Field Crops Research 304 (2023) 109133

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Incorporating environmental covariates to explore genotype × environment × management (G × E × M) interactions: A one-stage predictive model

Michael H. Mumford [a,*], Clayton R. Forknall [a], Daniel Rodriguez [b], Joseph X. Eyre [b], Alison M. Kelly [c]

[a] Department of Agriculture and Fisheries, Leslie Research Facility, Toowoomba, Queensland 4350, Australia
[b] Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, Gatton Campus, Qld 4343, Australia
[c] Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, Hermitage Research Facility, Qld 4370, Australia

ARTICLE INFO

Keywords:
Linear mixed model
REML
Forward selection
Cross validation
Multi-environment trial
Agronomy
Sorghum

ABSTRACT

Context: Evaluating the genotype (G) by management practice (M) interaction in agronomic experimentation is essential to help grain growers optimise the desired trait of interest (e.g. grain yield). However, the approach is complicated by interaction effects with environmental factors that differ across sites and seasons. Popular statistical methods for modelling the genotype by environment (G × E) interaction are limited as they neither provide a biological understanding of how environmental factors impact on the G × E interaction, nor assess how different management practices influence the G × E interaction. These limitations may be addressed by incorporating environmental covariates (ECs) into the modelling process to better explain why differences exist in the optimal genotype by management practice combination across environments.
Objective: A novel statistical methodology is proposed that incorporates ECs to explore genotype by environment by management practice (G × E × M) interactions in agronomic multi-environment trial studies.
Methods: A predictive linear mixed model is proposed that incorporates site and season specific ECs into a commonly used G × E interaction framework. The model is extended to include the effect of continuously varying agronomic management practices, whilst allowing for non-linear trait responses and complex variance structures. The methodology is applied to a multi-environment dataset exploring yield response to established plant density in a series of sorghum agronomy trials.
Results: Results indicated that the grain yield of sorghum genotypes would be optimised in environments that have (i) high total plant available water and photo-thermal quotient around flowering, (ii) low pre-flowering radiation and evapotranspiration and (iii) achieved flowering at an optimal time. Under this set of optimal G × E conditions, a high established plant density further optimised grain yield.
Conclusions: The proposed methodology successfully incorporated ECs to better understand G × E × M interactions in agronomic field trials, enabling predictions to be made in an untested or future environment and linking the statistical analysis to crop-ecophysiology principles.
Implications: This work will improve the generalisations agronomists can draw from experimental studies, enhancing the biological understanding of the analysis results and allowing for the development of more targeted and robust recommendations for agronomic practices.

1. Introduction

Optimising the management (M) of genotypes (G) in an agricultural production system is essential to maximise grain yields and reduce production risks. However, consistently achieving high yield is difficult due to uncontrollable environmental factors and the differing effect of these factors on each genotype (Rotili et al., 2020). It is well established that genotypic performance is strongly influenced by the environment (E), resulting in the common practice of assessing genotype performance in trials across different seasons and in different geographic locations

* Corresponding author.
E-mail address: Michael.Mumford@daf.qld.gov.au (M.H. Mumford).

https://doi.org/10.1016/j.fcr.2023.109133
Received 29 March 2023; Received in revised form 13 August 2023; Accepted 13 September 2023
Available online 21 October 2023

(Cooper and Hammer, 1996). A series of trials conducted over multiple years and trial sites is collectively referred to as a multi-environment trial (MET) series, where the aim is to assess the genotype by environment (G × E) interaction. In the coming decades, it is anticipated that environmental drivers contributing to G × E interaction will become more extreme due to ongoing climate change (Hatfield and Walthall, 2015), making it even more crucial to better understand the environmental drivers in these G × E studies.

Existing statistical methods for modelling G × E interactions are typically formulated in a linear mixed model (LMM) framework. For datasets in which large numbers of genotypes are tested, a multiplicative framework in the form of a factor analytic (FA) structure is used to model the G × E interaction effects (Piepho, 1998; Smith et al., 2001; Cullis et al., 2010). In terms of the E component, the models are limited by the observed covariance between sampled environments. Therefore, these models neither seamlessly allow for prediction at a 'new' untested environment, nor provide an understanding of the environmental drivers influencing the G × E interaction pattern (Heslot et al., 2014). For the G component, the FA models have performed well for large numbers of genotypes (Kelly et al., 2007), but may not be appropriate for datasets in which a small number of genotypes are tested. This is because estimates of complex interactions through the FA form may be unreliable (Macdonald, 2018).

Predictive models for the G × E interaction can be formed explicitly through the inclusion of environmental drivers, also commonly referred to as environmental covariates (ECs). We have defined an EC as a measured environmental parameter that may have an impact on the overall environmental conditions, and subsequently the phenotypic trait being investigated for the crop. An EC may be associated with the weather (e.g. precipitation, radiation, air temperature), the soil (e.g. soil type, soil fertility level) or both (e.g. soil temperature). By assessing the interaction of ECs with genotypes, the model can account for G × E interaction effects in a biologically meaningful manner. This will be denoted as a genotype × environmental covariate (G × EC) interaction to differentiate from the definition of a G × E interaction.

A number of methods have been proposed in the statistical literature to incorporate ECs when analysing MET data for large numbers of genotypes. Factorial regression (Denis, 1988; Van Eeuwijk, 1996) uses an analysis of variance (ANOVA) framework that partitions the sums of squares of the G × E interaction effect into a 'lack of fit' G × E sums of squares term and a G × EC sums of squares term for each EC. The advantage of the ANOVA framework is that it can partition the G × E (and G × EC) effects from blocking terms that are defined by the randomisation of genotypes to experimental units. This results in improved precision and reduced bias of the estimated G × E interaction effects whilst ensuring the correct numerator and denominator degrees of freedom are used for hypothesis testing. However, these early regression models are limited by their inability to account for extraneous spatial field trend and heterogeneity of residual variance across trials that have been shown to be important for MET data (Gilmour et al., 1997; Smith et al., 2001). Furthermore, these regression models are restricted to polynomial regression terms for non-linear trait responses to ECs, and thus, may not always provide an adequate fit to the data.

Extensions of factorial regression to a LMM framework have also been proposed (Malosetti et al., 2004; Boer et al., 2007). These extensions provide many of the advantages of the LMM including the flexibility to model variance heterogeneity and spatial field trend. However, both Malosetti et al. (2004) and Boer et al. (2007) assumed linear responses to ECs, and this is a limitation (Boer et al., 2007). It would be a useful generalisation if the non-linear trait response to ECs could be "expressed with respect to a spline basis" in factorial regression (Van Eeuwijk et al., 2019).

In the work presented by Oliveira et al. (2020), the environmental loadings from a FA model (fit to capture the G × E interaction effects in a LMM framework) were correlated with ECs measured across environments. This exploratory approach is an extension of the methodology that used additive main effects and multiplicative interaction (AMMI) models in an ANOVA framework (Perkins, 1972; Van Eeuwijk and Elgersma, 1993; Vargas et al., 1999). However, the approach proposed by Oliveira et al. (2020) does not incorporate the ECs into the modelling process and, for this reason, does not allow for the prediction of genotype performance in an untested environment.

A reaction norm model for genotype selection via the use of environmental kinship matrices to capture quantitative trait loci (QTL) by EC (QTL × EC) interactions was proposed by Jarquín et al. (2014). This approach has been applied in a number of recent studies in both crop and animal science (Malosetti et al., 2016; Tiezzi et al., 2017; Krause et al., 2019). One limitation acknowledged by Jarquín et al. (2014) is that, when using the reaction norm approach, "the ECs may not fully describe differences across environments, perhaps because some relevant ECs were not measured or because of model misspecification (e.g. non-linear effects of ECs on the trait of interest)". To account for this, 'lack-of-fit' terms for the E main effect and the QTL × E interaction effect are also included in the Jarquín et al. (2014) model, where E represents any environment effects that are not accounted for by the EC effects, including non-linear EC effects. A further limitation of the approach proposed in Jarquín et al. (2014) is that it does not allow for the modelling of complex variance structures including heterogeneous residual variation, spatial field trend and blocking terms, unless a two-stage approach is implemented.

In studies that seek to obtain G × E predictions, a two-stage approach occurs when genotype means for each environment are first obtained from an analysis of individual trials (first stage), then the predicted means from the individual trial analyses are carried over to a LMM framework to model G × E (or G × EC) effects across trials in the second stage (Möhring and Piepho, 2009). The main weakness of using a two-stage approach to obtain G × E predictions as compared to a one-stage analysis is that there is a loss of information when estimating the non-genetic variance components in a two-stage model (Gogel et al., 2018). This limitation can be minimised if a weighted two-stage approach is employed such that environments with a larger genetic variance are given more weight and vice versa (Piepho et al., 2012). In a one-stage approach, individual plot data is included in a mixed model formulation, and has been shown to be the most accurate model for obtaining predictions of G × E effects (Welham et al., 2010).

Recently, Tolhurst et al. (2022) proposed a one-stage statistical model that extends the FA approach of Smith et al. (2001) to simultaneously incorporate observed ECs and 'latent' ECs into the MET analysis of G × E data. The 'latent' ECs in Tolhurst et al. (2022) are analogous to the 'lack-of-fit' term imposed in Denis (1988) and Jarquín et al. (2014). The approach proposed by Tolhurst et al. (2022) is appealing, as it simultaneously accounts for experimental design terms, spatial field trend and heterogeneity of genetic and residual variance, all whilst doing so using a one-stage approach. The incorporation of ECs in Tolhurst et al. (2022) allows for the potential to make predictions in a new untested environment which is not possible in Smith et al. (2001). One limitation is that the methodology in Tolhurst et al. (2022) is only applicable when the number of ECs is less than (or equal to) the total number of environments. Another limitation is that the methodology in Tolhurst et al. (2022) assumes a linear relationship between the trait of interest and an EC within each factor of the FA model.

An effective way to capture the non-linear trait response to an EC is through the use of smoothing splines (Craven and Wahba, 1978; Silverman, 1985). Parameter estimates from the cubic smoothing spline can be represented in the form of a best linear unbiased predictor (BLUP, Speed, 1991). This allows the natural cubic smoothing spline to be formulated within a LMM framework (Verbyla et al., 1999). Smoothing splines can be extended to capture the potential non-linear trait response to combinations of two (or more) continuous variables simultaneously using tensor products (Wood, 2006; Lee et al., 2013). Recently, the tensor cubic smoothing spline was also formulated within a LMM framework whilst simultaneously allowing for complex variance

structures at the residual level (Verbyla et al., 2018).

All of the statistical methods that incorporate ECs that have been mentioned thus far have been formulated within a plant improvement context. When working in agronomic research, the impact of management practice (M) on genotypic performance is also a key objective (Rodriguez et al., 2018). Hence, statistical models need to accommodate unique combinations of E and M, resulting in a three-way interaction effect for genotype by environment by management practice (G × E × M). Additionally, in agronomic research, trials are mostly comprised of small numbers of genotypes, in an attempt to understand genotypic adaptation, as well as the interaction effect between genotype and management practice (G × M). Hence, the methodology developed for the genetic assessment of large numbers of genotypes may not be directly transferable to agronomic research.

A number of studies have focused on the importance of G × E × M interactions (Rotili et al., 2020; Rodriguez et al., 2018; Clarke et al., 2019; Cooper et al., 2020; Kirkegaard and Hunt, 2010; Hammer et al., 2020), but these have not been extended to the more general and flexible framework that includes ECs. Moreover, the studies focusing on G × E × M interactions have all used simulation studies via crop growth models, and to our knowledge, there has been no attempt to explain the presence of G × E × M interaction from experimental work independent of assumptions about the functional relationships between ECs and the trait of interest. Research that incorporates ECs outside of any functional relationship assumptions have turned to exploratory data analysis, which is susceptible to confounding between G, E, and M effects, confounding between spatial field trend with G, E or M effects, imprecision, and multi-collinearity of ECs.

This paper presents a novel statistical model that incorporates ECs to better explain the G × E × M interaction effect for MET data in an agronomic setting with small numbers of genotypes. The proposed method is a one-stage approach formulated in a LMM framework and uses subset selection to determine the ECs that are contributing most to the G × E × M interaction. The proposed methodology can accommodate non-linear responses to ECs and their interaction effects. The focus is on a single response variable, with a motivating example used to demonstrate the model and highlight how predictions can be obtained for a continuous explanatory variable related to management practices.

2. Materials and Methods

The dataset used to demonstrate the application of the proposed statistical methodology consists of a series of six sorghum agronomy trials conducted across five highly contrasting sites (Breeza, Moree, Surat, Warra, Emerald) of the northern grain growing region of eastern Australia in the 2018–19 growing season (Table 1).

The factorial combination of (i) time of sowing (TOS), (ii) genotype (G) and (iii) target plant density (M) was assessed at each trial. The number of levels of the factors varied between trials with two or three times of sowing, six to nine genotypes, but always with four target plant densities (3, 6, 9 and 12 plants/m^2). Details of the varying TOS and genotype levels are given in Table 1. The proposed statistical methodology was applied using the eight genotypes present at five out of the six trials.

A split-split plot design was employed in four of the trials (Breeza 1, Breeza 2, Moree and Emerald) with three replicate blocks, such that TOS was randomly allocated to main-plots. In two of these trials (Breeza 2 and Emerald), target plant density levels were randomly assigned to sub-plots nested within main-plots and genotypes were randomly allocated to individual plots nested within sub-plots. In the other two trials where a split-split plot design was used (Breeza 1 and Moree), genotypes were randomly allocated to sub-plots nested within main-plots, and target plant densities were randomly assigned to individual plots nested within sub-plots.

For the remaining two trials (Surat and Warra), a split-plot design was employed, with three replicate blocks, such that TOS was randomly allocated to main-plots and the factorial combination of target plant density and genotype was randomly allocated to individual plots within main-plots.

All trials were managed so that nutrient limitations were eliminated by the use of fertilisers. Pests and diseases were prevented by using chemical controls, and all trials were sown at one metre row spacings using precision sowing technologies.

The response variable of interest is grain yield (t/ha) which was determined from hand-harvested areas and adjusted to 0% moisture. Establishment counts were also taken to ascertain the established number of plants in each plot. Since the target plant density is often different from the established plant density observed in the field, the established plant density covariate is often favoured in the analysis of such experiments, as it is a more accurate indicator of the true contribution of plant density to the trait of interest (e.g. grain yield). Moreover, this allows for the trait of interest to be estimated across an observed domain of established plant densities.

2.1. Definition of environment

In this study, an 'environment' is defined as a set of differing growing conditions (e.g. climate, soil characteristics) under which genotypes or management practices are tested, and these growing conditions can be induced by geographic location or TOS. Factors contributing to the environment term include the geographic location (i.e. site) and TOS, as crops sown at different times would be exposed to contrasting climate and soil conditions. Conversely, a management practice is defined as the set of variables that are manipulated under the same 'growing' conditions.

In this dataset, the combination of trial and TOS resulted in a total of 17 environments. Management practice manipulations arise from the differing plant density targets. The final established plant density in each trial was used as a continuous measured covariate, replacing target plant density in the analysis.

An untested environment is defined as an environment where no phenotypic information has been collected, and inferences for the untested environment are made using a combination of phenotypic and EC data collected at the 'tested' environments. An untested environment may be a new location, a future or different season or TOS, or a combination of any or all of these factors.

2.2. Environmental covariates

Weather data for each of the six trials was obtained from weather

Table 1
Summary of the sowing time (TOS), number of genotypes, experimental design and geographic location for each sorghum agronomy trial in the motivating dataset.

Trial	TOS 1	TOS 2	TOS 3	Number of Genotypes	Experimental Design	Latitude	Longitude
Breeza 1	6/09/2018	17/09/2018	23/10/2018	6	Split-split plot	− 31.18	150.42
Breeza 2	3/09/2018	18/09/2018	16/10/2018	8	Split-split plot	− 31.18	150.42
Moree	8/08/2018	12/09/2018	27/09/2018	8	Split-split plot	− 28.96	150.06
Surat	8/08/2018	28/08/2018	24/01/2019	9	Split-plot	− 27.16	149.06
Warra	27/07/2018	19/10/2018	9/11/2018	9	Split-plot	− 26.82	150.83
Emerald	26/07/2018	16/08/2018		8	Split-split plot	− 23.54	148.18

stations within each trial and consisted of information on rainfall, air temperature, air relative humidity, incoming solar radiation, and soil temperature. Calculated environmental covariates included potential evapotranspiration (ETo), and the photothermal quotient (PTq, Rodriguez et al., 2014). For each genotype, phenology data was measured for days to (i) emergence, (ii) flowering, and (iii) maturity. By combining phenology and weather data, a total of 18 ECs were derived for each environment and 17 of these ECs varied for each genotype in each environment. A summary of the ECs is provided in Table 2.

2.3. Baseline statistical model

A baseline model to capture the G × E × M interaction effects was fitted prior to incorporating ECs into the model. The mathematical formulation of the baseline model is provided in the appendix, while a symbolic representation of the model and its constituent terms is provided here. Using the notation of Wilkinson and Rogers (1973), the baseline model can be represented as:

$$\begin{aligned} \texttt{fixed} &= \texttt{1} + \texttt{Genotype} + \texttt{Management} \\ &\quad + \texttt{Genotype:Management}, \\ \texttt{random} &= \texttt{Env} + \texttt{Genotype:Env} + \\ &\quad \texttt{Genotype:Env:Management} \\ &\quad + \texttt{Design} + \texttt{Error}, \end{aligned} \tag{1}$$

where `1` denotes the overall mean, `Env` is the environment term, `Design` represents the blocking terms pertaining to the experimental designs within each trial, and `Error` denotes the residual variance. A separate residual variance is fitted for each environment to improve the goodness of model fit.

For the motivating dataset, the baseline model in symbolic form is as follows:

$$\begin{aligned} \texttt{fixed} &= \texttt{1} + \texttt{density} + \texttt{Genotype} + \texttt{density:Genotype}, \\ \texttt{random} &= \texttt{Trial} + \texttt{Env} + \texttt{Trial:Genotype} + \texttt{Env:Genotype} \\ &\quad + \texttt{density:Trial:Genotype} \\ &\quad + \texttt{density:Env:Genotype} \\ &\quad + \texttt{spl(density)} + \texttt{spl(density):Genotype} \\ &\quad + \texttt{at(Trial):(Rep/MainPlot/SubPlot)} + \texttt{Error}. \end{aligned} \tag{2}$$

In (1), `Management` could be a continuous variable or a factor. In the motivating dataset, the M component is represented by established plant density, which is denoted as `density` and fit as a continuous variable. Any terms in Wilkinson and Rogers (1973) notation which begin with a lower case letter represents a continuous variable, and any term that begins with a capital letter denotes a factor. In the motivating example, the baseline model captures the trait response to both environment and plant density via a random regression approach (Laird and Ware, 1982). The non-linear trait response to plant density was captured via the LMM representation of the natural cubic smoothing spline (Verbyla et al., 1999). The term `spl(.)` is a random spline component that models the smooth non-linear trait response to the term within `spl()` (Verbyla et al., 1999).

For the motivating example, a `Trial` term is also required in the model to ensure that the correct strata for the experimental design terms are recognised (Bailey, 2008). This ensures that the model fits a nested structure of sowing times within each trial. Note that, in symbolic form, `Env = Trial:TOS` and thus TOS is implicitly included in the model. There is no `TOS` main effect in the model because sowing times between trials can differ significantly. For example, the last TOS for the trial in Moree was the 27/09/2018 whilst for Surat it was the 24/01/2019 (see Table 1).

In (2), the `Design` terms are represented by `at(Trial):(Rep/MainPlot/SubPlot)` which fits separate variance components arising from the split-plot or split-split plot designs for each trial. This, in combination with the `Env` and `Trial` terms, enables the baseline model to respect the experimental design implemented at each trial.

It is important that the statistical model is translation invariant to ensure that the parameter estimates from the model are consistent regardless of the scale of the response variable (Wood, 2017). In random coefficient regression, a covariance parameter between the linear regression's intercept and slope variance components is required to ensure that the model is translation invariant (De Faveri et al., 2015; Forknall et al., 2019). To ensure that the baseline model for the motivating example was invariant to a change of spline basis (White et al., 1998), the spline coefficients were incorporated into the model such that the non-linear trait response to plant density differed with respect to genotype only and not environment.

2.4. Incorporating environmental covariates into the model

The baseline model (1) can be expanded to incorporate a single EC which, using the notation of Wilkinson and Rogers (1973), can be

Table 2
Summary statistics of the 18 environmental covariates incorporated into the multi-environment trial analysis.

Environmental covariate (EC)	Acronym	Mean	Min	Max	Observational unit for EC
Pre-flowering plant available water* (mm)	PrePAW	202	11	382	G × E
Post-flowering plant available water* (mm)	PostPAW	73	0	245	G × E
Initial soil water (mm)	ISW	177	97	280	Environment
Pre-flowering maximum temperature (°C)	PreMaxT	30.6	26.0	35.7	G × E
Post-flowering maximum temperature (°C)	PostMaxT	34.8	27.9	39.7	G × E
Pre-flowering minimum temperature (°C)	PreMinT	14.5	9.3	22.3	G × E
Post-flowering minimum temperature (°C)	PostMinT	18.8	11.8	22.4	G × E
Pre-flowering radiation (MJ/day)	PreFlwRad	20.5	17.0	24.0	G × E
Post-flowering radiation (MJ/day)	PostFlwRad	23.0	13.9	28.8	G × E
Pre-flowering cumulative solar radiation (MJ)	PreCumRad	1602	948	2196	G × E
Post-flowering cumulative solar radiation (MJ)	PostCumRad	995	586	1685	G × E
Pre-flowering evapotranspiration (mm)	PreFlwEvap	486	291	625	G × E
Post-flowering evapotranspiration (mm)	PostFlwEvap	336	181	582	G × E
Mean photo-thermal quotient (PTq)** around flowering (MJ/°C day)	ptq	0.89	0.68	1.00	G × E
Normalised PTq (NPTq) around flowering (MJ.kPa/°C day)	NPTq	0.52	0.28	0.79	G × E
Mean soil temperature (7 cm depth) during 7 days after sowing (°C)	AvgSoilTmp	18.5	12.6	29.2	G × E
Water deficit index (mm)***	Deficit	545	171	811	G × E
Potential seed set (%)****	SeedSet	74.4	19.8	99.4	G × E

* Is equal to the rainfall + irrigation
** For the calculation of NPTq, see Rodriguez et al. (2014)
*** Water deficit index calculated as in Hammer et al. (2014)
**** Potential seed set was calculated as in Singh et al. (2015)

written symbolically as:

$$\begin{aligned} \texttt{fixed} &= \texttt{1} + \texttt{ec} + \texttt{Genotype} + \texttt{Management} \\ &\quad + \texttt{Genotype:Management} + \texttt{ec} \\ &\quad + \texttt{ec:Genotype} + \texttt{ec:Management} \\ &\quad + \texttt{ec:Genotype:Management}, \\ \texttt{random} &= \texttt{Env} + \texttt{Genotype:Env} \\ &\quad + \texttt{Genotype:Env:Management} + \texttt{spl(ec)} \\ &\quad + \texttt{spl(ec):Genotype} + \texttt{spl(ec):Management} \\ &\quad + \texttt{spl(ec):Genotype:Management} \\ &\quad + \texttt{Design} + \texttt{Error}. \end{aligned} \tag{3}$$

The LMM representation of the natural cubic smoothing spline requires the EC term in (3) to be modelled using a linear and non-linear component. This also includes the corresponding interaction effects with G, M and G × M. Since a separate residual variance is fitted to each environment, more weight is given to environments with smaller residual variation (Patterson and Silvey, 1980; Crossa, 1990) when assessing the impact of an EC. Note that an EC could be either a continuous variable or a factor. If an EC is a factor, then (3) simplifies as all the spline terms for that EC would not need to be fitted in the model. For the motivating data, all the ECs are continuous.

The motivating data with the addition of a single EC can be written in Wilkinson and Rogers (1973) notation as:

$$\begin{aligned} \texttt{fixed} &= \texttt{1+density+ec+Genotype} \\ &\quad \texttt{+density:Genotype+density:ec} \\ &\quad \texttt{+ec:Genotype+density:ec:Genotype}, \\ \texttt{random} &= \texttt{Trial+Env+Trial:Genotype} \\ &\quad \texttt{+Env:Genotype+density:Trial:Genotype} \\ &\quad \texttt{+density:Env:Genotype} \\ &\quad \texttt{+spl(density)+spl(density):Genotype} \\ &\quad \texttt{+spl(ec)+spl(ec):Genotype} \\ &\quad \texttt{+spl(density):ec+spl(density):ec:Genotype} \\ &\quad \texttt{+density:spl(ec)+density:spl(ec):Genotype} \\ &\quad \texttt{+spl(density):spl(ec)} \\ &\quad \texttt{+spl(density):spl(ec):Genotype} \\ &\quad \texttt{+at(Trial):(Rep/MainPlot/SubPlot)+Error}. \end{aligned} \tag{4}$$

The interaction effect between plant density and an EC is captured through tensor cubic smoothing splines (Verbyla et al., 2018), allowing for a three-dimensional surface to be fitted for the trait response to plant density and an EC simultaneously.

The model in (3) can be expanded to include multiple ECs in the model. The interaction effect between pair-wise combinations of ECs was not considered here. Thus, the inclusion of multiple ECs into the model is in an additive form. The mathematical form of the EC model is provided in the appendix.

2.5. Forward selection procedure

A forward selection procedure is used for incorporating ECs into the model and identifying the most important ECs contributing to the G × E × M interaction at each iteration. Firstly, each EC is incorporated individually via the full model (3). Then, for each EC, a *k*-fold cross validation scheme is implemented. For the motivating dataset, a leave-one-trial-out cross validation scheme was performed such that one trial was removed from the MET dataset, and then the full model for the EC being considered in the current iteration was refitted. Predictions for each individual plot were obtained for the 'missing' trial using information from the EC that is currently being fitted in the model. This process is repeated another five times (one for each trial) until a full set of predictions for a 'quasi-untested environment' are obtained for each plot at all environments, denoted as y^*_i for the i^{th} plot.

The mean squared error of prediction (MSEP) and the root-MSEP (RMSEP) are then calculated as:

$$\begin{aligned} \text{MSEP} &= \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{baseline} - y_i^*\right)^2 \\ \text{RMSEP} &= \sqrt{\text{MSEP}}, \end{aligned} \tag{5}$$

where y_i^{baseline} is the predicted value for the i^{th} plot obtained from the baseline model. The calculation of RMSEP via cross validation is an effective way to assess predictive accuracy that simultaneously takes precision and bias into account (Hastie et al., 2009). It is better to use y_i^{baseline} instead of y_i which is more commonplace in cross-validation schemes (see for example Equation (4.1) in Montesinos López et al., 2022) since y_i^{baseline} has been adjusted for spatial field trend within each trial, providing a more accurate baseline to compare with. This process is repeated for each EC and the EC that minimises the RMSEP is identified as being the most important EC from the forward selection procedure.

2.6. Backward selection procedure

Once an important EC is identified via the forward selection procedure, backwards selection is performed on the full model in (B.5) to remove any non-significant EC terms (main effect and corresponding interaction effects) to achieve a parsimonious model. For the motivating data, the backwards selection procedure starts with the non-linear three-way interaction effects (between an EC, established plant density and genotype) and then works backwards to test the non-linear main effects following the principle of marginality (Nelder, 1965). For the random spline terms, the Akaike information criterion (AIC, Akaike, 1973) is derived using the full log-likelihood at the REML parameter estimates (Verbyla, 2019) to determine which splines terms should be retained in the final model. For fixed effect terms, Wald tests with a conditional F-statistic (Kenward and Roger, 1997) were used for significance testing, allowing for adjustments in the denominator degrees of freedom depending on the observational unit that each EC was measured at (Table 2). Following the principle of marginality, only linear terms without a corresponding spline term can be tested and thus removed from the final model.

Once a parsimonious model is identified for a single EC, the RMSEP is recalculated via *k*-fold cross validation to ensure that the predictive performance of the single EC in an untested environment is better than the baseline model (or the current EC model being considered, see Fig. 1). If the RMSEP is lower, the 'current model' is updated to incorporate a single EC. The subset selection procedure is then repeated until a subset of important ECs are identified. The stopping criterion for incorporating ECs is when the RMSEP is no longer reduced by the inclusion of additional ECs. A flow chart summarising the subset selection procedure is provided in Fig. 1. The final model is an extension of the baseline model that consists of each of the key ECs identified such that all of the terms corresponding to each of the key ECs that are non-significant are omitted from the final model.

2.7. Software and implementation

All model parameters were estimated using residual maximum likelihood (REML, Patterson and Thompson, 1971) via the `ASRemlR` package (Butler et al., 2017) in the `R` software environment (R Core Team, 2022). Predictions for the trait response to an EC are empirical best linear unbiased estimators (eBLUEs) when the trait response to an EC is linear and empirical best linear unbiased predictors (eBLUPs) when there is a significant non-linear trait response to an EC.

For the motivating example, all spline terms were set to have six knot points to ensure that the subset selection procedure could be completed within a reasonable time frame (<one hour to identify an important EC via the subset selection procedure for the motivating data), whilst having enough knot points to accurately capture potential non-linear

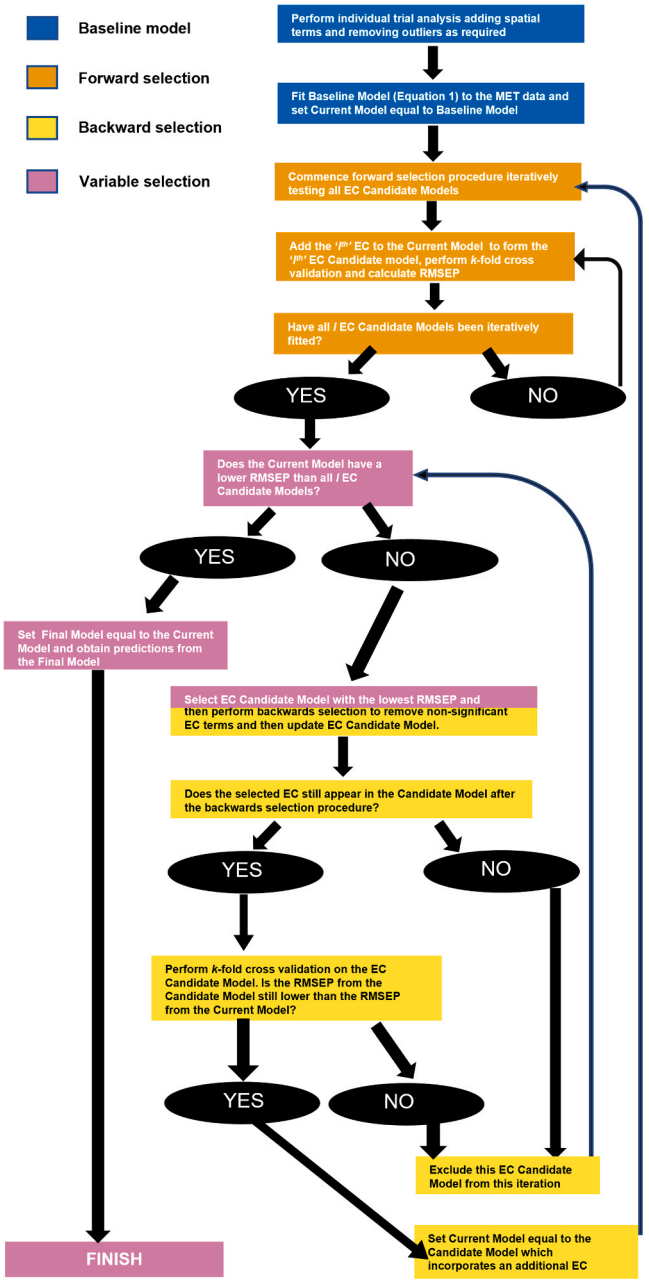

Fig. 1. A flow chart summarising the subset selection procedure implemented to multi-environment trial data to identify the most important environmental covariates contributing to the genotype × environment × management practice interaction.

trends. Once the final model was implemented, the number of knot points was updated to be the default value in `ASReml-R`, which is the minimum of 50 and the number of unique values for the EC term. All ECs in the motivating data were zero-centred prior to commencing the subset selection procedure. The `R`-script to implement the (i) baseline and (ii) EC models using the `ASReml-R` syntax for the motivating dataset are provided in a GitHub repository accessible using the following link: https://github.com/michaelhm-daf/ModellingGxExMwithECs.

3. Results

3.1. Baseline model

In the baseline model for the motivating data, the spline (i.e. non-linear) variance components for (i) established plant density and (ii) the non-linear interaction effect between established plant density and genotype were positive. There was also a positive variance component for the linear three-way interaction between established plant density, genotype and environment, providing evidence of a G × E × M interaction effect.

The variance components for the `Trial` terms were always larger than the equivalent variance components for `Env`. This suggests that there was more environmental variation due to environment differences across trials as opposed to differences in sowing times within a trial.

An additional step was taken to include residual variance heterogeneity for sowing times within trials, as visual diagnostics indicated that some sowing times had more residual variation than others. In most cases where residual variance heterogeneity across sowing times within a trial occurred, the earliest sowing time had more residual variation, which is hypothesised to be due to earlier sowing times having more variation in emergence date, resulting in less uniformity in response to environmental stressors compared to later, more conventional sowing times.

3.2. Environmental covariate model

Post-flowering plant available water was the first EC identified when implementing the forward selection procedure (Table 3). After post-flowering plant available water was incorporated into the model, the AIC (Verbyla, 2019) was calculated for numerous random effect models to determine the significance of the spline terms for the (i) post-flowering plant available water main effect, (ii) interaction effect with established plant density, (iii) interaction with genotype, and (iv) three-way interaction with established plant density and genotype, by working backwards. This resulted in all spline terms being non-significant except for the terms `spl(density):PostPAW`, `spl(PostPAW):Genotype` and `density:spl(PostPAW)` which were retained in the next iteration of the subset selection procedure (Table 4).

The EC fixed effects were then assessed via Wald tests (Kenward and Roger, 1997). The term for the linear three way interaction effect `density:PostPAW:Genotype` was non-significant and was thus dropped from the model (Table 5). The term `PostPAW:Genotype` was statistically significant and hence was retained in the model. The term `density:PostPAW` was not statistically significant, however, the spline term for the interaction effect between established plant density and post-flowering plant available water was significant (Table 4) and thus the corresponding linear term was retained in the model. Note that all terms aligned with the interaction effect between established plant density and post-flowering plant available water later became non-significant once additional ECs were included in the model, indicating that the E × M interaction was better captured by other ECs.

The subset selection procedure was then repeated to identify any additional important ECs. A summary of the EC terms in the final model with respect to genotype and plant density is provided in Table 6. The final model consisted of seven ECs that were identified as explaining a significant proportion of the G × E × M variation (Table 6). After identifying seven key ECs, the forward selection procedure was unable to detect any ECs that further reduced the RMSEP (Fig. 1).

For each EC, the interaction effect between genotype, plant density and the corresponding EC was explored. Predictions from the final model for the (a) genotype × post-flowering plant available water and

Table 3
Summary of the root mean square error of prediction (RMSEP, as defined in (5)) for each environmental covariate (EC) during the first iteration of the forward selection procedure applied to the sorghum multi-environment trial data. Post-flowering plant available water was selected as the first EC since its inclusion in the model resulted in the lowest RMSEP.

Environmental covariate (EC)	Acronym	Full log-likelihood	RMSEP (t/ha)
Post-flowering plant available water (mm)*	PostPAW	− 621.74	**2.03**
Post-flowering maximum temperature (°C)	PostMaxT	− 601.64	2.30
Water deficit index (mm)***	Deficit	− 616.98	2.31
Post-flowering minimum temperature (°C)	PostMinT	− 605.80	2.35
Pre-flowering radiation (MJ/day)	PreFlwRad	− 629.23	2.43
Pre-flowering minimum temperature (°C)	PreMinT	− 610.89	2.71
Pre-flowering evapotranspiration (mm)	PreFlwEvap	− 636.52	2.71
Post-flowering evapotranspiration (mm)	PostFlwEvap	− 644.02	2.71
Potential seed set (%)****	SeedSet	− 637.40	2.71
Pre-flowering maximum temperature (°C)	PreMaxT	− 621.29	2.72
Pre-flowering cumulative solar radiation (MJ)	PostCumRad	− 639.82	2.76
Initial soil water (mm)	ISW	− 623.00	2.78
Pre-flowering cumulative solar radiation (MJ)	PreCumRad	− 617.06	2.90
Pre-flowering plant available water (mm)*	PrePAW	− 617.81	2.91
Mean soil temperature (7 cm depth) during 7 days after sowing (°C)	AvgSoilTmp	− 605.87	2.93
Irrigation (pre-sowing) (mm)	Irrig	− 646.27	2.97
Mean photo-thermal quotient (PTq) around flowering (MJ/°C day)	ptq	− 603.46	3.03
Post-flowering radiation (MJ/day)	PostFlwRad	− 605.85	3.09
Normalised PTq (NPTq)** around flowering (MJ.kPa/°C day)	NPTq	− 629.31	3.21

* Is equal to the rainfall + irrigation
** For the calculation of NPTq, see Rodriguez et al. (2014)
*** Water deficit index calculated as in Hammer et al. (2014)
**** Potential seed set was calculated as in Singh et al. (2015)

Table 4
Summary of the spline models fitted for the environmental covariate post-flowering plant avaliable water.

Model	Full log-likelihood	*p*	*q*	*b*	AIC
Full environmental covariate (EC) model	− 617.87	32	37	9	1373.74
− `spl(density):spl(PostPAW):Genotype`	− 617.87	32	37	8	1373.74
− `spl(density):spl(PostPAW)`	− 617.89	32	36	8	1371.80
− `spl(density):PostPAW:Genotype`	− 617.88	32	36	7	1371.77
− **density:spl(PostPAW):Genotype**	− 617.88	32	36	6	**1371.77**
− `spl(density):PostPAW`*	− 619.50	32	35	6	1373.00
− `spl(PostPAW):Genotype`*	− 622.74	32	34	7	1377.47
− `density:spl(PostPAW)`*	− 619.80	32	35	6	1373.60

The characters *p*, *q* and *b* denote the total number of fixed parameters, variance parameters and boundary variance parameters respectively. The spline terms were sequentially tested using a backwards selection procedure starting with the full environmental covariate (EC) model, where the corresponding model term was dropped if the Akaike information criterion (AIC) value was higher. The AIC was derived using the full log-likelihood at the residual maximum likelihood parameter estimates (Verbyla, 2019). The minus sign (−) preceding each model term denotes that the term was removed from the model in the preceding row of the table prior to refitting. The * symbol indicates a term that was identified as important as per the AIC criterion and was therefore retained in the succeeding models. The bolded model indicates the spline model identified as the best fit.

Table 5
Conditional Wald tests for the post-flowering plant available water fixed effect terms after the non-significant spline terms associated with post-flowering plant available water were removed from the model. The denominator degrees of freedom were calculated using the methodology proposed in Kenward and Roger (1997).

Model term	Numerator degrees of freedom	Denominator degrees of freedom	Conditional *F*-statistic	*P*-value
`density`	1	34.1	33.9	< 0.001
`PostPAW`	1	83.7	6.7	0.011
`Genotype`	7	33.1	4.3	0.002
`density:Genotype`	7	22.8	0.4	0.860
`PostPAW:Genotype`	7	48.8	3.6	< 0.001
`density:PostPAW`	1	15.8	2.2	0.154
`density:PostPAW:Genotype`	7	82.0	0.1	0.997

(b) established plant density × post-flowering plant available water interaction effects are presented in Figs. 2 and 3 respectively. Predictions for the (a) G × EC and (b) M × EC interaction effects for the remaining ECs included in the final model can be seen in Supplementary Figs. A.1-A.6.

The predictions from the final model are displayed in Fig. 4 for a subset of two genotypes and two trials (Moree and Surat) at the earliest and latest sowing times.

Table 6
Model terms for the environmental covariates included in the final model.

Model term	Post-flowering plant available water (mm)	Initial soil water (mm)	Pre-flowering cumulative radiation (MJ)	Pre-flowering plant available water (mm)	Photo thermal quotient around flowering (MJ/°C day)	Pre-flowering evapo-transpiration (mm)	Post-flowering maximum temperature (°C)
`ec`	✓	✓	✓	✓	✓	✓	✓
`Genotype:ec`	✓				✓		
`density:ec`		✓	✓	✓		✓	✓
`spl(ec)`	✓				✓		✓
`spl(ec):Genotype`	✓				✓		
`density:spl(ec)`							✓
`spl(density):ec`		✓				✓	

The term `spl()` denotes a spline term in the model. The corresponding spline terms were included in the final model if their inclusion minimised the Akaike information criterion (AIC). Linear terms were included in the final model if the term was statistically significant using a Wald test with an approximate F-statistic. Terms that do not appear in the final model for any of the environmental covariates were excluded from the list of model terms. The environmental covariate terms are presented in order of inclusion (left to right) as identified by the forward selection procedure.

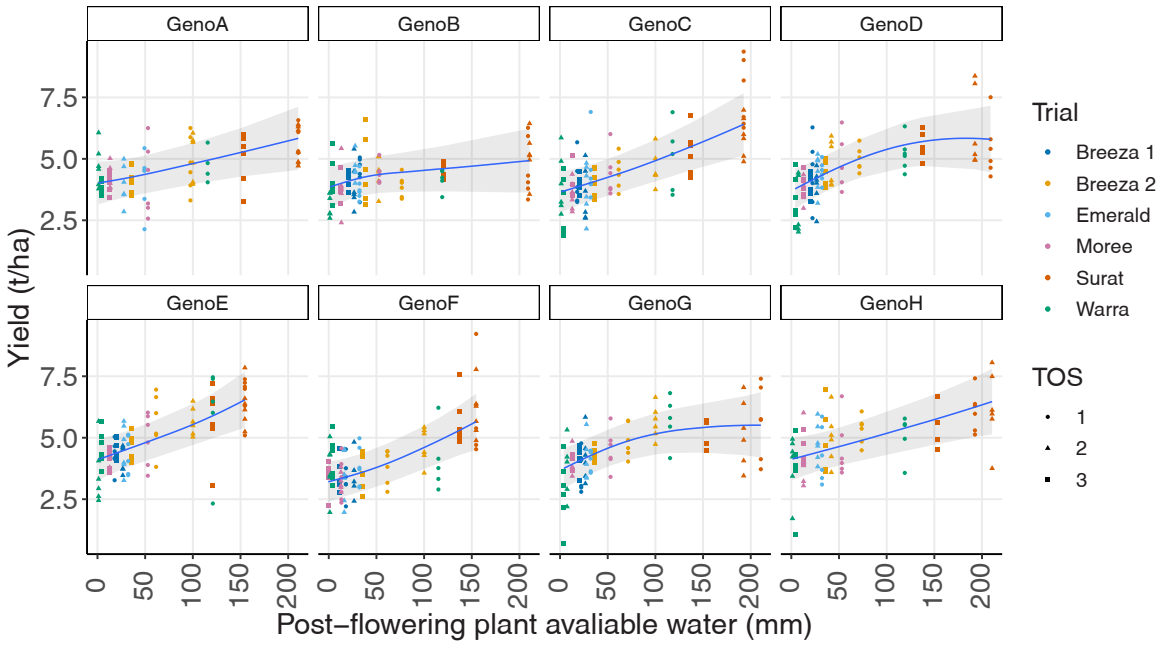

Fig. 2. Yield predictions for the genotype × post-flowering plant available water interaction (empirical best linear unbiased predictors). The final model consisted of a non-linear genotype × post-flowering plant available water interaction term. The raw data points are from field plots that targeted a plant density of 6 and 9 plants/m^2 and are adjusted for all other environmental covariates and 'lack of fit' effects. The shaded regions denote the 95% prediction interval for each genotype. Predictions were taken at the average plant density of 6.6 plants/m^2 observed across all environments.

4. Discussion

A new methodology is proposed which is the first to incorporate ECs into the analysis of MET data arising from agronomic field trials featuring G × E × M interaction effects for small numbers of genotypes (< 10). The results from the motivating example in Fig. 3 and Supplementary Figs. A.1-A.6 highlight how this methodology enables simultaneous modelling of the yield response to established plant density and the identified important ECs respectively. All previous research into the development of methodologies to improve predictive performance using ECs has focused on plant improvement programs. The proposed methodology allows agronomic researchers to identify key ECs contributing to the G × E × M interaction on measured experimental trial data, independent of any assumptions made via a crop growth model.

A novel aspect of the proposed methodology is that it facilitates modelling the non-linear trait response to an EC. This feature is highlighted in Fig. 2 which captures the non-linear yield response to post-flowering plant avaliable water. Moreover, the methodology can capture non-linear interaction effects between M and an EC. An example of this is the non-linear interaction effect between established plant density and post-flowering maximum temperature (Supplementary Fig. A.6). These figures summarise the ECs in a desirable format similar to that seen in Fig. 4 of Van Eeuwijk et al. (2019), without requiring any underlying assumptions of the relationship between traits. The most recently developed methods that incorporate ECs to explain G × E interactions (see for example Boer et al., 2007; Jarquín et al., 2014; Hadasch et al., 2020) all assume a linear or quadratic trait response to an EC. The proposed methodology uses the LMM formulation of the (tensor) cubic smoothing spline (Verbyla et al., 2018) to capture non-linear trait responses to ECs in a flexible and objective manner, independent of any underlying biophysical assumptions.

The incorporation of smoothing splines (Verbyla et al., 1999) also enables the EC model terms to be partitioned into linear and non-linear components (De Faveri et al., 2022), allowing for significance testing of these components to be performed independently (Tables 5 and 4 respectively). In the motivating example, two of the seven ECs (pre-flowering cumulative radiation and pre-flowering plant available water) had significant linear effects exclusively whilst the remaining five ECs had at least one significant non-linear interaction term with either genotype or plant density (Table 6). This ensures that the interpretation

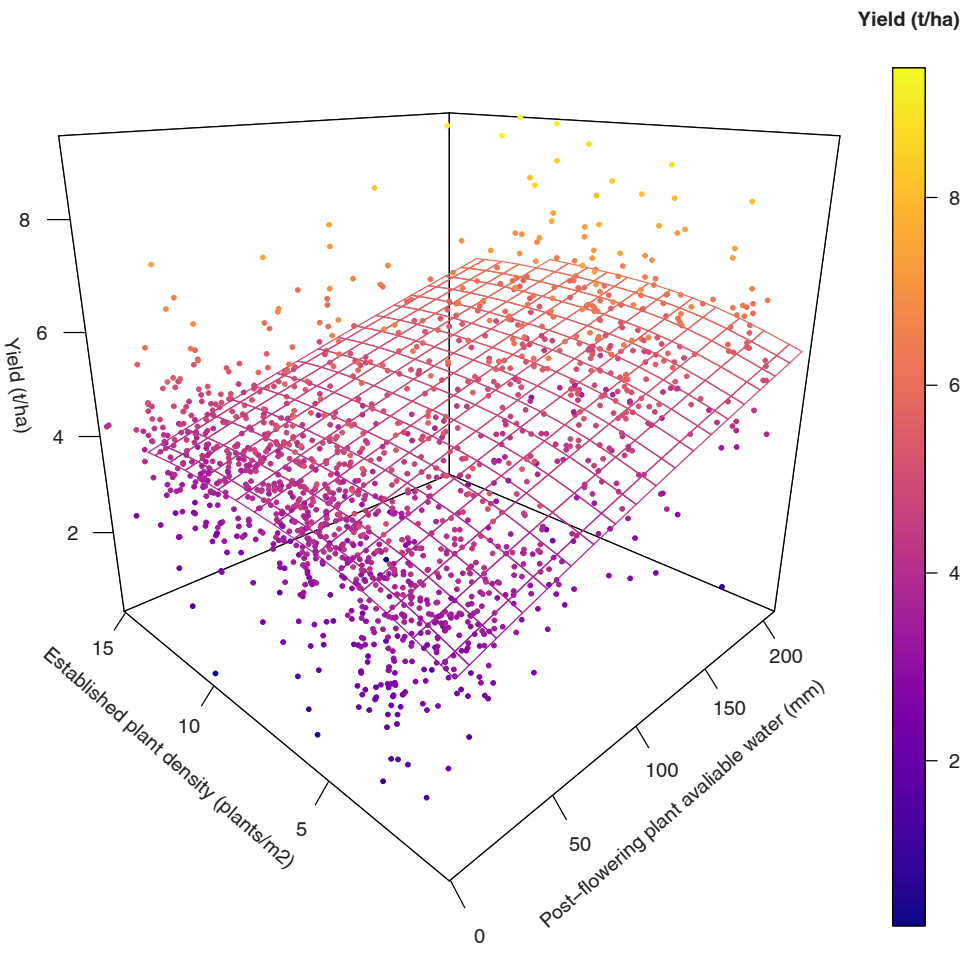

Fig. 3. Three dimensional surface highlighting how the predictions of grain yield change with respect to established plant density and post-flowering plant available water simultaneously. Predictions are averaged across all eight genotypes and the other additional environmental covariates included in the final model. There were no terms in the final model for the interaction effect between established plant density and post-flowering plant available water. The data points in the figure are grain yield values adjusted for all other environmental covariates and 'lack of fit' effects.

of results is succinct and practically useful when the trait response to an EC or management practice is linear.

Moreover, the methodology can also partition the G × EC × M interaction into a linear and non-linear component, allowing for a unique three dimensional surface for each genotype when there is a significant G × EC × M interaction effect (see for example Fig. 3). Note that there were no significant G × EC × M interaction effects in the motivating data for each of the key ECs identified and thus the shape of the yield response curve in Fig. 3 is representative of all genotypes. When there is a significant G × EC × M interaction effect, the relationship between the trait of interest, M and an EC for each genotype can be summarised graphically.

Interaction effects for the linear and non-linear EC terms with genotype and management practice can also be obtained, providing accurate information on how each EC interacts with genotype and management. For the motivating data, two of the important ECs, post-flowering plant available water and photo thermal quotient around flowering, were identified as having an interaction effect with genotype but not established plant density. Conversely, the remaining five ECs were determined to have an interaction effect with established plant density but not with genotype (Table 6). When all of the important ECs are used to obtain predictions for the overall yield response to established plant density, the collective output observed is a G × E × M interaction (Fig. 4). By incorporating ECs into the statistical model, the methodology can partition complex G × E × M interaction effects into the corresponding G × EC and EC × M interaction effects that comprise of the G × E × M interaction effect.

The proposed methodology also allows modelling of complex variance structures arising from experimental design terms, spatial field trend and residual variance heterogeneity across environments. The inclusion of these additional terms in the mixed model formulation is known to improve predictive accuracy of G × E interaction effects in plant improvement studies (Cullis et al., 1998). Thus, the proposed methodology provides accuracy improvements in the predictions obtained from G × E × M MET data compared to methodologies that incorporate ECs but do not account for complex variance structures (De Faveri et al., 2022). Using a methodology that provides the most accurate predictions for G × E × M effects in the baseline model will result in a higher likelihood of correctly identifying the key ECs contributing to the G × E × M interaction.

The proposed methodology is implemented using a one-stage approach, which ensures that variation across all strata (e.g. between and within replicate blocks) are taken into consideration when obtaining predictions for the G × E × M interaction effects. This results in further improvements in predictive accuracy (Welham et al., 2010) compared to most current statistical methodologies incorporating ECs that either (a) do not account for complex variance structures, or (b) account for complex variance structures, but do so via a two-stage approach (Gogel et al., 2018). Furthermore, the use of a one-stage approach ensures that the standard errors of the predictions are not underestimated which can occur in two stage approaches that incorporate ECs. By using a one-stage approach, predictive accuracy is further

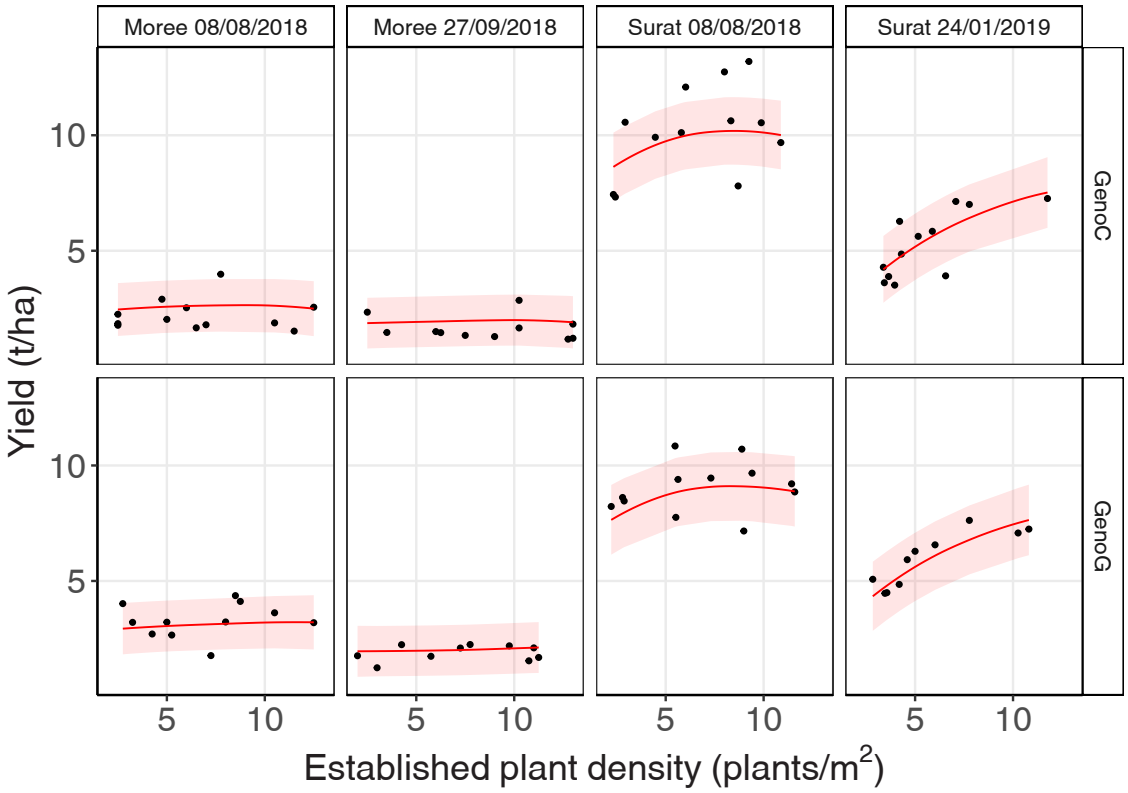

Fig. 4. Predictions of yield response to established plant density for a subset of two genotypes and four environments. The four environments selected come from the combination of two trials with high and low yield at the earliest and latest sowing times. The regression lines denote the predictions obtained for a tested genotype in a tested environment. The shaded regions denote the corresponding 95% prediction intervals. Black dot points in the figure represent raw yield data adjusted for experimental design and extraneous field trend effects identified in the baseline model.

improved whilst ensuring that no information regarding variation at different strata is lost.

As well as being statistically rigorous, it is also important that the methodology produces results that can be interpreted and presented as succinctly as possible. For the motivating example, the yield response to an EC can be clearly summarised and displayed graphically, including interaction effects between genotype and an EC (Fig. 2 and Supplementary Fig. A.4). Similarly, the yield response to established plant density and an EC simultaneously can be presented using a three dimensional plot (for example, see Fig. 3 and Supplementary Figs. A.1 and A.2). Summarising this information, the results indicated that grain yield performance of sorghum genotypes would be optimised in environments that have (i) high total plant available water and photothermal quotient around flowering, (ii) low pre-flowering radiation and evapotranspiration and (iii) achieved flowering at an optimal time. Under this set of optimal G $\times$ E conditions, a high established plant density further optimised grain yield. The proposed methodology can turn complex statistical models into clear, concise and practically useful information for researchers, industry, and the broader agricultural production system, especially under the challenges of a changing and volatile climate.

One of the most exciting aspects of the proposed methodology is the potential to use the key ECs identified to make predictions in a future or untested environment. If the environmental data pertaining to the key ECs is available in an untested environment, then assessment of the model's ability to predict in an untested environment is possible using the proposed methodology. It is also possible to obtain predictions for an untested management practice when management is a continuous variable by interpolation. For the motivating dataset, it was possible to predict for any value of established plant density within the domain of plant densities observed across all environments, which was between 1 and 14 plants/m^2. It is important not to extrapolate beyond this domain (e.g. $>$ 14 plants/m^2 in the motivating data) without additional experimental data collected for established plant densities $>$ 14 plants/m^2. This is the first methodology that the authors are aware of that allows for predictions of the trait response to management practice in an untested environment to be obtained using EC information.

It is important to note that the confounding of ECs is unavoidable in agricultural field trials. Thus, some ECs will inevitably be highly correlated with other ECs. The forward selection procedure in the proposed methodology is designed to minimise the impact of multicollinearity, ensuring that two ECs that are highly correlated do not both appear in the final model.

Whilst a significant EC may be identified during the forward selection procedure as a good fit to the data, that EC may not necessarily be describing a causal relationship with the trait of interest. By incorporating a larger number of environments into the MET analysis, it is expected that the key ECs identified will be more likely to be causal. This is analogous to incorporating a greater number of genotypes (that show sufficient genetic diversity) in a genome wide association study (Meuwissen et al., 2001) which typically require hundreds of genotypes in order to make valid inferences about the population. Identifying causal ECs may not be as much of a concern if the objective is to obtain accurate predictions in an untested or future environment. To improve confidence that the key ECs identified are causal, it is important to include as many environments in the MET analysis as possible spanning the range of possible environmental conditions within the target population of environments (TPE).

When compared to a machine learning approach, one advantage of the proposed methodology is that it provides meaningful information on how the predictions were obtained from a tested or untested environment. For example, in the motivating dataset, the information about the interaction effects between an EC with genotype or plant density (Figs. 2 and 3 and Supplementary Figs. A.1–A.6) were combined to obtain the

predictions of the yield response to plant density for each environment (Fig. 4). A benefit of this in practice is that it allows researchers to directly evaluate whether the yield response to an EC compares to existing eco-physiological frameworks. If a model includes an EC such that the trait response to that EC does not appear to be biologically feasible, then there is the option to re-prioritise that EC in practice and continue with the subset selection procedure described in Fig. 4.

It is also important that the sample of environments in the MET analysis are representative of the range of possible environmental conditions within the TPE. The motivating example contained 17 unique environments which is small when attempting to make inference about the TPE. Including additional environments for the motivating example that span the diverse range of environmental conditions possible within the TPE would further improve confidence that the identified ECs are causal, improving predictive performance in a future or untested environment.

It is difficult to conduct a large number of field trials for a MET analysis. This is because conducting a large number of field trials is expensive and time consuming. Thus, it is of the utmost importance to develop statistical methods that can generalise inference to untested environments whilst minimising the number of tested environments required to identify important ECs. One way this is achieved is by developing a rigorous statistical methodology that maximises the accuracy and precision of the predictions obtained from the model. Another important way to achieve this is to ensure that there is no confounding between E (and thus ECs) with G or M. This is accomplished by ensuring that the same genotypes and the same management practices are applied at all trials included in the MET analysis with ECs incorporated. In order to adopt the methodology proposed in this paper, it is vital that the field trials conducted remove as much confounding between E and M as possible to ensure that differences in M are not incorrectly accounted for by the ECs included in the final model. This is rarely taken into consideration in field crop studies that incorporate ECs and is one of the reasons why prediction in a future or untested environment is so difficult. Therefore, when the aim is to identify key ECs contributing to the G $\times$ E $\times$ M interaction, it is vital to ensure that management practice remains consistent across environments.

One limitation of the proposed methodology is that it does not consider the possibility of interaction effects between pair-wise combinations of ECs. This is because with a small number of environments (17) in the motivating example, exploring interaction effects between two (or more) ECs resulted in severe over-fitting. With a large number of environments constituting the dataset, it may be possible to explore interaction effects between pair-wise combinations of ECs. The model in (3) could also be generalised to include multiple management practices such as row spacing and nitrogen levels as long as the experiments are designed with the appropriate factor levels of M. The proposed methodology may be able to capture interaction effects between pair-wise combinations of ECs with more environments included in the analysis, as well as multiple management practices if they are present in all environments.

It is also important to obtain as much phenology data as possible when incorporating ECs into the MET analysis. This would allow for ECs such as rainfall, temperature, and radiation to be partitioned into key phenology periods, allowing for a better understanding of when ECs such as rainfall and temperature have the largest impact on the trait of interest. By matching phenology with daily weather information, most of the ECs in the motivating example had unique values for each genotype, allowing for improved partitioning of trait differences into their respective G, E and M components. If phenology data is available at the field plot stratum, then the ECs can be further partitioned to distinguish EC effects from noise as well as E and M effects. Most current methodologies are unable to accommodate for ECs that are measured at different strata, as the bulk of proposed methods in the literature only allow for one unique EC value for each environment. Not only does the data in the motivating example have different flowering days for each genotype (see for example Fig. 2), but the proposed methodology also accounts for the different strata in the Wald test by approximating the denominator degrees of freedom (Kenward and Roger, 1997). Therefore, it is important that a methodology that incorporates ECs can accommodate ECs measured at different scales in the experiment.

A key difference between our model and most G $\times$ E models in plant improvement studies is the decision to fit genotype as fixed and environment (as well as the G $\times$ E interaction term) as random. In a plant improvement program, it is more common to see environment fitted as fixed and genotype as random (Smith et al., 2005) which ties in with quantitative genetics theory (Falconer and Mackay, 1996). However, Smith et al. (2005) also states that the decision to fit genotype as fixed or random is situation dependent. When incorporating ECs into the statistical model, it is assumed that the environments in the analysis are a subset of the TPE (Piepho, 1998). Moreover, once ECs are incorporated into the model, all environment terms take on the role of 'lack of fit' terms in the LMM. Therefore, when incorporating ECs into the model, it makes sense to fit genotype as fixed and environment as random.

With some modifications, it should be possible to apply the methodology proposed in this paper to plant improvement studies. This is currently an area of ongoing research. Once developed, it will be important to compare the proposed methodology with other methodologies that incorporate ECs in the plant improvement literature to assess and determine the best methodology for obtaining predictions in a future or untested environment.

5. Conclusion

A new methodology is proposed to incorporate ECs into a MET analysis which is applicable to G $\times$ E $\times$ M data and provides an accurate estimate of the contribution of each EC to the trait of interest. The methodology accounts for complex sources of variation such as experimental design terms, spatial field trend and residual variance heterogeneity for each environment. Moreover, the proposed methodology is completed via a one-stage approach and allows ECs to be measured at different strata. This study is the first step towards developing one-stage statistical models that identify key environmental drivers of G $\times$ E $\times$ M interactions, enhancing the biological understanding of the experimental results and allowing for the development of more targeted and robust recommendations for agronomic practices.

CRediT authorship contribution statement

Michael H. Mumford: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Clayton R. Forknall**: Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration. **Daniel Rodriguez**: Investigation, Data Curation, Writing – review & editing, Project administration, Funding acquisition. **Joseph X. Eyre**: Investigation, Data Curation, Writing – review & editing. **Alison M. Kelly**: Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors have no competing interests to declare.

Data Availability

The data/code to run the analysis are provided in a GitHub repository accessible in the following link: https://github.com/michaelhm-daf/ModellingGxExMwithECs. .

Acknowledgements

The authors would like to acknowledge the Grains Research

Development Corporation for the funding of sorghum agronomy research as well as the Statistics for the Australian Grains Industry North (SAGI-North) project. The authors also acknowledge the Queensland Department of Agriculture and Fisheries for financial support. We thank the technical staff that conducted each of the trials presented in this manuscript. Finally, we would like to thank each of the reviewers for their constructive feedback which has resulted in an improved final version of the manuscript.

Appendix A. Supporting information

Supplementary figures associated with this article can be found in the online version at doi:10.1016/j.fcr.2023.109133.

Appendix B. Mathematical form of the models

B.1. Mathematical form of the baseline model

It is assumed that there are a total of v genotypes, t environments, and that the data is ordered as environments nested within trial, and then genotypes within environments. Under these assumptions, the general form of the baseline LMM for the i^{th} genotype in the j^{th} environment for the motivating example data can be written as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_{\mathrm{g}}\boldsymbol{u}_{\mathrm{g}} + \boldsymbol{Z}_{\mathrm{m}}\boldsymbol{u}_{\mathrm{m}} + \boldsymbol{Z}_{\mathrm{o}}\boldsymbol{u}_{\mathrm{o}} + \boldsymbol{\epsilon}, \tag{B.1}$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of measurements for each plot ($n = \sum_{j=1}^{t} n_j$), $\boldsymbol{\beta} = [\boldsymbol{\beta}'_0\ \boldsymbol{\beta}'_1]'$ is a $2v \times 1$ vector of fixed effects such that the sub-vectors $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$ are $v \times 1$ vectors of regression intercept and slope coefficients for v genotypes. The design matrix $\boldsymbol{X} = [\boldsymbol{X}_0\ \boldsymbol{X}_1]$, is of dimension $n \times 2v$ such that the $n \times v$ indicator matrix $\boldsymbol{X}_0 = [\boldsymbol{X}'_{0_1}\ \boldsymbol{X}'_{0_2}\ \ldots\ \boldsymbol{X}'_{0_t}]'$ can be further partitioned into sub-matrices $\boldsymbol{X}_{0_j}$ for each environment (dimension $n_j \times v$) with elements equal to one if genotype i was sown in each plot and zero otherwise. If all genotypes are present in all environments then

$$\boldsymbol{X}_{0_j} = \oplus_{i=1}^{v} \boldsymbol{I}_{n_{ji}},$$

where $\boldsymbol{I}_{n_{ji}}$ is a $n_{ji} \times 1$ vector of ones. If genotype i was not sown in environment j, then the column vector $\boldsymbol{I}_{n_{ji}}$ is replaced with a column vector of zero length.

Similarly, the matrix $\boldsymbol{X}_1 = [\boldsymbol{X}'_{1_1}\ \boldsymbol{X}'_{1_2}\ \ldots\ \boldsymbol{X}'_{1_t}]'$ of dimension $n \times v$ can also be partitioned into sub-matrices $\boldsymbol{X}_{1_j}$ of dimension $n_j \times v$ for each environment such that

$$\boldsymbol{X}_{1_j} = \oplus_{i=1}^{v} \boldsymbol{x}_{ji},$$

where $\boldsymbol{x}_{ji} = [x_{ji_1}\ x_{ji_2}\ \ldots\ x_{ji_{n_{ji}}}]'$ is the plant density observed for each plot in environment j containing genotype i. Under this definition, $\boldsymbol{X}_1$ is equivalent to $\boldsymbol{X}_0$ except that each vector of ones ($\boldsymbol{I}_{n_{ji}}$) is replaced with $\boldsymbol{x}_{ji}$.

To ensure that $\boldsymbol{X}$ is of full column rank, $\boldsymbol{X}_{0_1}$ and $\boldsymbol{X}_{1_1}$ are replaced with $\boldsymbol{1}$ and $\boldsymbol{x}$ respectively, where $\boldsymbol{x}$ is the plant density observed on each plot. As a result, the coefficients $\boldsymbol{\beta}_{0_1}$ and $\boldsymbol{\beta}_{1_1}$ are the intercept and slope for the first variety respectively and the regression coefficients for the remaining varieties are differences in the intercept and slope from the first variety. Thus, in (2), the term $\boldsymbol{\beta}_0$ corresponds to `1 + Genotype` and $\boldsymbol{\beta}_0$ to `density + density:Genotype`.

The column vector $\boldsymbol{u}_{\mathrm{g}}$, with length $2tv$, of random G × E interaction effects can be partitioned into (i) a random trial main effect $\boldsymbol{u}_{\mathrm{h}}$ (dimension $d \times 1$) corresponding to `Trial` in (2), (ii) an environment main effect $\boldsymbol{u}_{\mathrm{e}}$ (dimension $t \times 1$) corresponding to `Env` in (2), (iii) a genotype by trial interaction effect $\boldsymbol{u}_{\mathrm{hg}}$ (dimension $dv \times 1$) corresponding to `Trial:Genotype + density:Trial:Genotype` in (2), and (iv) a G × E interaction effect $\boldsymbol{u}_{\mathrm{eg}}$ (dimension $2tv \times 1$) corresponding to `Env:Genotype + density:Env:Genotype` in (2), such that

$$\begin{aligned}
\boldsymbol{u}_{\mathrm{g}} &= \boldsymbol{Z}_{\mathrm{h}}\boldsymbol{u}_{\mathrm{h}} + \boldsymbol{Z}_{\mathrm{e}}\boldsymbol{u}_{\mathrm{e}} + \boldsymbol{Z}_{\mathrm{hg}}\boldsymbol{u}_{\mathrm{hg}} + \boldsymbol{Z}_{\mathrm{eg}}\boldsymbol{u}_{\mathrm{eg}},\\
\boldsymbol{u}_{\mathrm{g}} &= \boldsymbol{Z}_{0_{\mathrm{hg}}}(\boldsymbol{I}_d \otimes \boldsymbol{I}_v)\boldsymbol{u}_{\mathrm{h}} + \left(\begin{bmatrix}1\\0\end{bmatrix} \otimes \boldsymbol{I}_t \otimes \boldsymbol{I}_v\right)\boldsymbol{u}_{\mathrm{e}} + \begin{bmatrix}\boldsymbol{Z}_{0_{\mathrm{hg}}} & \boldsymbol{Z}_{1_{\mathrm{hg}}}\end{bmatrix}\boldsymbol{u}_{\mathrm{hg}} + \begin{bmatrix}\boldsymbol{Z}_{0_{\mathrm{eg}}} & \boldsymbol{Z}_{1_{\mathrm{eg}}}\end{bmatrix}\boldsymbol{u}_{\mathrm{eg}},
\end{aligned}$$

where $\begin{bmatrix}1\\0\end{bmatrix}$ is an indicator vector ensuring that the environment main effects are adjusted for the regression intercepts only and not the regression slopes, $\boldsymbol{Z}_{0_{\mathrm{hg}}} = \left(\begin{bmatrix}1\\0\end{bmatrix} \otimes (\oplus_{p=1}^{d}\boldsymbol{I}_{f_p}) \otimes \boldsymbol{I}_v\right)$, d denotes the total number of trials, f_p is the number of sowing times in trial p and $t = \sum_{p=1}^{d} f_p$.

The vector $\boldsymbol{u}_{\mathrm{eg}}$ can be further partitioned so that $\boldsymbol{u}_{\mathrm{eg}} = [\boldsymbol{u}'_{0_{\mathrm{eg}}}\ \boldsymbol{u}'_{1_{\mathrm{eg}}}]'$, where $\boldsymbol{u}_{0_{\mathrm{eg}}}$ and $\boldsymbol{u}_{1_{\mathrm{eg}}}$ are $tv \times 1$ vectors of random regression intercept and slope coefficients respectively, $\boldsymbol{u}_{0_{\mathrm{eg}}} = [b_{0_{11}}\ b_{0_{12}}\ \ldots\ b_{0_{1i}}\ b_{0_{21}}\ \ldots\ b_{0_{ji}}]'$, and $\boldsymbol{u}_{1_{\mathrm{eg}}} = [b_{1_{11}}\ b_{1_{12}}\ \ldots\ b_{1_{1i}}\ b_{1_{21}}\ \ldots\ b_{1_{ji}}]'$. A similar partitioning can be done for the term $\boldsymbol{u}_{\mathrm{hg}}$.

The matrix $\boldsymbol{Z}_{\mathrm{g}} = \begin{bmatrix}\boldsymbol{Z}_{\mathrm{h}} & \boldsymbol{Z}_{\mathrm{e}} & \boldsymbol{Z}_{\mathrm{hg}} & \boldsymbol{Z}_{\mathrm{eg}}\end{bmatrix}$ is a $n \times 2tv$ full column rank matrix such that $\boldsymbol{Z}_{\mathrm{eg}} = [\boldsymbol{Z}_{0_{\mathrm{eg}}}\ \boldsymbol{Z}_{1_{\mathrm{eg}}}]$ where $\boldsymbol{Z}_{0_{\mathrm{eg}}}$ and $\boldsymbol{Z}_{1_{\mathrm{eg}}}$ are of dimension $n \times tv$ and $\boldsymbol{Z}_{0_{\mathrm{eg}}} = \oplus_{j=1}^{t}\oplus_{i=1}^{v}\boldsymbol{I}_{n_{ji}}$. Similar to $\boldsymbol{X}_1$, the matrices $\boldsymbol{Z}_{1_{\mathrm{hg}}}$ and $\boldsymbol{Z}_{1_{\mathrm{eg}}}$ are equivalent to $\boldsymbol{Z}_{0_{\mathrm{hg}}}$ and $\boldsymbol{Z}_{0_{\mathrm{eg}}}$ respectively, except that the vector of ones are replaced with the established plant density $\boldsymbol{x}_{ji}$. When genotype i is missing from environment j, the vectors $\boldsymbol{I}_{n_{ji}}$ and $\boldsymbol{x}_{ji}$ within $\boldsymbol{Z}_g$ are replaced with a zero column vector.

The column vector $\boldsymbol{u}_{\mathrm{m}} = [\boldsymbol{u}'_{\mathrm{m}_0}\ \boldsymbol{u}'_{\mathrm{m}_{\mathrm{g}}}]'$ of length $(r-2)(1+v)$ contains the regression coefficients from the natural cubic smoothing spline for the non-linear trait response to plant density. The design matrix corresponding to $\boldsymbol{u}_{\mathrm{m}}$ is $\boldsymbol{Z}_{\mathrm{m}} = [\boldsymbol{Z}_{\mathrm{m}_0}\ \boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}}}]$ of dimension $n \times (r-2)(1+v)$. Specifically, $\boldsymbol{u}_{\mathrm{m}_0}$ is a $(r-2) \times 1$ vector of random coefficients that corresponds to `spl(density)` in (2) and captures the overall non-linear trait response to plant density

where r is the total number of knot points used in fitting the cubic smoothing spline (Verbyla et al., 1999). The $(r-2)v \times 1$ vector $\boldsymbol{u}_{\mathrm{m_g}}$ corresponding to `spl(density):Genotype` in (2), allows for non-linear deviations from the overall non-linear trait response to plant density for each genotype.

The vector $\boldsymbol{u}_{\mathrm{o}}$ contains the peripheral random effects such as experimental design and extraneous spatial field trend terms (Gilmour et al., 1997) with corresponding design matrix Z_{o}. In (1), $\boldsymbol{u}_{\mathrm{o}}$ is represented by the term `design` and is represented by `at(Trial):(Rep/MainPlot/SubPlot)` in (2). The vector $\boldsymbol{\epsilon}$ is a $n \times 1$ vector of residuals for each plot within an environment and is represented by the `Error` term in (2). Heterogeneity of residual variation across environments was incorporated within the residual variance structure such that $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{R})$ where $\boldsymbol{R} = \oplus_{j=1}^{t} \sigma^2_{\epsilon_j} \boldsymbol{I}_{n_j}$.

The random effects and residuals are assumed to be normally distributed with zero mean and variance-covariance matrix

$$\mathrm{var}\left(\begin{bmatrix} \boldsymbol{u}_{\mathrm{g}} \\ \boldsymbol{u}_{\mathrm{m}} \\ \boldsymbol{u}_{\mathrm{o}} \\ \boldsymbol{\epsilon} \end{bmatrix}\right) = \begin{bmatrix} \boldsymbol{G}_{\mathrm{g}} & & & \\ \mathbf{0} & \boldsymbol{G}_{\mathrm{m}} & & \\ \mathbf{0} & \mathbf{0} & \boldsymbol{G}_{\mathrm{o}} & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{R} \end{bmatrix},$$

where $\mathrm{var}(\boldsymbol{y}) = \boldsymbol{Z}_{\mathrm{g}}\boldsymbol{G}_{\mathrm{g}}\boldsymbol{Z}' + \boldsymbol{Z}_{\mathrm{m}}\boldsymbol{G}_{\mathrm{m}}\boldsymbol{Z}' + \boldsymbol{Z}_{\mathrm{o}}\boldsymbol{G}_{\mathrm{o}}\boldsymbol{Z}' + \boldsymbol{R}$. The variance of the (non-spline) G × E random effects $\boldsymbol{G}_{\mathrm{g}}$ can be partitioned into:

$$\mathrm{var}(\boldsymbol{u}_{\mathrm{g}}) = \sigma^2_{\mathrm{h}}\left(\boldsymbol{Z}_{0_{\mathrm{hg}}}[\boldsymbol{I}_d \otimes \boldsymbol{I}_v \boldsymbol{I}']\boldsymbol{Z}'\right) + \sigma^2_{\mathrm{e}}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \boldsymbol{I}_t \otimes \boldsymbol{I}_v \boldsymbol{I}'\right) + \left(\boldsymbol{G}_{\mathrm{h}} \otimes \left(\oplus_{p=1}^{d} \boldsymbol{I}_{f_p} \boldsymbol{I}'\right) \otimes \boldsymbol{I}_v\right) + (\boldsymbol{G}_{\mathrm{b}} \otimes \boldsymbol{I}_t \otimes \boldsymbol{I}_v),$$

where $\boldsymbol{G}_{\mathrm{h}} = \begin{bmatrix} \sigma^2_{\mathrm{h_0}} & \\ \sigma_{\mathrm{h_0,h_1}} & \sigma^2_{\mathrm{h_1}} \end{bmatrix}$ and $\boldsymbol{G}_{\mathrm{b}} = \begin{bmatrix} \sigma^2_{\mathrm{b_0}} & \\ \sigma_{\mathrm{b_0,b_1}} & \sigma^2_{\mathrm{b_1}} \end{bmatrix}$ denote the variance-covariance matrix for the random regression genotype × trial and G × E interaction effects respectively. The variance components σ^2_{h} and σ^2_{e} are the variance components for the random `Trial` and `Env` effects such that

$$\boldsymbol{u}_{\mathrm{h}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{h}} \boldsymbol{I}_d\right), \quad \boldsymbol{u}_{\mathrm{hg}_{pi}} = \begin{bmatrix} \mathrm{h}_{0_{pi}} \\ \mathrm{h}_{1_{pi}} \end{bmatrix} \sim N(\mathbf{0}, \boldsymbol{G}_{\mathrm{h}}),$$

$$\boldsymbol{u}_{\mathrm{e}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{e}} \boldsymbol{I}_t\right), \quad \boldsymbol{u}_{\mathrm{eg}_{ji}} = \begin{bmatrix} b_{0_{ji}} \\ b_{1_{ji}} \end{bmatrix} \sim N(\mathbf{0}, \boldsymbol{G}_{\mathrm{b}}).$$

The variance of $\boldsymbol{u}_{\mathrm{g}}$ becomes such that $\boldsymbol{G}_{\mathrm{h}}$ and $\boldsymbol{G}_{\mathrm{b}}$ denote the variance-covariance matrix for the random regression genotype × trial and G × E interaction effects respectively, and

$$\boldsymbol{u}_{\mathrm{h}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{h}} \boldsymbol{I}_d\right), \quad \boldsymbol{u}_{\mathrm{hg}_{pi}} = \begin{bmatrix} \mathrm{h}_{0_{pi}} \\ \mathrm{h}_{1_{pi}} \end{bmatrix} \sim N(\mathbf{0}, \boldsymbol{G}_{\mathrm{h}}).$$

The definition of the spline basis in Verbyla et al. (1999) ensures that the spline random effects are independent, as given by

$$\boldsymbol{u}_{\mathrm{m_0}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{m_0}} \boldsymbol{I}_{(r-2)}\right), \quad \boldsymbol{u}_{\mathrm{m_g}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{m_g}} \boldsymbol{I}_{(r-2)v}\right). \tag{B.2}$$

B.2. Incorporating environmental covariates into the model

The model that extends the baseline model (B.1) to incorporate a single EC can be expressed as:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{X}_{\mathrm{c}_l}\boldsymbol{\beta}_{\mathrm{c}_l} + \boldsymbol{Z}_{\mathrm{g}}\boldsymbol{u}_{\mathrm{g}} + \boldsymbol{Z}_{\mathrm{m}}\boldsymbol{u}_{\mathrm{m}} + \boldsymbol{Z}_{\mathrm{c}_l}\boldsymbol{u}_{\mathrm{c}_l} + \boldsymbol{g}_l + \boldsymbol{Z}_{\mathrm{o}}\boldsymbol{u}_{\mathrm{o}} + \boldsymbol{\epsilon}, \tag{B.3}$$

where $\boldsymbol{X}$, $\boldsymbol{\beta}$, $\boldsymbol{Z}_{\mathrm{g}}$, $\boldsymbol{u}_{\mathrm{g}}$, $\boldsymbol{Z}_{\mathrm{m}}$, $\boldsymbol{u}_{\mathrm{m}}$, $\boldsymbol{Z}_{\mathrm{o}}$, $\boldsymbol{u}_{\mathrm{o}}$ and $\boldsymbol{\epsilon}$ are as defined in (B.1), and $\boldsymbol{\beta}_{\mathrm{c}_l}$ is a column vector of length v allowing for a separate linear trait response to the l^{th} EC for each genotype pertaining to `ec + ec:Genotype` in (4). The design matrix $\boldsymbol{X}_{\mathrm{c}_l}$ of dimension $n \times v$ is equivalent to $\boldsymbol{X}_1$ except that $\boldsymbol{x}_{ji}$ is replaced with $\boldsymbol{c}_{l_{ji}} = \begin{bmatrix} c_{l_{j i_1}} & c_{l_{j i_2}} & \ldots & c_{l_{j i_{n_{ji}}}} \end{bmatrix}'$ where $\boldsymbol{c}_{l_{ji}}$ denotes the value of the l^{th} EC within a particular plot featuring genotype i in environment j.

The term $\boldsymbol{u}_{\mathrm{c}_l} = [\boldsymbol{u}'_{\mathrm{c}_{l_0}} \; \boldsymbol{u}'_{\mathrm{c}_{l_g}}]'$ is a vector of dimension $(q_l - 2)(1 + v)$ such that q_l denotes the number of knot points for the l^{th} EC. The column vector $\boldsymbol{u}_{\mathrm{c}_{l_0}}$ of length $(q_l - 2)$, corresponding to `spl(ec)` in (4) represents the overall non-linear trait response to the l^{th} EC while $\boldsymbol{u}_{\mathrm{c}_{l_g}}$ of length $(q_l - 2)v \times 1$ which aligns with `spl(ec):Genotype` in (4) allows for a separate non-linear trait response to the l^{th} EC for each genotype. Similar to (B.2), the EC spline terms are assumed to be independent such that

$$\boldsymbol{u}_{\mathrm{c}_{l_0}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{c}_{l_0}} \boldsymbol{I}_{(q_l-2)}\right), \quad \boldsymbol{u}_{\mathrm{c}_{l_g}} \sim N\left(\mathbf{0}, \sigma^2_{\mathrm{c}_{l_g}} \boldsymbol{I}_{(q_l-2)v}\right).$$

The spline design matrix corresponding to $\boldsymbol{u}_{\mathrm{c}_l}$ is $\boldsymbol{Z}_{\mathrm{c}_l} = [\boldsymbol{Z}'_{\mathrm{c}_{l_0}} \; \boldsymbol{Z}'_{\mathrm{c}_{l_g}}]'$, where $\boldsymbol{Z}_{\mathrm{c}_{l_0}}$ is of dimension $n \times (q_l - 2)$ and $\boldsymbol{Z}_{\mathrm{c}_{l_g}}$ of dimension $n \times (q_l - 2)v$.

The form of $\boldsymbol{g}_l$, which contains all of the terms pertaining to the interaction effect between plant density and the l^{th} EC is

$$\boldsymbol{g}_l = (\boldsymbol{x} \otimes_{\mathrm{r}} \boldsymbol{c}_l)\beta_{12_l} + (\boldsymbol{X}_1 \otimes_{\mathrm{r}} \boldsymbol{c}_l)\boldsymbol{\xi}_{12_l} + (\boldsymbol{Z}_{\mathrm{m_0}} \otimes_{\mathrm{r}} \boldsymbol{c}_l)\boldsymbol{u}_{\mathrm{s2}_l} + \left(\boldsymbol{Z}_{\mathrm{m_g}} \otimes_{\mathrm{r}} \boldsymbol{c}_l\right)\boldsymbol{\nu}_{\mathrm{s2}_l} + \left(\mathbf{x} \otimes_{\mathrm{r}} \boldsymbol{Z}_{\mathrm{c}_{l_0}}\right)\mathbf{u}_{1\mathrm{s}_l} + \left(\mathbf{x} \otimes_{\mathrm{r}} \boldsymbol{Z}_{\mathrm{c}_{l_g}}\right)\boldsymbol{\nu}_{1\mathrm{s}_l} + \left(\boldsymbol{Z}_{\mathrm{m_0}} \otimes_{\mathrm{r}} \boldsymbol{Z}_{\mathrm{c}_0}\right)\mathbf{u}_{\mathrm{ss}_l} + \left(\boldsymbol{Z}_{\mathrm{mc}_{l_g}}\right)\boldsymbol{\nu}_{\mathrm{ss}_l}, \tag{B.4}$$

where $\otimes_{\mathrm{r}}$ denotes the 'row-wise' Kronecker product (as defined in Lee and Durbán, 2011; Wood et al., 2013) and $\boldsymbol{x}$ and $\boldsymbol{c}_l$ are column vectors of length n containing the values of established plant density and the l^{th} EC for each individual plot respectively. For the tensor spline terms in (B.4), the subscripts 1 and 2_l for $\beta, \boldsymbol{\zeta}, \boldsymbol{u}$ and $\boldsymbol{\nu}$ correspond to $\boldsymbol{X}_1$ and $\boldsymbol{X}_{\mathrm{c}_l}$ respectively (i.e. the linear component). The subscript s for $\boldsymbol{u}_{\mathrm{s}.}$ and $\boldsymbol{\nu}_{\mathrm{s}.}$ refers to $\boldsymbol{Z}_{\mathrm{m}_0}$ and $\boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}}}$ respectively, while the subscript s for $\boldsymbol{u}_{.\mathrm{s}_l}$ and $\boldsymbol{\nu}_{.\mathrm{s}_l}$ refers to $\boldsymbol{Z}_{\mathrm{c}_{l_0}}$ and $\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}}}}$ respectively. The constant β_{12_l} captures the linear interaction effect between M and the l^{th} EC, while the vector $\boldsymbol{\zeta}_{12_l}$ (of length v) denotes a separate linear surface for each genotype between the trait of interest, M and the corresponding EC. The terms β_{12_l} and $\boldsymbol{\zeta}_{12_l}$ correspond with `density:ec` and `density:ec:Genotype` in (4) respectively.

The vectors $\boldsymbol{u}_{\mathrm{s}2_l}$ (dimension $(r-2)\times 1$), $\boldsymbol{u}_{1\mathrm{s}_l}$ (dimension $(q_l-2)\times 1$) and $\boldsymbol{u}_{\mathrm{ss}_l}$ (dimension $(r-2)(q_l-2)\times 1$) jointly capture the non-linear interaction effect between M and EC$_l$ whilst $\boldsymbol{\nu}_{\mathrm{s}2_l}$ (dimension $(r-2)v\times 1$), $\boldsymbol{\nu}_{1\mathrm{s}_l}$(dimension $(q_l-2)v\times 1$) and $\boldsymbol{\nu}_{\mathrm{ss}_l}$ (dimension $(r-2)(q_l-2)v\times 1$) jointly represent the three-way non-linear interaction effect between genotype, M and EC$_l$. The spline design matrix $\boldsymbol{Z}_{\mathrm{mc}_{l_{\mathrm{g}}}}$ is of dimension $n\times(r-2)(q_l-2)v$ such that

$$\boldsymbol{Z}_{\mathrm{mc}_{l_{\mathrm{g}}}} = \oplus_{i=1}^{v}\left(\boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}_i}}\otimes_{\mathrm{r}}\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}_i}}}\right),$$

where $\boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}_i}}$ and $\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}_i}}}$ are subsetted versions of $\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}}}}$ (dimension $n_v\times(r-2)$) and $\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}_i}}}$ (dimension $n_v\times(q_l-2)$) respectively where if $\boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}}}$ and $\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}}}}$ were reordered in terms of genotypes and then environment nested within genotypes then $\boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}}} = \oplus_{i=1}^{v}\boldsymbol{Z}_{\mathrm{m}_{\mathrm{g}_i}}$ and $\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}}}} = \oplus_{i=1}^{v}\boldsymbol{Z}_{\mathrm{c}_{l_{\mathrm{g}_i}}}$. The variance for each of these terms, in accordance with Verbyla et al. (2018) are

$$\begin{aligned}
\boldsymbol{u}_{1\mathrm{s}_l} &\sim N\left(\boldsymbol{0}, \sigma^2_{1\mathrm{s}_l}\boldsymbol{I}_{(q_l-2)}\right), & \boldsymbol{\nu}_{1\mathrm{s}_l} &\sim N\left(\boldsymbol{0}, \sigma^2_{1\mathrm{s}_l}\boldsymbol{I}_{(q_l-2)v}\right),\\
\boldsymbol{u}_{\mathrm{s}2_l} &\sim N\left(\boldsymbol{0}, \sigma^2_{\mathrm{s}2_l}\boldsymbol{I}_{(r-2)}\right), & \boldsymbol{\nu}_{\mathrm{s}2_l} &\sim N\left(\boldsymbol{0}, \sigma^2_{\mathrm{s}2_l}\boldsymbol{I}_{(r-2)v}\right),\\
\boldsymbol{u}_{\mathrm{ss}_l} &\sim N\left(\boldsymbol{0}, \sigma^2_{\mathrm{ss}_l}\boldsymbol{I}_{(r-2)(q_l-2)}\right), & \boldsymbol{\nu}_{\mathrm{ss}_l} &\sim N\left(\boldsymbol{0}, \sigma^2_{\mathrm{ss}_l}\boldsymbol{I}_{(r-2)(q_l-2)v}\right).
\end{aligned}$$

The terms $\boldsymbol{u}_{1\mathrm{s}_l}, \boldsymbol{\nu}_{1\mathrm{s}_l}, \boldsymbol{u}_{\mathrm{s}2_l}, \boldsymbol{\nu}_{\mathrm{s}2_l}, \boldsymbol{u}_{\mathrm{ss}_l}, \boldsymbol{\nu}_{\mathrm{ss}_l}$ align with `density:spl(ec)`, `density:spl(ec):Genotype`, `spl(density):ec`, `spl(density):ec:Genotype`, `spl(density):spl(ec)`, and `spl(density):spl(ec):Genotype` in (4) respectively.

The inclusion of multiple (say w) ECs into the model in an additive form is given by

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_{\mathrm{g}}\boldsymbol{u}_{\mathrm{g}} + \boldsymbol{Z}_{\mathrm{m}}\boldsymbol{u}_{\mathrm{m}} + \sum_{l=1}^{w}\left(\boldsymbol{X}_{\mathrm{c}_l}\boldsymbol{\beta}_{\mathrm{c}_l} + \boldsymbol{Z}_{\mathrm{c}_l}\boldsymbol{u}_{\mathrm{c}_l} + \boldsymbol{g}_l\right) + \boldsymbol{Z}_{\mathrm{o}}\boldsymbol{u}_{\mathrm{o}} + \boldsymbol{\epsilon}. \tag{B.5}$$

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Selected papers of Hirotugu Akaike. Second International Symposium on Information Theory, pp. 199–213.

Bailey, R.A., 2008. Design of Comparative Experiments. Cambridge University Press.

Boer, M.P., Wright, D., Feng, L., Podlich, D.W., Luo, L., Cooper, M., van Eeuwijk, F.A., 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics 177, 1801–1813. https://doi.org/10.1534/genetics.107.071068.

Butler, D., Cullis, B.R., Gilmour, A.R., Gogel, B.J., Thompson, R., 2017. ASReml-R Reference Manual Version 4. Technical Report. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.⟨www.vsni.co.uk⟩.

Clarke, S., McLean, J., George-Jaeggli, B., McLean, G., De Voil, P., Eyre, J., Rodriguez, D., 2019. Understanding the diversity in yield potential and stability among commercial sorghum hybrids can inform crop designs. Field Crop. Res. 230, 84–97. https://doi.org/10.1016/j.fcr.2018.10.010.

Cooper, M., Hammer, G.L., 1996. Plant adaptation and crop improvement. IRRI.

Cooper, M., Tang, T., Gho, C., Hart, T., Hammer, G., Messina, C., 2020. Integrating genetic gain and gap analysis to predict improvements in crop productivity. Crop Sci. 60, 582–604. https://doi.org/10.1002/csc2.20109.

Craven, P., Wahba, G., 1978. Smoothing noisy data with spline functions. Numer. Math. 31, 377–403. https://doi.org/10.1007/BF01404567.

Crossa, J., 1990. Statistical analyses of multilocation trials. Advances in Agronomy. Elsevier, pp. 55–85. https://doi.org/10.1016/S0065-2113(08)60818-4.

Cullis, B., Gogel, B., Verbyla, A., Thompson, R., 1998. Spatial analysis of multi-environment early generation variety trials. Biometrics 1–18. https://doi.org/10.2307/2533991.

Cullis, B.R., Smith, A.B., Beeck, C.P., Cowling, W.A., 2010. Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome/Natl. Res. Counc. Can. 53, 1002–1016. https://doi.org/10.1139/G10-080.

D. Falconer T. Mackay Introd. Quant. Genet. Essex 1996.

De Faveri, J., Verbyla, A.P., Pitchford, W.S., Venkatanagappa, S., Cullis, B.R., 2015. Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. Crop Pasture Sci. 66, 947–962. https://doi.org/10.1071/CP14312.

De Faveri, J., Verbyla, A., Rebetzke, G., 2022. Random regression models for multi-environment, multi-time (MEMT) data from crop breeding selection trials. Crop Pasture Sci. 74, 271–283. https://doi.org/10.1071/CP21732.

Denis, J.B., 1988. Two way analysis using covariates. Statistics 19, 123–132. https://doi.org/10.1080/02331888808802080.

Forknall, C.R., Simpfendorfer, S., Kelly, A.M., 2019. Using yield response curves to measure variation in the tolerance and resistance of wheat cultivars to fusarium crown rot. Phytopathology 109, 932–941. https://doi.org/10.1094/PHYTO-09-18-0354-R.

Gilmour, A.R., Cullis, B.R., Verbyla, A.P., 1997. Accounting for natural and extraneous variation in the analysis of field experiments. J. Agric. Biol. Environ. Stat. 2, 269–293. https://doi.org/10.2307/1400446.

Gogel, B., Smith, A., Cullis, B., 2018. Comparison of a one-and two-stage mixed model analysis of Australia's National Variety Trial Southern Region wheat data. Euphytica 214, 44. https://link.springer.com/article/10.1007/s10681-018-2116-4.

Hadasch, S., Laidig, F., Macholdt, J., Bönecke, E., Piepho, H., 2020. Trends in mean performance and stability of winter wheat and winter rye yields in a long-term series of variety trials. Field Crop. Res. 252, 107792 https://doi.org/10.1016/j.fcr.2020.107792.

Hammer, G.L., McLean, G., Chapman, S., Zheng, B., Doherty, A., Harrison, M.T., van Oosterom, E., Jordan, D., 2014. Crop design for specific adaptation in variable dryland production environments. Crop Pasture Sci. 65, 614–626. https://doi.org/10.1071/CP14088.

Hammer, G.L., McLean, G., van Oosterom, E., Chapman, S., Zheng, B., Wu, A., Doherty, A., Jordan, D., 2020. Designing crops for adaptation to the drought and high-temperature risks anticipated in future climates. Crop Sci. 60, 605–621. https://doi.org/10.1002/csc2.20110.

Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2. Springer.

Hatfield, J.L., Walthall, C.L., 2015. Meeting global food needs: Realizing the potential via genetics × environment × management interactions. Agron. J. 107, 1215–1226. https://doi.org/10.2134/agronj15.0076.

Heslot, N., Akdemir, D., Sorrells, M.E., Jannink, J.L., 2014. Integrating environmental covariates and crop modelling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 127, 463–480. https://doi.org/10.1007/s00122-013-2231-5.

Jarquín, D., Crossa, J., Lacaze, X., DuCheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., de los Campos, G., 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 127, 595–607. https://link.springer.com/article/10.1007/s00122-013-2243-1.

Kelly, A.M., Smith, A.B., Eccleston, J.A., Cullis, B.R., 2007. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. Crop Sci. 47, 1063–1070. https://doi.org/10.2135/cropsci2006.08.0540.

Kenward, M.G., Roger, J.H., 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 983–997. https://doi.org/10.2307/2533558.

Kirkegaard, J., Hunt, J., 2010. Increasing productivity by matching farming system management and genotype in water-limited environments. J. Exp. Bot. 61, 4129–4143. https://doi.org/10.1093/jxb/erq245.

Krause, M.R., González-Pérez, L., Crossa, J., Pérez-Rodríguez, P., Montesinos-López, O., Singh, R.P., Dreisigacker, S., Poland, J., Rutkoski, J., Sorrells, M., Gore, M.A., Mondal, S., 2019. Hyperspectral reflectance-derived relationship matrices for genomic prediction of grain yield in wheat. G3: Genes, Genomes, Genet. 9, 1231–1247. https://doi.org/10.1534/g3.118.200856.

Laird, N.M., Ware, J.H., 1982. Random-effects models for longitudinal data. Biometrics 963–974. https://doi.org/10.2307/2529876.

Lee, D.J., Durbán, M., 2011. P-spline ANOVA-type interaction models for spatio-temporal smoothing. Stat. Model. 11, 49–69. https://doi.org/10.1177/1471082X1001100104.

Lee, D.J., Durbán, M., Eilers, P., 2013. Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested bases. Comput. Stat. Data Anal. 61, 22–37. https://doi.org/10.1016/j.csda.2012.11.013.

Macdonald, B., 2018. How low can you go? Performance of factor analytic models in the analysis of multi-environment trials with small numbers of varieties. Honours Thesis. University of Southern Queensland. https://sear.unisq.edu.au/37210/.

Malosetti, M., Voltas, J., Romagosa, I., Ullrich, S., Van Eeuwijk, F., 2004. Mixed models including environmental covariables for studying QTL by environment interaction. Euphytica 137, 139–145. https://doi.org/10.1023/B:EUPH.0000040511.46388.ef.

Malosetti, M., Bustos-Korts, D., Boer, M.P., van Eeuwijk, F.A., 2016. Predicting responses in multiple environments: issues in relation to genotype × environment interactions. Crop Sci. 56, 2210–2222. https://doi.org/10.2135/cropsci2015.05.0311.

Meuwissen, T., Hayes, B., Goddard, M., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829. https://doi.org/10.1093/genetics/157.4.1819.

Möhring, J., Piepho, H.P., 2009. Comparison of weighting in two-stage analysis of plant breeding trials. Crop Sci. 49, 1977–1988. https://doi.org/10.2135/cropsci2009.02.0083.

Montesinos López, O.A., MontesinosLópez, A., Crossa, J., 2022. Multivariate statistical machine learning methods for genomic prediction. Springer Nature,.

Nelder, J.A., 1965. The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. Proc. R. Soc. Lond. Ser. A. Math. Phys. Sci. 283 (1393), 163–178. https://doi.org/10.1098/rspa.1965.0013.

Oliveira, I.C.M., Guilhen, J.H.S., de Oliveira Ribeiro, P.C., Gezan, S.A., Schaffert, R.E., Simeone, M.L.F., Damasceno, C.M.B., de Souza Carneiro, J.E., Carneiro, P.C.S., daCostaParrella, R.A., Pastina, M.M., 2020. Genotype-by-environment interaction and yield stability analysis of biomass sorghum hybrids using factor analytic models and environmental covariates. Field Crops Res. 257, 107929 https://doi.org/10.1016/j.fcr.2020.107929.

Patterson, H., Silvey, V., 1980. Statutory and recommended list trials of crop varieties in the united kingdom. J. R. Stat. Soc.: Ser. A (Gen.) 143, 219–240. https://doi.org/10.2307/2982128.

Patterson, H.D., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545–554. https://doi.org/10.1093/biomet/58.3.545.

Perkins, J.M., 1972. The principal component analysis of genotype-environmental interactions and physical measures of the environment. Heredity 29, 51–70. https://doi.org/10.1038/hdy.1972.64.

Piepho, H.P., 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Theor. Appl. Genet. 97, 195–201. https://doi.org/10.1007/s001220050885.

Piepho, H.P., Möhring, J., Schulz-Streeck, T., Ogutu, J.O., 2012. A stage-wise approach for the analysis of multi-environment trials. Biom. J. 54, 844–860. https://doi.org/10.1002/bimj.201100219.

R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨https://www.R-project.org/⟩.

Rodriguez, D., Cox, H., deVoil, P., Power, B., 2014. A participatory whole farm modelling approach to understand impacts and increase preparedness to climate change in Australia. Agric. Syst. 126, 50–61. https://doi.org/10.1016/j.agsy.2013.04.003.

Rodriguez, D., De Voil, P., Hudson, D., Brown, J., Hayman, P., Marrou, H., Meinke, H., 2018. Predicting optimum crop designs using crop models and seasonal climate forecasts. Sci. Rep. 8, 1–13. https://doi.org/10.1038/s41598-018-20628-2.

Rotili, D.H., de Voil, P., Eyre, J., Serafin, L., Aisthorpe, D., Maddonni, G.Á., Rodríguez, D., 2020. Untangling genotype × management interactions in multi-environment on-farm experimentation. Field Crops Res. 255, 107900 https://doi.org/10.1016/j.fcr.2020.107900.

Silverman, B.W., 1985. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. J. R. Stat. Soc.: Ser. B Methodol. 47, 1–21. https://doi.org/10.1111/j.2517-6161.1985.tb01327.x.

Singh, V., Nguyen, C.T., van Oosterom, E.J., Chapman, S.C., Jordan, D.R., Hammer, G.L., 2015. Sorghum genotypes differ in high temperature responses for seed set. Field Crop. Res. 171, 32–40. https://doi.org/10.1016/j.fcr.2014.11.003.

Smith, A., Cullis, B., Thompson, R., 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57, 1138–1147. https://doi.org/10.1111/j.0006-341X.2001.01138.x.

Smith, A., Cullis, B.R., Thompson, R., 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. J. Agric. Sci. 143, 449–462. https://doi.org/10.1017/S0021859605005587.

Speed, T., 1991. That BLUP is a good thing: the estimation of random effects. Stat. Sci. 6, 42–44. https://www.jstor.org/stable/2245699.

Tiezzi, F., de Los Campos, G., Gaddis, K.P., Maltecca, C., 2017. Genotype by environment (climate) interaction improves genomic prediction for production traits in US Holstein cattle. J. Dairy Sci. 100, 2042–2056. https://doi.org/10.3168/jds.2016-11543.

Tolhurst, D.J., Gaynor, R.C., Gardunia, B., Hickey, J.M., Gorjanc, G., 2022. Genomic selection using random regressions on known and latent environmental covariates. Theor. Appl. Genet. 135, 3393–3415. https://doi.org/10.1007/s00122-023-04417-8.

Van Eeuwijk, F., Elgersma, A., 1993. Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. Heredity 70, 447–457. https://doi.org/10.1038/hdy.1993.66.

Van Eeuwijk, F.A., 1996. Between and beyond additivity and non-additivity: the statistical modelling of genotype by environment interaction in plant breeding. Wageningen University and Research.

Van Eeuwijk, F.A., Bustos-Korts, D., Millet, E.J., Boer, M.P., Kruijer, W., Thompson, A., Malosetti, M., Iwata, H., Quiroz, R., Kuppe, C., Muller, O., Blazakis, K.N., Yu, K., Tardieu, F., Chapman, S.C., 2019. Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. Plant Sci. 282, 23–39. https://doi.org/10.1016/j.plantsci.2018.06.018.

Vargas, M., Crossa, J., van Eeuwijk, F.A., Ramírez, M.E., Sayre, K., 1999. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype × environment interaction. Crop Sci. 39, 955–967. https://doi.org/10.2135/cropsci1999.0011183X003900040002x.

Verbyla, A.P., 2019. A note on model selection using information criteria for general linear models estimated using REML. Aust. N. Z. J. Stat. 61, 39–50. https://doi.org/10.1111/anzs.12254.

Verbyla, A.P., Cullis, B.R., Kenward, M.G., Welham, S.J., 1999. The analysis of designed experiments and longitudinal data by using smoothing splines. J. R. Stat. Soc.: Ser. C. (Appl. Stat.) 48, 269–311. https://doi.org/10.1111/1467-9876.00154.

Verbyla, A.P., De Faveri, J., Wilkie, J.D., Lewis, T., 2018. Tensor cubic smoothing splines in designed experiments requiring residual modelling. J. Agric. Biol. Environ. Stat. 23, 478–508. https://doi.org/10.1007/s13253-018-0334-9.

Welham, S.J., Gogel, B.J., Smith, A.B., Thompson, R., Cullis, B.R., 2010. A comparison of analysis methods for late-stage variety evaluation trials. Aust. N. Z. J. Stat. 52, 125–149. https://doi.org/10.1111/j.1467-842X.2010.00570.x.

White, I., Cullis, B., Gilmour, A., Thompson, R., 1998. Smoothing biological data with splines, in: Invited Papers 19th International Biometric Conference, Cape Town, 14–18 December 1998, 57–66.

Wilkinson, G., Rogers, C., 1973. Symbolic description of factorial models for analysis of variance. J. R. Stat. Soc.: Ser. C. (Appl. Stat.) 22, 392–399. https://doi.org/10.2307/2346786.

Wood, S.N., 2006. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. Biometrics 62, 1025–1036. https://doi.org/10.1111/j.1541-0420.2006.00574.x.

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R. CRC Press.

Wood, S.N., Scheipl, F., Faraway, J.J., 2013. Straightforward intermediate rank tensor product smoothing in mixed models. Stat. Comput. 23, 341–360. https://doi.org/10.1007/s11222-012-9314-z.