standard deviation of 150–200 mm (Chapman et al. 2000b). Therefore, it is of interest to dryland farmers to access wheat cultivars that can both tolerate periods of stress, and also perform well during wetter years when profits are maximised (yields of 3–5 t/ha) for the same cost of inputs. This presents a challenge for breeders who cannot reliably sample the range of environments in their on-farm evaluation trials each year due to inter-annual variation in rainfall. A dataset of 76 wheat 'probe' trials was collected over 12 years and 22 locations using standard management inputs as well as management regimes (additional irrigation and fertiliser) as reported by Cooper et al. (1997) and Qiao et al. (2000). The same 18 genotypes were used in each trial, being a sample of the local germplasm as well as additional superior yielding (but lower grain quality) lines from CIMMYT. Several CIMMYT lines had a substantial yield advantage (10–20%), but not in all testing environments. The research determined that, in evaluation of potential germplasm, it was necessary to include one or more 'high-input' trials in the established trialling system to ensure that lines with potential in wetter years were not accidentally overlooked through poor sampling of the variable on-farm environments.

For each of the 76 experiments, the daily weather record (temperature, radiation, rainfall), soil water holding characteristics and record of sowing and management was assembled (Chapman et al. 2001). These were used as input to the APSIM-Wheat model (Version 2.1; Wang et al. 2002). For each trial, the model parameters were modified to match the predicted flowering date with the observed flowering date for a standard check genotype, Hartog. In a subset of more detailed trials, the model predicted well the anthesis biomass dry weight and final grain yield (Fig. 1). Each day, the model outputs a water stress index, which indicated the degree to which potential water supply from the volume of soil explored by roots is able to match the water demand as determined by radiation and temperature impacting the crop canopy. The index drops below 1 as stress begins to affect plant growth and across the season indicates the temporal pattern of water stress that would have been experienced by cultivars in the trial. Before further analysis of each of the 76 trials, the stress index was averaged for every 100°Cd of thermal time from sowing to harvest.

Across all of the trials, clustering identified two groups of trials (Chapman et al. 2000b) as determined from the seasonal water stress patterns for Hartog (Fig. 2). In 56% of the environments (trials), the crops experienced a moderate degree of water stress from just before flowering to mid-grainfill (ET1). This group of trials included several that had received full or supplementary irrigation (Cooper et al. 1997). A moderate to severe water stress in the period prior to flowering and through grain filling occurred in the remaining 44% of the trials.

Using the genotype means (Best Linear Unbiased Estimates) and relative weights from existing analyses (Qiao et al. 2000), the variance components for genotype and GEI were estimated using the ASREML software (Gilmour et al. 1997) and the methods described by Cullis et al. (1996) and Chapman et al. (2000a). Across the 76 trials, the GEI variance component was larger than that for genotype, as has been typically observed in these environments (Table 1). This leads to substantial confounding of genotype rankings and is characteristic of the low across-trial heritabilities typical of this region of

Fig. 1 Performance of APSIM-Wheat model in simulating (**a**) anthesis biomass (g m^{-2}) and (**b**) grain yield (g m^{-2}) of cultivar Hartog in 16 breeding trials
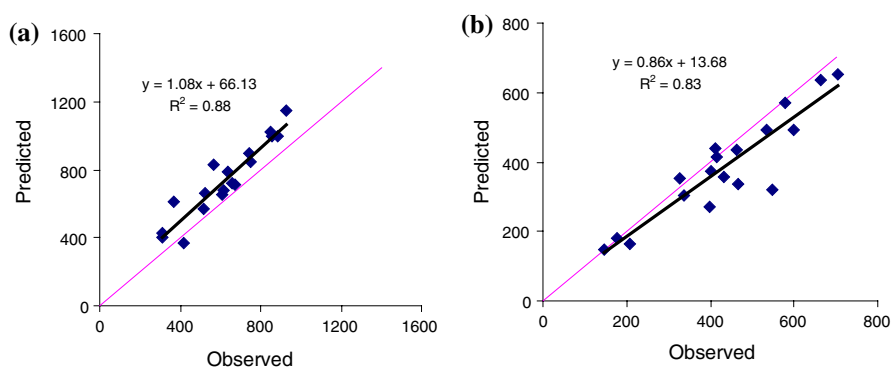

Fig. 2 Mean water stress index (ratio of canopy water demand to root supply capacity) for two environment types averaged across 76 trials in 23 locations and 12 years. The matrix of the index (averaged over 17 periods each 100°Cd long) across trials was clustered into two environment types before averaging. Arrow indicates averaged flowering date across trials
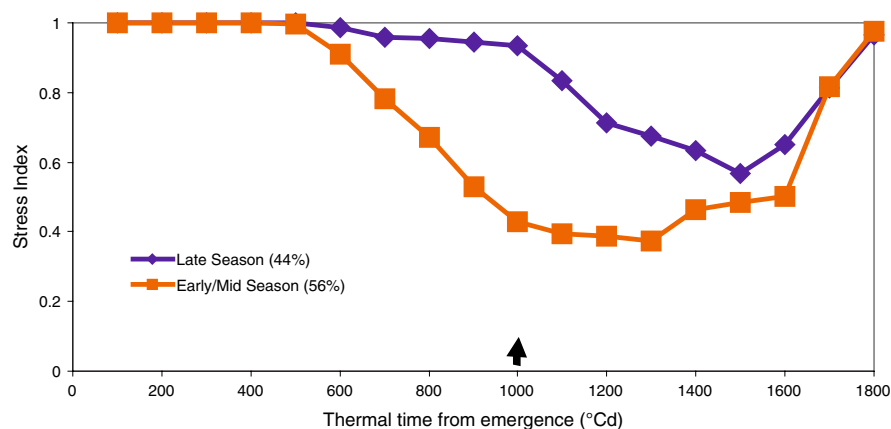
Table 1 Genotype by environment analyses of yield in 76 wheat trials and for trials classified within two drought environment types

Source	Variance component	Standard error
Gen	0.0830	0.0270
Gen.Env	0.1197	0.0065
Gen within ET1	0.1415	0.0472
Gen within ET2	0.0379	0.0139
Gen.ET group	0.1146	0.0063
Residual Variance	0.0960	

variable rainfall (Cooper et al. 1997; Mathews et al. 2002). Both of these potentially impede selection and genetic advance. The analyses were repeated using the environment types from the drought pattern classification to explain the main effect of genotype. The genotypic variance in the mild stress environment types (ET1) was three times as large as the component within ET2 (Table 1). The interaction component for Genotype by Environment Type interaction was similar in size to that for the GEI in the standard analysis. Thus, the use of the model as an independent measure of the environment drought stress pattern was able to confirm:

(1) The speculation of (Cooper et al. 1997) (based on rainfall patterns and soil types) that a substantial amount of the GEI observed in wheat trials was a result of GEI related to the timing of drought stress, especially in the period around flowering.

(2) That environment sampling needed to include 'managed low stress' trials to allow more reliable identification of genotypes adapted to the infrequent but profitable wetter years.

(3) That genotypes existed (cv. Cunningham) which contributed substantially to the GEI through its relatively poorer performance in droughted environments (ET2) even though it performed moderately well in mild stress environments (data not presented).

Spatial variation in GEI of sorghum is related to frequency of drought at different locations
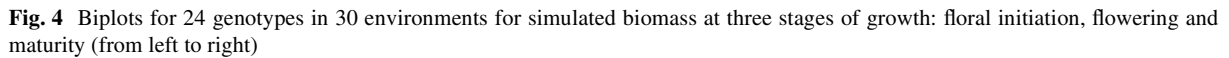
In summer, sorghum is the major crop in the northern wheat-growing region, averaging about 3 million t of production. Sorghum crops depend on both stored soil water and in-season rainfall to achieve average yields of ca 2.5 t/ha in this region, with yields of up to 8 of 10 t/ha in the best cases. Due to spatial and temporal variability in rainfall, sorghum and wheat breeders face similar challenges in being able to sample environments appropriately over time. Applying the method of DeLacy et al. (1996) to unbalanced multi-environment trial data collected over 17 years, Chapman et al. (2000a) first computed a proximity matrix for the locations, based on their discrimination among sorghum genotypes tested, and then applied principal component analysis to this matrix. Unlike the wheat dataset, the within-season rainfall data were not available for these trials. Using the long-term weather records for nearby stations, and a similar method to determine the water stress index pattern using the APSIM-Sorghum model (Hammer et al. 1996b), the long-term frequency of different seasonal patterns of drought was determined for six locations in the dataset. Across these locations, a clear negative relationship ($r = -0.904$) was established between the PCA1 scores from the analysis of trial yields, and the long-term frequency of 'mild stress' at each location. Similarly, a positive relationship was found between the PCA scores for axis 1 and the frequency of severe terminal drought stress (due to lack of rainfall from mid to late season). Lower stress seasons were more common in the southern part of the region and a different set of genotypes were favoured when compared with the higher stress seasons in the warmer more northern locations, which also had soils of a lower water holding capacity. Although the methodology was necessarily different to that used in the wheat example, the stress index integration of the model again showed that stresses around flowering were contributing to GEI for yield for the crop, and needed to be managed through sampling and/or trial management to ensure that hybrids with good potential in either type of environment were not discarded prematurely.

Model simulation of traits in multi-environment trials

'Time-slices' of crop growth

Using APSIM-Sorg (Ver 2.1), a multi-environment sorghum trial was simulated based on 30 environments (three locations by 10 sequential years 1982–1991) with 24 hypothetical genotypes that varied in four

traits, each with a range of variation that was commensurate with that observed in the local germplasm. The traits were varied in a factorial combination of maturity type (**e**arly, **m**edium, **l**ate); stay-green (**n**ormal, **s**tay-green), transpiration efficiency (**n**ormal, **h**igh) and tillering (**n**ormal, **a**bsent) and the genotypes named in four letters (one for each level of trait expression assigned). The yields from these simulations were previously analysed by Hammer et al. (1996a). They showed that the ratio of genotypic to GEI variance components was ca 0.3 and that the magnitudes of the variance components were similar to that determined from multi-environment trials conducted with sorghum in this region during these same years (Chapman et al. 2000a). Typically, to study GEI, we only have yield data at final harvest and cannot interpret how that GEI has arisen from biological interactions among component traits and their responses to environment. Here, the analyses are expanded to apply principal component analyses to traits during the season, and comparing these analyses to that for yield. This method illustrates an advantage of simulation in that the model output is daily and is retained for many of the component traits (such as biomass) that change through the season and contribute to yield.

Three stress environment types were identified by cluster analysis (method described above for wheat; for sorghum see Chapman et al. 2000b) and could be classified late terminal stress that was either severe (E82 or E85) or mild (E83, E84, E86, E89, E91, G86, W85, W84) or as terminal stress beginning prior to flowering (remaining trials) (data not shown). Similar stress types were found using an improved crop model applied to a much larger dataset of 200+ locations and 90–100 years for this region (Chapman et al. 2000c).

Applying a conventional pattern analysis (DeLacy et al. 1996), the simulated yield matrix was standardised by environment and principal component analysis was applied to produce a biplot (Fig. 3). It can be seen that the majority of the trials classified independently as late terminal stress were partitioned to the left of the biplots by the analysis of yields, while the remaining trials were partitioned to the right of the biplot. The angles between the environment vectors indicate almost zero (90°) or negative (>90°) correlations between the two types of environments. Most of the genotypes that were best adapted to the late terminal stress (LHS) were of early or medium maturity

Fig. 3 Biplot of simulated grain yield for trials over 10 years (numbers) at three locations (Gatton, Warwick and Emerald) with 24 hypothetical genotypes (see text for abbreviations)

and had 'high' expression of transpiration efficiency trait and 'normal' expression of tillering. In early and medium maturing types, genotype pairs differing in stay-green (e.g. esha versus enha) were largely adjacent in terms of yield, but were more different when expressed in a late-maturing background. Late maturing types could still perform well in environments with early, terminal drought especially if they have other water-conserving traits like high TE and reduced tillering.

From the simulation results, the total biomass at each of three stages of growth (floral initiation, flowering and maturity) was extracted and analysed in the same way as yield (Fig. 4). The PCAs below show the divergence of isoline adaptation at different 'time slices' of GxE for total biomass at floral initiation, flowering and maturity. At the initial harvest, most of the environments were strongly correlated, maturity responsible for most of the variation among genotypes. By flowering, there had been more (vertical) divergence between the effects of sub-traits within maturity types. These effects increased as some of the late terminal stress environments began to diverge from other environments in terms of genetic correlations for biomass production. While the methods were not available at the time of this work, Malosetti et al. (2006) have recently described non-linear mixed

Fig. 4 Biplots for 24 genotypes in 30 environments for simulated biomass at three stages of growth: floral initiation, flowering and maturity (from left to right)

model analyses (applied to curves of senescence in potato) that may be suitable to capture the progressions of change in variables in these datasets. As the model outputs daily data for all of the traits of interest, it may be possible to develop multi-variate methods that utilise these data-traces to explain how the final differences among genotypes arose. Appropriate use of these methods could help guide analyses of real-world data to indicate what traits need to be measured at what time of the season to identify underlying causes of GEI.

Studies of long-term multi-environment trials

In one of the examples in the first section of this paper, Chapman et al. (2000a) reported on variation in genotypic and GEI variance components over years of multi-environment trials. In the variable climate of the northern Australia grain region, simulation of such trials allows the interpretation of how weather sequences might affect the evaluation of genotypes and prediction of their future performance. Using an improved sorghum crop model, Hammer et al. (2001) and Chapman et al. (2002a) simulated the yields of 54 genotypes varying for expression of four traits in a manner similar to that of Hammer et al. (1996a), albeit with a slightly different set of trait combinations. By this time, more research had been done to characterise the physiological models of traits and the approximate range of genetic variation for trait values that existed in the breeding populations. These results over 6 locations and 108 years were analysed as a sequence of 54 2-year multi-environment trials. Figure 5 shows genotypes ranked by mean yield

across the entire simulation, having a range of 3–4.5 t ha^{-1}. However, in any given 2-year trial, the genotypes could have been ranked quite differently, simply due to variation in the impact of the weather on growth and development, i.e. excluding any confounding variation due to experimental error and biotic effects. In years like 1919 and 1939 where the rank changes were greater, the trials had poor correlations with over-all performance, associated with a high GEI variance component (Chapman et al. 2002a). Correlations between multi-environment trials and long-term performance were shown to be substantially poorer (30% were < 0.8) when only 1 year of testing was used to estimate performance (Fig. 6).

For the recent period (1980s and 1990s) of real-world trials studied by Chapman et al. (2000a) the simulations showed decreased correlations for the years 1981/2; 1984/5 and 1993/4 (Fig. 6). These all included seasons of extreme drought which had low genotypic variance and/or high GEI variance in the real world trials. The simulations show that these experiences are 'not unusual' and that there were periods of even poorer correlations during years around 1901–1903, 1913–1915 and 1931/2. Chapman et al. (2000b) showed that these years were characterised by an increased frequency of drought, according to the stress index produced by the model (data not shown). There has been much discussion to determine that: while the 1980s/1990s were a time of frequent droughts, this was not out of order for the 100-year climate record and these simulations allow the opportunity to discuss this in the context of the potential impacts on evaluation trials in breeding programmes. One option for plant breeders would be to 'weight'
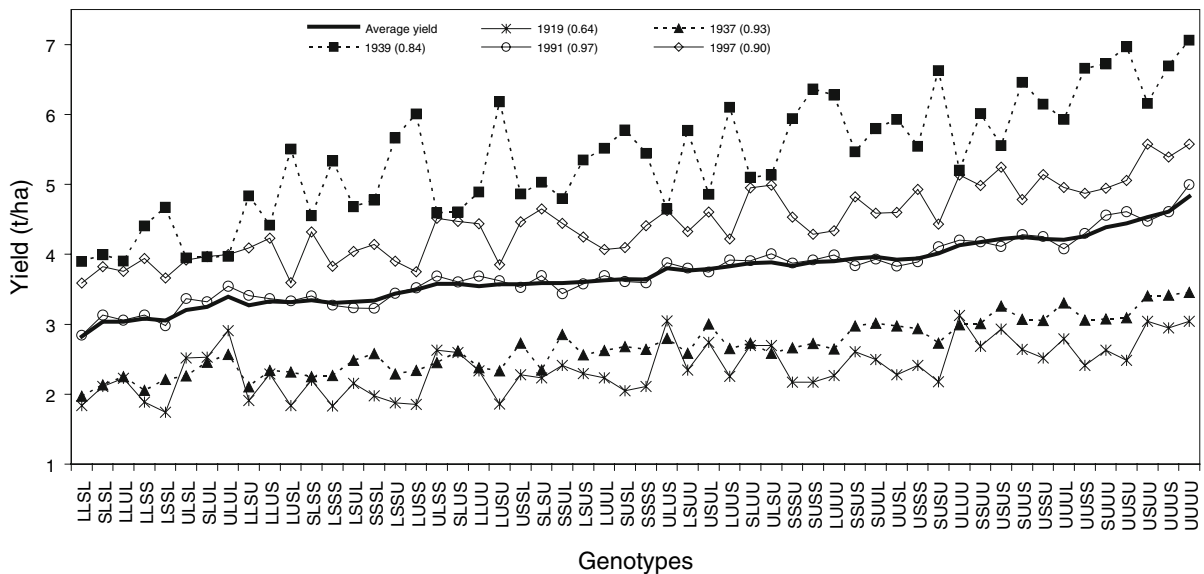
Fig. 5 Fifty-four genotypes differing in their combination of adaptive traits (names are a representation of the "dosage" of genes (L = 1, S = 2, U = 3) for each of four traits (transpiration efficiency, flowering time, stem retranslocation to grain, stay-green) are ranked by simulated mean yield over 547 trials conducted over six locations and 108 years. Other lines indicate their average yield in five 2-year by six locations multi-environment trials, and the correlation with long-term yields. (Chapman et al. 2002a) with permission

the data from each trial, depending on how representative the stress experienced is of the longer-term frequency of the stress in the target environments. Using simulation methodology, Podlich and Cooper (1998) have demonstrated for a large number of starting populations and several generations of selection, that genetic advance would be greater using this type of "weighted selection" than using the trial means only.

Changes in GEI with selection

The approaches described above to study GEI have been extended to studies of changes during selection. Standard genetic simulation models of plant breeding programmes have no direct bio-physical connection between the gene effects associated with a trait and the yield phenotype of resulting genotypes as modulated by dynamic biological interactions among traits and the abiotic environment. Typically, in a genetic simulation, the effects of these gene and physiological trait interactions on yield would normally be derived from field experiments as stochastic parameters (estimates of variance components and heritability) and from direct or approximate knowledge of the allelic effects of genes on traits and/or yield. Our approach was to define the actions of genes and their interactions with other genes (epistasis) and with environments (Gene × Environment interactions) for simulation parameters controlling the four crop traits discussed above (see Fig. 1, Chapman et al. (2003) for a framework of the modelling system).

A quantitative genetics simulation model, QU-GENE (Podlich and Cooper 1998; Cooper et al. 2002) was used to generate inbred genotypes that varied based on two alleles at each of 15 'genes' associated with the 4 traits, i.e. there were 3^{15} possible genotypes (including heterozygotes). The genes were assumed to be additive and have the same size in their effects on a trait parameter and to account for 100% of the variation. These assumptions resulted in there being only 4235 'expression states', i.e. genotypes that had the same net value for a trait parameter (once gene effects are added) have the same expression state, and would have the same phenotype in a given environment (Chapman et al. 2003). This 'granulation' of the dataset was necessary to reduce the number of simulations required to convert the genotypes to phenotypes, based on simulating the growth of each genotype in trials at six locations using 100 years of weather data. This resulted in 547 environments with yields > 0. These data were assembled into a Genotype by Environment database. From this point, we
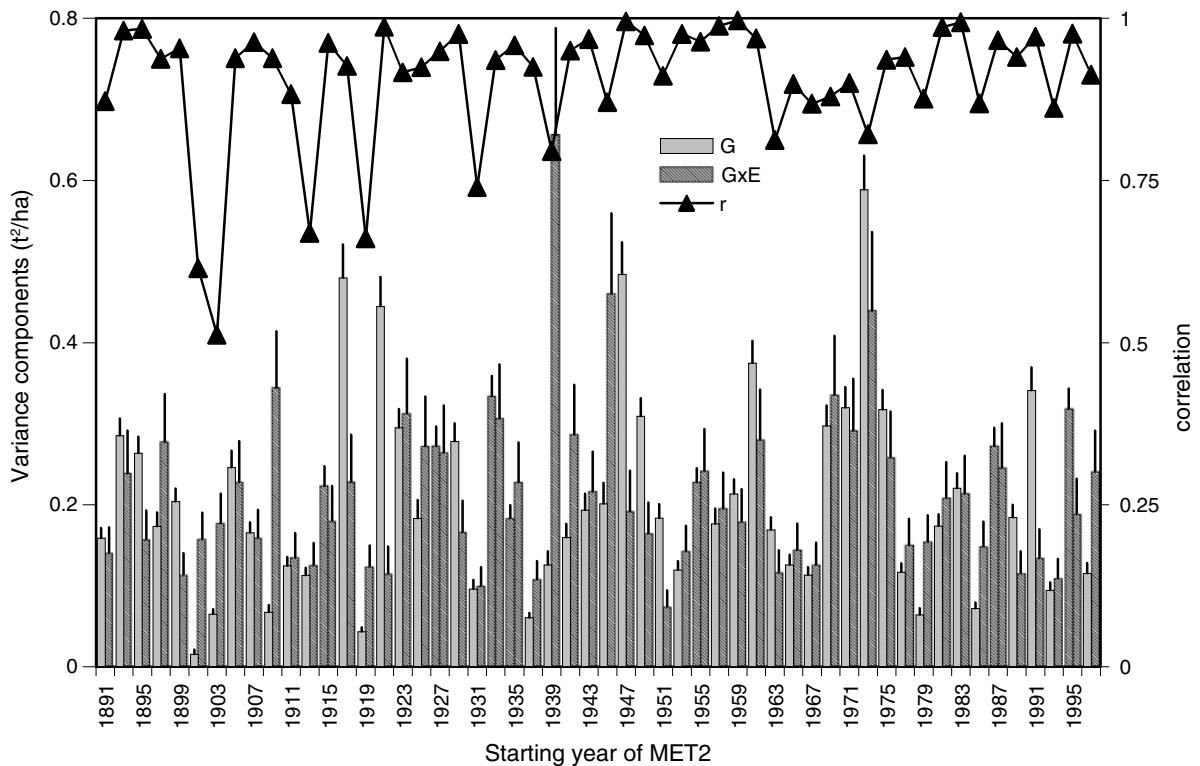
Fig. 6 For a sequence of 54 simulated 2-year METs of 54 geno-types, variance components for genotypic (G) and genotype by environment (G × E) effects and correlations (*r*) between geno-type means for the METs and the genotype averages for 547 tri-als (*n* = 54, *P* < 0.05, signif *r* = 0.27)). (Chapman et al. 2002a) with permission

consider first an analysis of the statistical variation in the entire dataset, and then of the changes in these characteristics when selection is applied to the population.

A staged analysis of the data is described below (adapted from Hammer et al. 2005) to demonstrate how additional 'knowledge' about sources of genetic and/or trait variation contributes to the understanding of GEI for a complex trait like yield. Such an approach could be used to decide how to best use additional information on gene and trait segregation to contribute to real-world analyses. Although there are several generous assumptions about gene effects and 'lack of error' in the model described, the GEI observed for yield in the results was substantial, with near equal variance components for G and GEI effects (top panel, Fig. 7), i.e. the additive effects incorporated in relating allelic variation to variables influencing response and control equations in the crop model had generated the significant GEI at the level of grain yield. An analysis of the 15 gene effects (as measured by the contribution of the positive allele) on yield

indicated a fairly similar effect on phenotypic varia-tion for yield for most of the genes, with three genes having a large effect (second panel, Fig. 7). Such a result could be expected from a powerful QTL analy-sis for the trait of yield, but does not indicate the eco-physiological basis of the associations. Next, the drought types of the environments were determined, based on simulating in each trial, a genotype that was intermediate for all trait values (see above). When the drought type information was incorporated in the analysis of gene effects, it was immediately clear that average gene effects varied substantially among envi-ronments (third panel, Fig. 7). In fact, genes with strong positive effects overall and in MTS (mild terminal stress) and MSS (mid-season stress), had negative average effects in STS (severe terminal stress) environments. Hence, the environment typing helped to resolve the effects of particular genes within a general type of environment. Finally, the associa-tion of genes with their broad physiological basis was incorporated in the analysis (bottom panel, Fig. 7). The average yield effects varied substantially among
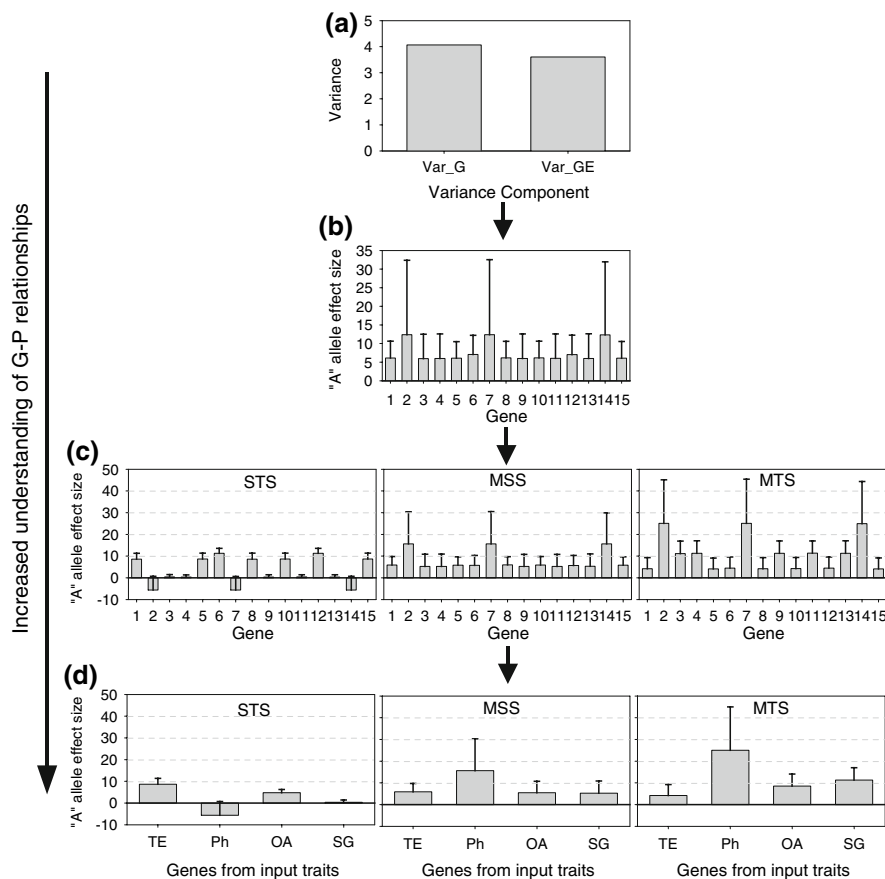
Fig. 7 Analysis of genetic and environmental effects with increasing understanding of gene-to-phenotype relationships. The top panel shows variance components analysis from conventional phenotypic analysis on yield for the whole dataset. The second panel extends this to analysis of gene effects for yield, similar to QTL analysis. The bars represent average yield effects across all environments and the line indicates the standard deviation of effect size. The third panel extends this to analysis of gene yield effects by environment type, which must be defined by simulation. The lower panel extends this by grouping gene yield effects by underpinning physiological traits, which requires enhanced knowledge of trait physiology and genetics. In all cases, gene yield effects are defined in relation to the positive allele for increasing trait value. A negative allele yield effect indicates that the opposite allele is defined as favourable for yield. Adapted from Hammer et al. (2005)

traits in the different environment types. It becomes clear that in this case all negative yield effects in STS environments are associated with late phenology, whereas major positive yield effects in that environment type are associated with TE, and OA (implemented as an effect on increased availability of stem CHOs for grain filling). Hammer et al. (2005) go on to describe how the individual gene effects varied within every environment and trait combination. Thus, there were still some environment types where the direction of the gene effect was reversed, but these were rare and the environment typing was overall quite effective in explaining GEI in terms of an interaction between trait value of different genotypes and the relative timing of drought stress. Hammer et al.

(2005) explored the implications of this as the effect of poor (i.e. unrepresentative) environment sampling on the ability of QTL studies to assist in marker-assisted genetic improvement for yield.

Given that maturity is an important part of the adaptation of sorghum to drought environments, Chapman et al. (2003) considered the comparison of three selection programmes where the phenotypic expression of maturity was used as a constraint prior to selection for yield. This would be the case where a breeder was targeting a specific maturity group for the market or expected target environment. Using QU-Gene, an initial cross was used to create a population of lines within which F2 selections were made and lines were recombined. Rounds of recurrent selection

Fig. 8 Changes in components of variance when recurrent ▶
selection was simulated for a sorghum population in northern
Australia. Phenotypic selection for yield (determined from
genotypes grown together in random environments sampled
from six locations and 100 years) was undertaken with/without
constraints on maturity 'group'. With constraints, the popula-
tion was first constrained to those with the sets of genes related
to early, medium or late flowering time, before selection for
yield was undertaken (Chapman et al. 2003). Note that data in
(**c**) are poorly estimated beyond cycle 8 due to low values of the
components

(a) Genotypic variance

(b) GxE variance

(c) Variance ratio: GxE/G

were undertaken based on the phenotypes extracted
from the GxE table. In this case, the lines were 'eval-
uated' in a random sample of five locations over two
random years from the 100+ years in the dataset. We
calculated the variance components for genotype and
GEI at each round of selection, from 200 runs of the
genetic simulation (Fig. 8). Over 10 cycles of selec-
tion, with no maturity constraints, the genotypic and
GEI components of variance began at about the same
level and decreased at about the same rate. When the
maturity of the selected group was constrained (i.e.
with initial selection on maturity—early, medium or
late prior to yield selection), there was a more rapid
decrease in the GEI if the lines were constrained to be
late rather than early or medium in maturity. This
result arose because the late lines were generally
more favourably adapted to a greater number of envi-
ronments in the dataset (Fig. 9). Hence, selection that
constrained maturity to be late decreased the GEI
more quickly by eliminating the early lines that were
likely to be less well adapted across all environments.
This would not have been a suitable strategy for the
severe terminal stress environments, where the early-
maturing lines were favoured (Chapman et al. 2003).

Conclusions

As genetic gains become more difficult to realise, new
tools are needed to understand the three interacting
factors of environments, traits and genes in plant
breeding experiments.

Crop simulation models have an increasingly
established role in dealing with the first factor by
explaining components of GEI that are observed in
plant breeding evaluation trials. Models can be used
to provide environmental indices or 'virtual' entries
that are then incorporated into mixed model or
multivariate analyses. The wheat and sorghum
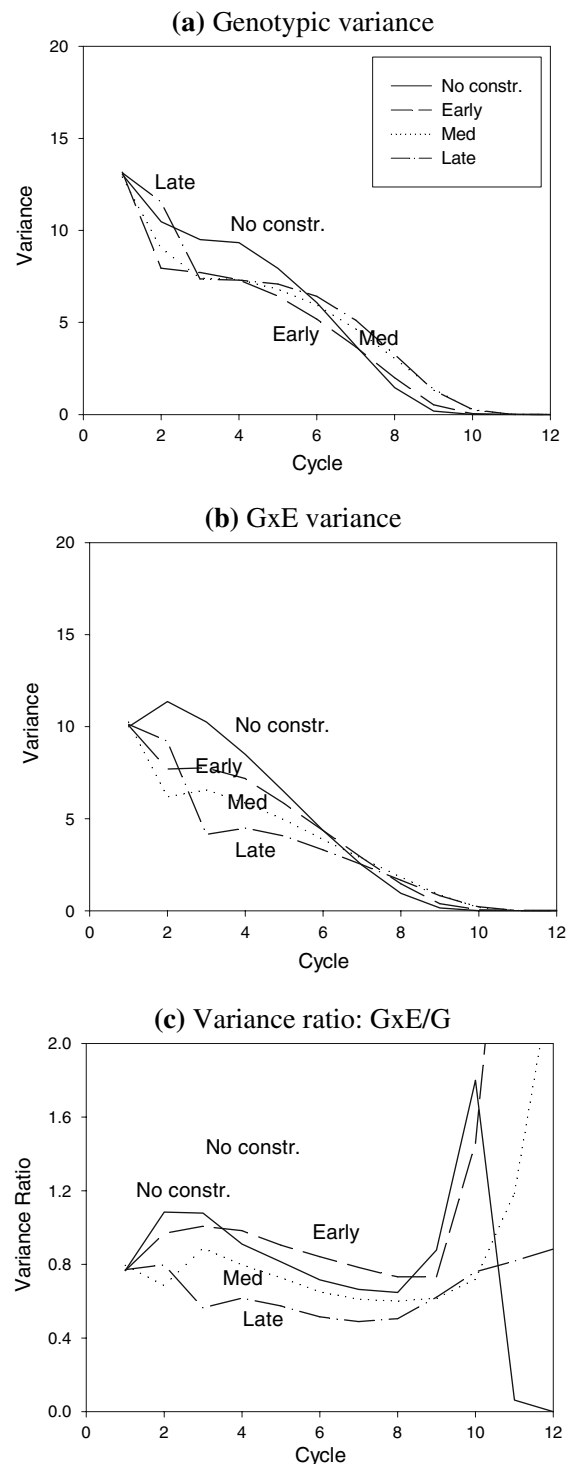
examples here show how the timing of drought, rela-
tive to different growth stages of the crop, impacts on
size of the genotype and GEI variance components
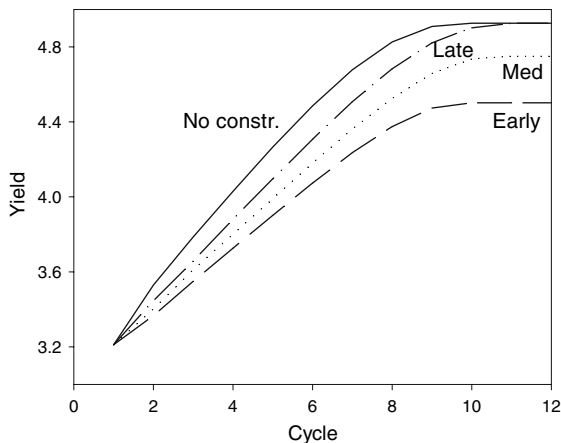and on the rankings of lines for yield. By using

Fig. 9 Effect of maturity constraints on recurrent selection when selection was undertaken using random samples across all types of environments. Adapted from Chapman et al. (2003)

classifications of drought patterns in statistical analyses, breeders can attempt to identify lines that have either broad or narrow adaptation to different drought environments. If the production system is such that these environments are identifiable as geographic regions or can be predicted using seasonal forecasting tools, then farmers are able to make use of specifically adapted lines. In a practical sense, Loffler et al. (2005) showed that these types of indices (modelled for heat in their case) could assist in the interpretation of GEI in commercial breeding programmes.

Understanding the second contributing factor to GEI (interactions among traits and the environment), requires the construction of appropriate models that capture the physiological processes of growth and predict yield. In real-world experiments, it is impossible to measure the values of all potential traits. Using output from a complete simulation of MET experiments, a statistical analysis can be tested for its ability to identify the likely best traits to measure and when to measure them with the objective of better predicting yield variation in sets of evaluated lines.

Hammer et al. (2006) discuss how the third factor, gene action information, can be linked to traits and input to crop simulation models. The simulations of METs and breeding programmes show that these tools can generate the indirect effects of trait genes on yield as mediated by crop growth and development processes and interactions with the soil and aerial environments. Such tools allow incorporation of new understanding of crop-trait-environment interactions

and the accommodation of more complex gene action relationships and selection methods. To effectively apply such tools to increase genetic gain, we need to develop statistical methods that can provide further insight into plant breeding programmes.

References

Bartels D, Furini A, Ingram J, Salamini F, Belhassen E (1996) Response of plants to dehydration stress: a molecular analysis. First INTERDROUGHT international conference 1995 Drought tolerance in higher plants: genetical, physiological and molecular biological analysis. CABI, Montpellier, pp 111–118

Bidinger FR, Hammer GL, Muchow RC (1996) The physiological basis of genotype by environment interaction in crop adaptation. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, Wallingford, pp 329–348

Campos H, Cooper A, Habben JE, Edmeades GO, Schussler JR (2004) Improving drought tolerance in maize: a view from industry. Field Crops Res 90:19–34

Chapman SC, Crossa J, Edmeades GO (1997) Genotype by environment effects and selection for drought tolerance in tropical maize. 1. Two mode pattern analysis of yield. Euphytica 95:1–9

Chapman SC, Cooper M, Butler DG, Henzell RG (2000a) Genotype by environment interactions affecting grain sorghum. I. Characteristics that confound interpretation of hybrid yield. Aust J Agric Res 51:197–207

Chapman SC, Cooper M, Hammer GL, Butler DG (2000b) Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. Aust J Agric Res 51:209–221

Chapman SC, Hammer GL, Butler DG, Cooper M (2000c) Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population of environments. Aust J Agric Res 51:223–233

Chapman SC, Mathews KL, Cooper M, Jensen N, Wang E, Butler D, Sheppard J, Sahama T (2001) Using environment characterization to interpret wheat yield in water-limited environments. In: Proceedings of the 10th wheat breeding assembly, Mildura, 16–21 September, 2001, pp 136–139

Chapman SC, Cooper M, Hammer GL (2002a) Using crop simulation to generate genotype by environment interaction

effects for sorghum in water-limited environments. Aust J Agric Res 53:379–389

Chapman SC, Hammer GL, Podlich DW, Cooper M (2002b) Linking biophysical and genetic models to integrate physiology, molecular biology and plant breeding. In: Kang MS (ed) Quantitative genetics, genomics, and plant breeding. CABI, Wallingford, pp 167–187

Chapman SC, Cooper M, Podlich D, Hammer GL (2003) Evaluating plant breeding strategies by simulating gene action and dryland environment effects. Agron J 95:99–113

Chenn K, Chapman SC, Hammer GL, McLean G, Tardien F (2008) Short term responses of leaf growth rate to water deficit scale up to whole-plant and crop levels. An integrated modelling approach in maize. Plant, Cell and Environment (in press)

Comstock RE (1977) Quantitative genetics and the design of breeding programs. In: Proceedings of the international conference on quantitative genetics. Iowa State University Press, Ames, USA, August 16–21, 1976, pp 705–718

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD (2004) Breeding for high water-use efficiency. J Exp Bot 55:2447–2460

Cooper M, Woodruff DR, Eisemann RL, Brennan PS, DeLacy IH (1995) A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat: managed-environments for selection among genotypes. Theor Appl Genet 90:492–502

Cooper M, Stucker RE, DeLacy IH, Harch BD (1997) Wheat breeding nurseries, target environments, and indirect selection for grain yield. Crop Sci 37:1168–1176

Cooper M, Podlich DW, Micallef KP, Smith OS, Jensen NM, Kruger NL (2002) Complexity, quantitative traits and plant breeding: A role for simulation modelling in the genetic improvement of crops. In: Kang MS (ed) Quantitative genetics, genomics, and plant breeding. CABI, Wallingford, pp 143–166

Cooper M, Podlich DW, Löffler CM, Van Eeuwijk F, Chapman SC (2006) Genotype-by-environment interactions under water-limited conditions. In: Ribaut J-M (ed) Drought adaptation in cereals. Food Products Press, New York, pp 51–96

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, Wallingford, pp 39–124

Edmeades GO, Bolaños J, Chapman SC, Lafitte HR, Banziger M (1999) Selection improves drought tolerance in tropical maize populations: I. Gains in biomass, grain yield, and harvest index. Crop Sci 39:1306–1315

Edmeades GO, McMaster GS, White JW, Campos H (2004) Genomics and the physiologist: bridging the gap between genes and crop response. Field Crops Res 90:5–18

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

Hammer GL, Jordan DR (2007) An integrated systems approach to crop improvement. In: Spiertz JHJ, Struik PC, van Laar HH (eds) Scale and complexity in plant systems research: gene–plant–crop relations. Springer, The Netherlands, pp 45–61

Hammer GL, Butler DG, Muchow RC, Meinke H (1996a) Integrating physiological understanding and plant breeding via crop modelling and optimization. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, Wallingford, pp 419–441

Hammer GL, Chapman SC, Muchow RC (1996b) Modelling sorghum in Australia: the state of the science and its role in the pursuit of improved practices. In: Proceedings of the third Australian sorghum conference. AIAS Occasional Publication, Tamworth, 20–22 February 1996, pp 43–61

Hammer GL, van Oosterom E, Chapman SC, Mclean G (2001) Supply and demand economics applied to crop growth. In: Borrell AK, Henzell RG (eds) Proceedings of the fourth Australian sorghum conference. CD-rom format. Range Media Pty Ltd. ISBN: 0–7242–2163–8, Kooralbyn, 5–8 February 2001

Hammer GL, Chapman S, Van Oosterom E, Podlich DW (2005) Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. Aust J Agric Res 56:947–960

Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D (2006) Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci 11:587–593

Hsiao TC (1973) Plant responses to water stress. Ann Rev Plant Physiol 24:519–570

Loffler CM, Wei J, Fast T, Gogerty J, Langton S, Bergman M, Merrill B, Cooper M (2005) Classification of maize environments using crop simulation and geographic information systems. Crop Sci 45:1708–1716

Ludlow MM, Muchow RC (1990) A critical evaluation of traits for improved crop yields in water-limited environments. Adv Agron 43:107–153

Malosetti M, Visser RGF, Celis-Gamboa C, van Eeuwijk FA (2006) QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato. Theor Appl Genet 113:288–300

Mathews KL, Chapman SC, Butler D, Cooper M, DeLacy IH, Sheppard J, Sahama T (2002) Inter-annual changes in genotypic and genotype by environment variance components for different stages of the Northern Wheat Improvement Program. In: McComb JA (ed) Plant breeding for the 11th millennium proceedings 12th Australasian plant breeding conference. Perth, 15–20 September 2002, pp 650–654

Podlich D, Cooper M (1998) QU-GENE: a simulation platform for quantitative analysis of genetic models. Bioinformatics 14:632–653

Qiao CG, Basford KE, DeLacy IH, Cooper M (2000) Evaluation of experimental designs and spatial analyses in wheat breeding trials. Theor Appl Genet 100:9–16

Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiol 131:664–675

Rizza F, Badeck FW, Cattivelli L, Lidestri O, Di Fonzo N, Stanca AM (2004) Use of a water stress index to identify barley genotypes adapted to rainfed and irrigated conditions. Crop Sci 44:2127–2137

Saulescu NN, Kronstad WE (1995) Growth simulation outputs for detection of differential cultivar response to environmental factors. Crop Sci 35:773–778

Thornley JHM, France J (2004) Mathematical models in agriculture: quantitative methods for the plant, animal and ecological sciences, 2nd edn. CABI Pub, Wallingford

Voltas J, Van Eeuwijk FA, Araus JL, Romagosa I (1999) Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley II. Grain growth. Field Crops Res 62:75–84

Wang E, Robertson MJ, Hammer GL, Carberry PS, Holzworth D, Meinke H, Chapman SC, Hargreaves JNG, Huth NI, McLean G (2002) Development of a generic crop model template in the cropping system model APSIM. Eur J Agronomy. Elsevier Science B.V., Amsterdam Netherlands, pp 121–140