Genotype specific P-spline response surfaces assist interpretation of regional wheat adaptation to climate change

Daniela Bustos-Korts¹, Martin P. Boer¹, Karine Chenu², Bangyou Zheng³, Scott Chapman², Scott Chapman², Fred van Eeuwijk¹

¹ Biometris, Wageningen University and Research Centre, Wageningen, The Netherlands
² The University of Queensland, Queensland Alliance for Agriculture and Food Innovation, Queensland, Australia
³ CSIRO Agriculture and Food, Brisbane, Queensland, Australia.
⁴ School of Agriculture and Food Sciences, The University of Queensland, Queensland, Australia

*Corresponding author: fred.vaneuuwijk@wur.nl
Abstract

Yield is a function of environmental quality and the sensitivity with which genotypes react to that. Environmental quality is characterized by meteorological data, soil and agronomic management, whereas genotypic sensitivity is embodied by combinations of physiological traits that determine the crop capture and partitioning of environmental resources over time. This paper illustrates how environmental quality and genotype responses can be studied by a combination of crop simulation and statistical modeling. We characterized the genotype by environment interaction for grain yield of a wheat population segregating for flowering time by simulating it using the APSIM cropping systems model. For sites in the NE Australian wheat-belt, we used meteorological information as integrated by APSIM to classify years according to water, heat and frost stress. Results highlight that the frequency of years with more severe water and temperature stress has largely increased in recent years. Consequently, it is likely that future varieties will need to cope with more stressful conditions than in the past, making it important to select for flowering habits contributing to temperature and water stress adaptation. Conditional on year-types, we fitted yield response surfaces as functions of genotype, latitude and longitude to virtual multi-environment trials. Response surfaces were fitted by two-dimensional P-splines in a mixed-model framework to predict yield at high spatial resolution. Predicted yields demonstrated how relative genotype performance changed with location and year-type and how genotype-by-environment interactions can be dissected. Predicted response surfaces for yield can be used for performance recommendations, quantification of yield stability and environmental characterization.

Keywords: GxE, wheat, APSIM, P-splines, adaptation landscape, climate change, breeding strategy.
1. Introduction

Genotypes vary in their sensitivity to the environmental conditions, which is the basis for their improvement by plant breeding. These differences in sensitivity lead to genotype by environment interaction (G×E), potentially changing the genotypic ranking across levels of environmental quality. To understand G×E, and to make genotype recommendations, plant breeders evaluate their candidate varieties (genotypes) in a set of multi-environment trials (METs). METs aim to represent the growing conditions that varieties are likely to encounter when grown by farmers. This set of conditions is usually described as the target population of environments, TPE (Chapman et al., 2000b; Chenu, 2015; Comstock and Moll, 1963; Hammer et al., 2019).

Locations used for the METs are a sample of possible locations that belong to the TPE. Hence, breeders and farmers are not only interested in characterizing adaptation to those specific locations, but across the whole latitude and longitude range encompassed by the TPE. With this aim, genotype adaptation needs to be predicted across the whole range of geographies in which genotypes will potentially be grown (Cooper et al., 2014). Across the TPE, a given genotype shows adaptation to the region in which it realizes highest yields, and for a given region, the highest yielding genotype shows the best adaptation to that region (van Eeuwijk et al., 2016). Predictions of genotype performance across the TPE can be made for subsets of locations that are internally homogeneous, called mega-environments (Atlin et al., 2000; Chauhan et al., 2017; Piepho and Möhring, 2005). Mega-environments can be geographically defined by subsets of locations that share weather or soil characteristics, and these locations may not be contiguous or adjacent, especially when considered in a national or global context. In dry environments as those in the Australian wheat-belt, soil characteristics can largely determine the level of water stress, as they influence water retention capacity (Chenu et al., 2013). Besides describing environments as instances of mega-environments, environments can be described as functions of explicit environmental gradients, represented by latitude and longitude (Lowry et al., 2019). As an extension of this latter approach, latitude and longitude can be complemented and replaced by explicit environmental covariables related to weather or agronomic management (Malosetti et al., 2004; Millet et al., 2016).

There is a range of possible models to predict genotype adaptation across a gradient defined either by geographical coordinates or by explicit environment quality (Bustos-Korts et al., 2016; Piepho et al., 2014; Piepho and Möhring, 2005; Smith et al., 2005; van Eeuwijk et al., 2005, 2019). For example, the factorial regression models are a linear function of the genotypic sensitivities to environmental covariables, and are popular due to their simplicity and because their parameters offer a clear biological interpretation in terms of genotype average performance and sensitivity to the environment (Brancourt-Hulmel et al., 2000; Bustos-Korts et al., 2018; Cullis et al., 1996; Malosetti et al., 2004, 2013; Millet et al., 2016; Parent et al., 2017; Smith et al., 2005).
While factorial regression models are a convenient approach to predict adaptation across an environmental gradient, they may be restrictive in cases where gradient effects are non-linear, as it is often the case in plant breeding. To fit non-linear responses, factorial regression models can be extended to include quadratic or higher order polynomial terms, but this will require large numbers of degrees of freedom with consequent problems in fitting. Spline models offer a flexible alternative to model non-linear responses (Eilers et al., 2015; Eilers and Marx, 1996), and can be even extended to model variation across multiple dimensions (Lee et al., 2013; Rodríguez-Álvarez et al., 2015; Wood, 2017; Wood et al., 2013). Two-dimensional P-spline models are being used to separate genetic differences from spatial heterogeneity within trials (Boer et al., 2020; Rodríguez-Álvarez et al., 2018; Velazco et al., 2017). In this paper, we aim to illustrate the use of 2-dimensional P-spline approaches at larger scales than trials, i.e. to model spatially-dependent G×E variation at the level of the TPE, predicting the yield response surfaces of individual genotypes as a function of only latitude and longitude.

The choice of which environmental covariables to include in the prediction model largely depends on the environmental drivers for G×E. Within the TPE sample represented by METs, there may be recurring or repeating characteristics (i.e. that remain constant across years for a given location) that induce differential genotypic responses, which are an expression of G×E. Typical repeating characteristics are associated with latitude, longitude and to soil type. Latitude usually has a large effect on differential genotype adaptation via its effect on phenology (Zheng et al., 2012, 2013), whereas soil characteristics largely impact nutrient and water availability for the crop. Hence, soil differences between locations are usually increased in environments with low rainfall (Wang et al., 2017). Note however that in real-world METs, the effects of ‘location’ may be affected by the fact that a ‘location’ reference is usually associated with a nearby geographical reference (such as a town name), and the actual trials are not done in the exact same field each year even if breeders try to choose a typical soil type and management. In low rainfall environments like Australia where fallows and crop rotations are common, running a trial in exactly the same field and same place in successive years would be considered poor experimental practice due to potential carry-over effects of plot effects, impacts on soil water reserves and pressure of weeds and diseases.

Other characteristics of environmental quality are less predictable because they change from year to year with the weather fluctuations, but they are predictable from their frequencies estimated from long-term data. For example, the long-term frequency of water supply-demand ratio (Chapman et al., 2000b; Chenu et al., 2013) and the frequency of El Niño–Southern Oscillation (ENSO) events in the eastern Australian wheat-belt (Zheng et al., 2018) could potentially be used to estimate the probability of a certain stress level to occur at a particular location. If the probability of a certain year type (weather scenario) can be estimated from long-term data, it becomes attractive to predict response surfaces across latitude and longitude for each of the likely weather scenarios across years. In that
way, latitude and longitude effects become repeatable G\(\times\)E, conditional on a year type or weather scenario.

In Australia, national variety trials of wheat are conducted by the GRDC (Grains Research and Development Corporation) in cooperation with commercial breeding companies. Those companies also conduct their own research trials. However, in neither of these variety trials are the same varieties grown over large numbers of seasons. The limited replication of varieties over years represents a bottleneck in studying long-term adaptative responses. This bottleneck can be addressed by utilizing crop simulation models to construct synthetic/virtual breeding trial datasets that span a longer series of years. This approach has been widely adopted for multiple crops; e.g. sorghum (Chapman et al., 2002; Hammer et al., 2014, 2019); maize (Chenu et al., 2009; Harrison et al., 2014; Messina et al., 2011); wheat (Chenu et al., 2017), and soybean (Messina et al., 2006), including studies that look at flowering time effects in wheat for current (Zheng et al., 2015b) and future climates (Zheng et al., 2016).

In this paper, we present the use of P-splines embedded in mixed models to interpret G\(\times\)E and predict the adaptive responses of individual wheat genotypes from simulated data for a region of about 677 by 445 km in size. APSIM yield was simulated for 156 genotypes varying in flowering time and sown in 13 Australian locations across 39 years of weather records. We focus on the adaptive responses across latitude and longitude, and we examine how these response surfaces change depending on the level of drought and heat stress present across years. We also describe the adaptation landscape in terms of the traits contributing to adaptation across environmental conditions (i.e. sensitivity to photoperiod, vernalization requirements and thermal time requirements from floral initiation and flowering).

2. Methods

This section describes the main steps of our approach; data generation using simulations in APSIM Wheat, G\(\times\)E analysis of outputs generated by APSIM and construction of environmental indices using APSIM outputs to facilitate classification of years in year types. Conditional on year type, yield response surface models for individual genotypes were fitted as functions of longitude and latitude using P-spline methods within a mixed model context. The fitted response surfaces provided predictions of yield responses at any desired resolution level for all genotypes. The yield predictions were used to subdivide the area defined by longitude and latitude coordinates in adaptation zones. Our workflow is schematically represented in Figure 1.
2.1. Simulated data

Data corresponded to grain yield for 156 wheat genotypes simulated by Zheng et al. (2018) using the APSIM cropping system model (Holzworth et al., 2014) together with a phenology model (Zheng et al., 2013), frost impact module (Zheng et al., 2015a) and heat impact module (Lobell et al., 2015). In this data set, variation in APSIM genotype specific parameters was induced by allelic variation for the VRN-A1, VRN-B1, VRN-D1, and PPD-D1 genes, and the full range of values of additional thermal time requirements from floral initiation to flowering (from 425 to 1025 °Cd, Zheng et al., 2013). The set of genotypes included commercial varieties and virtual genotypes that could potentially be bred based on the flowering alleles present in the Australian germplasm pool (Zheng et al., 2013). Allelic combinations at vernalization and photoperiod genes produced variation for the APSIM parameters; Ppd, Vrn and Eps. Genotypes with the same phenology (but different allelic combinations) were disregarded, so that a total of 156 genotypes unique for their phenology were considered. Overall, the selected genotypes had APSIM parameters ranging from 0 to 1.2 for the photoperiod sensitivity (Ppd parameter, with values of 0, 0.3, 0.6, 0.9 and 1.2, with 0.6 for the reference genotype Janz), 0.9 to 1.7 for the vernalization sensitivity (Vrn parameter, with values of 0.9, 1.1, 1.3, 1.5 and 1.7, with 0.9 for Janz), and 425 to 1025°Cd for earliness-per-se (Eps parameter, with values of 425, 475, 525, 575, 625, 675, 725, 775, 825, 925, 975 and 1025°Cd, with 675°Cd for Janz). Genotypes were labelled by their flowering time parameters; the first number indicates the value for sensitivity to photoperiod, the second indicates vernalization requirement, and the third number indicates the minimum thermal time requirement from floral initiation to flowering. For example, ‘g1.2_0.9_425’ indicates a genotype with a sensitivity to photoperiod of 1.2, vernalization requirements of 0.9 and minimum thermal time requirement from floral initiation to flowering of 425°Cd. For most of the environments, the range for flowering time was around 50 days (Supplementary Figure S5). Note that this genotypic variation can be considered to be rather extreme compared to real-world conditions. Australian breeders tend to select wheats for early-season (slower maturing) or main-season (faster maturing) sowing times. Commercial wheats are often classified as ‘quick’, ‘medium’ or ‘slow’ and are usually compared to reference cultivars that are established on the market. Consequently, the range in flowering time within a typical breeders trial may be 3-4 weeks in early generation breeding, or 1 to 3 weeks in the type of MET we are considering here, i.e. much less than 50 days.

In this study, we focused on 13 out of the 15 locations used in Zheng et al. (2018), removing ‘Emerald’ and ‘Roma’ (Figure 2; Table 1). We dropped Emerald because it was geographically too distant, which doesn’t allow for a reliable surface estimation with P-splines. Roma had extreme stress conditions, leading to zero yield for many genotypes. For the 13 locations we
considered, yield was simulated for each season from 1978 to 2016. Zheng et al. (2018) simulated several sowing dates. In this study, we restricted ourselves to one sowing date per location (the same date was used across years). The selected sowing date per location was identified as the one leading to the largest yield for the average of the genotypes (‘optimal’ sowing date, Table 1). For a location, the starting soil conditions were the same in every year of simulation and represented the average starting condition for that location after the analysis of historical data in Zheng et al. (2018). Four environments with a large crop failure were removed, leaving in total 503 environments for the G×E analysis.

2.2. Using environmental covariables to classify years into year types

Water and temperature stress are common environmental drivers for grain yield in Australian wheat production systems (Ababaei and Chenu, 2020; Chenu et al., 2011, 2013). To help classifying years according to their water and temperature stress, we used APSIM to compute four environmental indices for each genotype (Table 2; Supplementary Figures S2, S3 and S4). These indices were related to water, frost and heat stress. Indices were calculated between 300°C before flowering and 100°C after flowering, coinciding with the critical period for grain number determination (Table 2).

After computing the indices per genotype, we averaged the values across all genotypes for each year-location combination. The average environmental indices across genotypes were used to characterize environmental quality. Then, data was arranged in a matrix with years as objects in rows and index-location combinations as variables in columns. Variables were centred and scaled, and were used to calculate Euclidean distances between years. These distances were used in a hierarchical clustering procedure to classify years according to the four environmental indices relating to water, frost and heat stress (Ward method; Zelterman, 2015) using the ‘hclust’ function in R (R Core Team, 2019). With this clustering procedure, years were classified into two classes; ‘mild’ years (with reduced water and temperature stress) and ‘hot and dry’ (with water, frost and heat stress).

2.3. Mixed model G×E analysis

2.3.1. Variance components model

To quantify the relative contribution of locations and years to the total GxE, we fitted the following mixed model to the APSIM yield:

\[ y_{ijk} = \mu + L_j + Y_k + LY_{jk} + G_i + GL_{ij} + GY_{jk} + GLY_{ijk} \] (1)
In Equation (1), $y_{ijk}$ is the phenotype (yield) of the $i^{th}$ genotype in the $j^{th}$ location and the $k^{th}$ year ($i=1,...,156; j=1,...,13, k=1,...,39$). $\mu$ is the general intercept, $L_j$ and $Y_k$ are the fixed effects of location and year and $LY_{jk}$ is the fixed interaction between location and year. $G_j$ is the random main effect of the $j^{th}$ genotype whereas $GL_{ij}$ is the random effect of genotype by location interaction. $GY_{ik}$ represents genotype by year interaction and $GLY_{ijk}$ corresponds to a residual term that contains the genotype by location by year interaction. Because of the absence of replicate information (APSIM output consisted of one yield observation per genotype-environment combination), the last term represents only interaction between genotype, location and year, whereas for real data the residual term would contain within trial error as well. Random effects were assumed to be independent and normally distributed with zero means and specific variances: $G_j \sim N(0, \sigma^2_{G})$, $GL_{ij} \sim N(0, \sigma^2_{GL})$, $GY_{ik} \sim N(0, \sigma^2_{GY})$, $(GLY_{ijk}) \sim N(0, \sigma^2_{E})$.

To characterize the main sources of variation and quantify the contribution of year types (‘s’, scenarios) to G×E variation, the model was expanded to model (2):

$$y_{ijk(s)} = \mu + L_j + T_s + Y(T)_{k(s)} + LY_{jk} +$$
$$G_j + GL_{ij} + GT_{js} + GY(T)_{ik(s)} + GLT_{jks} + GLY(T)_{ijk(s)} \quad (2)$$

Equation (2) substitutes the random terms $GY_{ik}$ and $GLY_{ijk}$ from equation (1) by the following random terms; $GT_{js}$, which is the random interaction between genotype and year type, $GY(T)_{ik(s)}$ that represents genotype by year (within year type) interaction, $GLT_{jks}$ that is the three way genotype by location by year type interaction, and $GLY(T)_{ijk(s)}$ that corresponds to a residual term that contains genotype by location by year interaction. As in model (1), all the random effects are assumed to be independent and normally distributed, with zero means and homogeneous variances. For the fixed part of the model, the year effect was partitioned into a year type ($T_s$) and year within year type ($Y(T)_{k(s)}$) effect.

As the set of genotypes used in this study was specifically segregating for APSIM parameters related to flowering time, we also assessed the relative importance of those parameters, using the following model:

$$y_{fgh(i)l} = \mu + E_l + Ppf_{f(i)} + Vrn_{g(i)} + Eps_{h(i)} + Ppd.\text{Env}_{f(i)l} + Vrn.\text{Env}_{g(i)l} + Eps.\text{Env}_{h(i)l} +$$
$$Ppd.\text{Vrn.}\text{Env}_{fgh(i)l} + Ppd.\text{Eps.}\text{Env}_{fgh(i)l} + Vrn.\text{Eps.}\text{Env}_{gh(i)l} + GE_{fgh(i)l} \quad (3)$$
In model (3), $y_{gith}t$ is the yield of genotype $i$ in environment $t$ (year-location combination or trial), $E_t$ is the fixed environment effect, $Ppd_{f(i)}$, $Vrn_{g(i)}$ and $Eps_{h(i)}$ are the random effects of the APSIM parameters that were used to generate the genotype $i$ (see section ‘simulated data’). These parameters regulate photoperiod response, vernalization requirements and thermal time requirements from floral initiation to flowering. For the $Ppd$, $Vrn$ and $Eps$ APSIM parameters, the parameters defined 5, 5 and 13 classes, respectively. The interactions between APSIM parameters regulating phenology and the environment were also included. The term $G\times E$ represents residual $G\times E$. All random terms were assumed to be independent and normally distributed, with zero means and homogeneous variances.

2.3.2. GGE biplot

To visualize the contribution of APSIM parameters regulating phenology to APSIM yield and to describe their relation to genotypic performance across environments, we used a genotype-genotype-by-environment model (GGE, Yan and Kang, 2002; Yan and Rajcan, 2002).

$$y_{it} = \mu + E_t + \sum_{m=1}^{M} b_{im}z_{tm} + \xi_{it} \quad (4)$$

In model (4), $y_{it}$ represents the mean yield of the $i^{th}$ genotype in the $t^{th}$ trial, $\mu$ stands for an intercept and $E_t$ is the fixed trial effect. The genotype main effect and the interaction are explained by $M$ multiplicative terms. Each multiplicative term is formed by the product of a genotypic sensitivity ($b_{im}$ genotypic score) and environmental scores $z_{tm}$. Finally, $\xi_{it}$ is a residual term. To visualize how sensitivity to photoperiod, vernalization requirements and earliness contribute to $G\times E$, the first two PCs estimated for $G+G\times E$ ($\sum_{m=1}^{2} b_{im}z_{tm}$ in model 4) were visualized as a GGE biplot, label-colouring genotypes according to their values for the APSIM parameters regulating phenology.

To understand the contribution of environmental conditions to genotypic performance, the standard GGE biplot was enriched with environmental information. The scaled environmental covariables were regressed against the environmental PC1 and PC2 from the GGE model. The coefficients of these regressions were used to describe directions of greatest change for these covariables in the biplot (Graffelman and Van Eeuwijk, 2005; Voltas et al., 1999). The direction is given by the regression coefficients and the origin. Furthermore, the angles between the direction vectors again give information about correlations: small angles between vectors mean high correlations, while angles $> 90^\circ$ indicate negative correlations between the vectors.
2.3.3. Predicting response surfaces with P-splines embedded in a mixed model

To describe the genotypic response across latitude and longitude, we used the following model for the APSIM yield data for each genotype separately, i.e., conditional on $i$:

$$
Y_{ijk} = \mu + Y_{k} + \left( \beta_1^i \right)^T s_1(lat)_j + \left( \beta_2^i \right)^T s_2(lon)_j + \left( \beta_3^i \right)^T s_3(lat,lon)_j + \epsilon_{ijk}
$$

In Equation (5), $Y_{ijk}$ is the yield of genotype $i$ in location $j$ and year $k$, $Y_{k}$ is the fixed year effect. The term $s_1(lat)_j$ defines the evaluation at location $j$ of a set of basis functions (in vector form) for the spline fit on the latitude of the trial, while $\left( \beta_1^i \right)^T$ is the corresponding set of genotype specific random coefficients (in row vector form). Similar spline terms are defined for longitude, $\left( \beta_2^i \right)^T$, and the interaction between latitude and longitude, $\left( \beta_3^i \right)^T$. The interaction is orthogonal to the main effect spline terms for latitude and longitude (Boer et al., 2020; Piepho et al., 2021; Wood, 2017; Wood et al., 2013). The residual term $\epsilon_{ijk}$ contains GxE not explained by the spline surfaces, and, for real data, within trial error. Smooth terms were fitted using first-degree P-splines with 10 segments and first-order penalties (Eilers et al., 2015). Higher degree P-splines and higher differences could also be used, but the advantage of using first-degree P-splines and first order differences is that this model is equivalent to a linear variance model (Boer et al., 2020; Williams, 1996). The latitude and longitude range of the spline segments coincides with the latitude and longitude range spanned by the trials. The P-spline mixed model was fitted with ASReml-R v. 4 (Butler et al., 2017).

As there were contrasting year types, it made sense to predict the genotypic response surfaces of each genotype across latitude and longitude, conditional on the year type (one predicted surface per genotype for ‘mild’ years and another predicted surface per genotype for ‘hot and dry’ years). Then, for each genotype we used the predicted surfaces to subtract ‘mild’ years from ‘hot and dry’ years. In this way, the difference in yield between ‘mild’ and ‘hot and dry’ years represents the sensitivity to year type, across the whole latitude and longitude span.

To characterize the surfaces for explicit environmental quality, we fitted model (5), replacing yield as a response by the environmental indices that were used to classify years; rainfall ($S2\_sum\_rain$), average maximum temperature ($S3\_avg\_maxt$), sum of frost temperatures ($S2\_frost\_sum2$) and vapour pressure deficit ($S2\_vpd2$) averaged across genotypes.

To quantify the contribution of year-to-year variation (within year type) to the total variation, we extended model (5) as follows:
\[ y_{ijk} = \mu + \gamma_k + \left( \beta_1^T \right)^T s_1(lat)_j + \left( \beta_2^T \right)^T s_2(lon)_j + \left( \beta_3^T \right)^T s_3(lat,lon)_j + \left( \beta_{4k}^T \right)^T \gamma_1(lat)_j + \left( \beta_{5k}^T \right)^T \gamma_2(lon)_j + \left( \beta_{6k}^T \right)^T \gamma_3(lat,lon)_j + \epsilon_{ijk} \] (6)

Here, the term \( \left( \beta_{4k}^T \right)^T \) represents the set of random coefficients for latitude (in row vector form) that are specific to each year, i.e., they are scenario dependent deviations of the genotype specific coefficients introduced above. Similar spline terms are defined for year-by-longitude variation, \( \left( \beta_{5k}^T \right)^T \), and the contribution of year to year variation to the interaction between latitude and longitude, \( \left( \beta_{6k}^T \right)^T \). We quantified the contribution of each model term by starting with a null model with only a fixed year main effect. Then, we sequentially added terms in model 6. We quantified the contribution of each term by calculating the difference in the residual variance between the null model and the model with spline terms, divided by the residual variance of the null model.

### 2.3.4. Finding patterns across the genotypic response surfaces - Location classification

#### 2.3.4.1. Highest yielding genotypes across latitude and longitude (method 1 in Figure 1)

We used model (5) to make predictions for each genotype over a grid of 140 by 100 points that covered a latitude range from -27° S to -33.5° S and a longitude range of 147° E to 151.5° E (these ranges corresponded to the latitude and longitude ranges for the locations used in the simulations, Figure 2). The covered area approximately spans 677 km from North to South and 445 km from East to West, making each pixel equivalent to about 4.8 by 4.8 km.

For each pixel on this grid (=location), we identified the five genotypes that had the highest yield in the tested conditions (i.e. 1 sowing date, 1 soil condition), as predicted by the 2-dimensional spline surfaces across years within year type. The location by genotype matrix contained a 1 if a genotype was in the top five of the genotypic ranking at a particular location and a 0 if it was not. This is equivalent to applying a selection intensity of 3.2%. This presence or absence matrix was used to calculate the Jaccard similarities between locations, based on which were the five highest yielding genotypes. This similarity matrix was used for a hierarchical clustering procedure of locations (Ward method) implemented in the ‘hclust’ function in R (R Core Team, 2019). The resulting regions have similar set of highest yielding genotypes. In such a way, these regions would reflect set of locations for which breeders might do the same genotype selection, at least in terms of maturity. We ran this procedure using the predictions for each year type separately (i.e. producing one classification for years with mild heat and water stress and one classification for ‘hot and dry’ years).
2.3.4.2. Using G+G×E to cluster environments in the whole adaptation landscape (method 2 in Figure 1)

Assigning locations to regions by commonality of the five winning genotypes is a simple and straightforward method to classify locations. However, if genotypes that are in the upper yield percentiles are very similar in yield, this might lead to frequent changes in genotypic ranking across latitude and longitude, associated to small yield differences. This potentially leads to frequent spatial discontinuities in the classification of environments. A more robust procedure assigns locations to regions by the full set of fitted yields for all genotypes. Hence, we created yield predictions for each genotype on a grid of 15 by 24. For computational convenience, we reduced the grid resolution, compared to the analysis above that looked at the winning genotype per pixel. We considered that each point in the grid defined by latitude and longitude defined a (virtual) location. In this part of the analysis, each pixel covered about 28 by 28 km.

As in the GGE model (Yan and Kang, 2002), we fitted a principal components models to the genotype by virtual location matrix of the environment-centred predicted yields from the 2-dimensional spline surfaces across years within each year type. As the first two principal components explained most of the relevant variation (87.1 and 87.6% for ‘mild’ and ‘hot’ years, respectively), we retained the scores of both of them to construct a location by location similarity matrix (Euclidean distances). This similarity matrix was used for a hierarchical clustering procedure of locations (Ward method) implemented in the ‘hclust’ function in R (R Core Team, 2019). For each year type, the clustering procedure resulted in location classes (regions).

2.3.5. Quantifying the contribution of sensitivity to photoperiod, vernalization requirements and earliness per se to the predicted adaptation landscape per year type

The contribution of sensitivity to photoperiod, vernalization requirements and earliness per se to the predicted adaptation surface was also quantified with a mixed model that was fitted to the spline-predicted yield across virtual locations. The following mixed model was fitted to predictions made for each year type separately, where the subscript \( r \) refers to the region as identified within year type by clustering on G+G×E of the predicted adaptation landscape;

\[
Y_{j(i)j(r)} = \mu + L_j + P_{pd(i)} + V_{rn(i)} + E_{ps(i)} + P_{pd.Region(i)r} + V_{rn.Region(i)r} + E_{ps.Region(i)r} + P_{pd.Vrn.Region(i)r} + P_{pd.Eps.Region(i)r} + V_{rn.Eps.Region(i)r} + GE_{fgh(i)r} \tag{6}
\]

Model (6) follows the logic of model (3), \( Y_{fghr} \) is the APSIM yield of the flowering time parameter combinations \( f, g, h, i \) in location \( j \), \( L_j \) is the fixed location effect, \( P_{pd(i)} \), \( V_{rn(i)} \) and
\( EPS_{h(i)} \) are the random main effects of the APSIM parameters regulating photoperiod response, vernalization requirements and thermal time requirement from floral initiation to flowering (with parameters fitted as factors), nested within genotype \( i \). The interactions between APSIM parameters regulating phenology and regions were also included. The term \( GE_{fph(i)j(c)} \) represents residual G\( \times \)E.

3. Results

3.1. Year classification

The clustering procedure on environment covariables suggested two clearly defined groups of years (Figure 3), which were contrasting in their levels of water and temperature stress. Year type 1 was characterized by larger rainfall (S2_sum.rain, supplementary figure 1), lower average maximum temperature (S3_avg.maxt, supplementary Figure 2), less accumulation of frost temperatures (S2_frost.sum, supplementary figure 3) and lower vapour pressure deficit (S2_vpd, supplementary Figures 1 to 4) than type 2 years. Given these differences, year type 1 can be interpreted as years with mild temperature and water stress (hereafter referred to as ‘mild’ years), and year type 2 as years with strong temperature (heat and frost) and water stress (hereafter referred to as ‘hot and dry’ years). Noteworthy, the relative frequency of ‘hot and dry’ years has greatly increased in the most recent decades. For example, 2 and 4 years were classified as ‘hot and dry’ in the decades 1978-1987 and 1988-1997, whereas 6 and 7 years were classified as ‘hot and dry’ in the periods 1998-2007 and 2008-2016 (Figure 5). This is consistent with the estimates of a trend increase of average temperature for August to November of about 0.05°C per year from 1985 to 2017 in this region (Fig. 3 in Ababaei and Chenu, 2020).

Within year type, there was also spatial variation for the environmental conditions (Figure 4). Locations in the North and East had higher rainfall (S2_sum.rain, Figures 4A and 4B), lower maximum temperature (S3_avg.maxt, Figures 4C and 4D) and vapour pressure deficit (S2_vpd, Figures 4E and 4F) than the other locations. Although the spatial pattern for S3_avg.maxt2 and S2_vpd was preserved across year types, the absolute values were very different between year types; ‘hot and dry’ years had lower S2_sum.rain, higher S3_avg.maxt and S2_vpd than ‘mild’ years. The spatial pattern of S2_frost.sum differed between year types; in ‘mild’ years, S2_frost.sum was larger in Southern locations than in the rest of the region, reflecting lower frost stress (Figure 4G). In contrast, frost temperatures were more important in ‘hot and dry’ years (Figure 4H).
3.2. Variation in flowering time and yield

The variation in sensitivity to photoperiod, vernalization requirements and earliness per se led to large variation in flowering time (Supplementary Figure 5). Locations differed in the means and range of days to flowering. The shortest mean duration between sowing and flowering was observed in Goondiwindi, with a long-term mean across genotypes of 107 days. The largest duration was in Wellington, with a long-term mean across genotypes of 156 days. Within location, days to flowering did not seem to vary much between year types (Supplementary Figure 6). Note that the majority of the flowering dates (25 and 75 percentiles) coincide with the flowering ranges observed in real breeding trials. However, the full range of flowering dates is ca. 10 to 25 days greater than in most breeding trials.

The influence of year types was much larger on yield than on days to flowering, where year types had a strong effect on genotypic performance, with ‘hot and dry’ years having in general lower yield across locations than ‘mild’ years (Figure 5), consistent with differences in environment quality. Given that ‘hot and dry’ years have increased in their frequency during the most recent years (Figure 5), a strategy to select for wheat varieties that are well-adapted to future climate conditions could be to select for varieties with adapted ‘flowering genetics’ or sow earlier (Collins and Chenu, 2021; Zheng et al., 2016).

3.3. Variance components for yield across locations and years

The variance components model indicates that there is large G×E for yield in this data set (Table 3). Most of the G×E is not consistent from one year to the next and is mainly related to the two-way genotype by year interaction and the three way interaction between genotypes, locations and years (captured in the residual in Table 3). For the two-way genotype by year interaction, we investigated the option to classify years in groups that are more internally homogeneous and predictable. For example, from explicit covariables, as we did here and as done in Chenu et al. (2011); Zheng et al. (2018), or from long-term frequencies, as done in Chenu et al. (2013). The genotype by location interaction will be examined by fitting smooth surfaces with P-splines across latitude and longitudes.

3.4. Variance components for yield between year types

Upon integrating year type in the G×E analysis (Table 4), we see that the variance component for genotype by location by year type interaction is about half the magnitude of the variance component for genotype by location interaction. Therefore, it will be interesting to model the
genotype by location interaction by explicit functions conditional on the year type. Variance components for individual effects in Table 3 are different from those of Table 4 (especially the genotype main effect) because of the absence of some locations in particular years (four year-location combinations were removed because of large crop failure in those environments).

3.5. Contribution of APSIM parameters regulating flowering time to G×E

We also estimated the contribution of APSIM parameters regulating flowering time on yield. Across the 503 environments considered in this analysis, the largest contribution to the yield G×E was made the three-way interactions, especially Ppd:Eps:Env and Vrn:Eps:Env (Table 5). This implies that specific combinations of photoperiod and vernalization alleles, in combination with different earliness levels make important contributions to wheat adaptation across environments. The two-way interactions between Eps:Env and Ppd:Env also made very important contributions to the total G×E variation (Table 5). The same result can be illustrated in Supplementary Figures 7 and 8 that describe the relationship between yield and flowering time for each of the environments. In most environments, this relationship shows an optimum, but the position of this optimum between yield and days to flowering depends on the year-location combination.

When assessing the relationship between yield and the APSIM parameters regulating phenology, it can be observed that for all environments there is an optimum Eps value, and that the position and height of this optimum depends on the specific combinations of Ppd and Vrn values (Figure 6). Furthermore, the effect of Ppd and Vrn on yield is much larger in mild than in hot years, leading to a larger phenotypic variance in ‘mild’ than in ‘hot and dry’ years (as already observed in Figures 4 and 5). There is a large interaction between earliness and Ppd/Vrn values, shown in the yield crossing overs in all locations and year types. However, the optimum combination of Eps and Ppd/Vrn depends on the location, and it is also highly influenced by year type. This explains the large and complex G×E observed in this data set. The Eps value that leads to optimum yield is in general lower for hot than for ‘mild’ years, coinciding with the general observation that long-cycle genotypes run out of water in ‘hot and dry’ years. In ‘hot and dry years’, the optimum combination of APSIM parameter values is an intermediate value of Eps with large values for Ppd and Vrn. However, when Eps is larger than 700°Cd, genotypes with a lower value of Ppd/Vrn have an advantage. This leads to large cross-over G×E in this data set (Figure 6). In contrast, for any Eps value in ‘mild’ years, having large Ppd and Vrn values is an advantage, compared to low values for Ppd and Vrn. When repeating the analysis for the periods 1978-1997 and 1998-2016, earlier-flowering genotypes, determined either by lower Eps values or by higher Eps values in combination with small values for Ppd/Vrn, are favoured in the most recent decades (Supplementary Figure 8).
A GGE analysis per year type characterized the effect of combinations of flowering time parameters on G×E for yield (Figure 7). In both years, the genotype main effect (Geno_mean) was more associated to larger values of Ppd (smaller angle between vectors), than to the Vrn and Eps. This result coincides with Figure 6, which showed that larger values for Ppd generally lead to larger yields. In contrast, Eps contributed more to G×E than to the genotype main effect. These results coincide with the variance components model in Table 5 and with Figure 6, which shows that the optimum value for Eps depends on the environment. In Figure 7, examining the relative positions of the vectors for environment indices and genotypic parameters regulating flowering time gives insight in the underlying physiological mechanisms of G×E in this data set. In ‗hot and dry‘ years, Eps induced a large G×E interaction with warmer and dryer environments (this can be observed by the almost opposite position of the vector for Eps with S2_avg.max and S2_vpd). In Figure 6, this can be observed as the position of the optimum for yield, in relation to Eps values, occurs at higher values for environments that are less drought-prone, like ‘Narrabri‘ or ‘Wellington’ (Figure 4). This indicates that when more water is available, it is advantageous to have larger values for Eps (hence, later flowering), compared to very dry environments. In ‘mild’ years, Eps also had a positive interaction with more favourable environments that had a larger rainfall (S2_sum.rain2). Larger values for Ppd and Vrn contributed to generate a negative interaction with warmer environments (i.e. those with larger values for S2_avg.max). In Figure 6, this can for example be observed as lower Ppd and Vrn values leading to a larger yield in drought-prone locations like ‘Meandarra’, ‘Goondiwindi’ and ‘Dalby’ (Figure 4).

3.6. Spline surfaces across latitude and longitude for each year type

As the interaction of genotype and year type was large (Table 4 and GTs in Equation 2), it is potentially useful to inspect the spline-fitted surfaces per year type. We quantified the contribution of each term in the P-spline model by sequentially adding those terms and computing the reduction in the residual after including an additional model term (equation 6, Supplementary Figure 9). The relative importance of latitude, longitude, their two-way interactions and their three-way interactions with year indicate that ‘hot and dry‘ years in general required a larger model complexity, as the interactions between latitude and longitude with year explained a larger proportion of the variation than in ‘mild‘ years. In ‘hot and dry‘ years, genotypes also differed more in the contribution of each model term (Supplementary Figure 9).

When inspecting the predicted response surfaces for each genotype and year type, ‘mild‘ years had an average yield that was 2.13 times larger than that of ‘hot and dry‘ years (2879 vs 1351 kg ha⁻¹, Figure 8). For both ‘mild‘ and ‘hot and dry‘ years, there was a yield gradient; average yield per virtual location (pixel) was larger in the East and South-East (close to locations Wellington and Gunnedah), than in the West (especially in locations Walgett and Meandarra). This aligns
with the rainfall isohyets which decrease from NW to SE in this part of the country (Chenu et al., 2011).

However, the average yield deviation for each virtual location (pixel) in the latitude-longitude range spanned by the trials, compared to the mean yield, was much less in ‘mild’ than in ‘hot and dry’ years. In ‘mild’ years, the best locations had average yields that were about 1.3 times the general mean (for the same year type), and the worst locations had average yields that were about 0.8 times the general mean. In contrast, in ‘hot and dry’ years, the best locations had average yields that were about 1.5 times the general mean (i.e. still lower than mean yield in a mild year) and the poorest locations had average yields that were about 0.4 times the general mean, i.e. the gradient of environmental quality was much stronger in ‘hot and dry’ than in ‘mild’ years. This coincides with Figure 8A, which show a larger difference in phenotypic variances in more rainy places in ‘mild’ years.

Besides inspecting the gradients in average yield per year type, we examined the response surfaces for individual genotypes and focused on the yield difference between ‘hot and dry’ and ‘mild’ years across latitude and longitude for each genotype (Figure 9). Most genotypes showed a large yield reduction when comparing ‘hot and dry’ and ‘mild’ years. Only few of them (e.g. ‘g0_0.9_425’, Figure 9A) had a similar yield (or even larger yield for some locations) in ‘hot and dry’ than in ‘mild’ years, causing strong G×E within and between year types (Figure 9A). These exceptional genotypes were very early flowering, with small values for the three flowering time parameters and a very low mean yield (e.g. mean yield of ‘g0_0.9_425’ in mild years was 773.9 kg ha⁻¹, Figure 9A). Genotypes with small Eps values, but larger sensitivity to photoperiod (e.g. ‘g1.2_0.9_425’, Figure 9B) had much larger average yield than ‘g0_0.9_425’ (2053.5 vs 773.9 kg ha⁻¹ in mild years, Figure 9A) and showed an intermediate behaviour; maintaining or increasing yield in ‘hot and dry’ compared to ‘mild’ years in South-Eastern locations, and reducing yield in Western locations, in ‘hot and dry’ compared to ‘mild’ years (Figure 9). The locations that led to a yield increase in ‘hot and dry’ years for genotypes similar to ‘g0_0.9_425’ (Figure 9A) were characterized by lower temperatures and milder drought, compared to the Western locations (Figure 4), which are more prone to drought (Chenu et al., 2011 and 2013) and heat stress (Ababaei and Chenu, 2020).

Genotypes with large Eps values (e.g. ‘g0_0.9_1025’, Figure 9D; ‘g1.2_0.9_1025’, Figure 9E and ‘g1.2_1.7_1025’, Figure 9F) were much more sensitive to year type, with yield reductions between ‘mild’ and ‘hot and dry’ years of about 90% for some locations. For these genotypes, yield reduction in ‘hot and dry’ compared to ‘mild’ years was especially strong in Eastern locations, which have lower temperatures and milder drought (Figure 4). This apparent contradiction can be explained because genotypes with large Eps values express their larger yield
potential in ‘mild’ years and non-stressing conditions (i.e. Eastern locations). However, their yield is rapidly reduced under water limitation, heat and frost stress, and their high yield is no longer expressed in those locations during ‘hot and dry’ years, leading to a large proportional reduction. In contrast, Western locations, even for ‘mild’ years, have less favourable conditions than Eastern locations (Figure 3). Hence, their yield reduction between mild and ‘hot and dry’ years is less strong than for Eastern locations. In both year types, genotypes with intermediate Eps values had larger general mean than those with very large values for Eps, and had a smaller yield reduction between ‘mild’ and ‘hot and dry’ years (Figure 9G-I). Especially for those genotypes with intermediate values for photoperiod sensitivity (e.g. ‘g0.3_1.1_825’ and ‘g0.6_1.1_675’), the range of reduction was also smaller between locations, corresponding to more stable genotypes (Figures 7 and 9).

3.7. Within-year type location classification based on the five highest-yielding genotypes

We predicted the yield response surfaces for each genotype and year type. The yield predictions per pixel in the latitude-longitude grid were used to classify pixels, defining adaptation regions that share the same five highest-yielding genotypes. Figure 10 shows that in both year types, locations in the North-West had a different set of top genotypes than in the South-East. However, the pattern was more marked for ‘hot and dry’ than for ‘mild’ years. In ‘mild’ years (Figure 10A), most of the locations belonged to the same region and the clustering procedure clearly indicated only two regions. In contrast, the response surface in ‘hot and dry’ years had a more complex geographical separation and led to three regions (Figure 10B). This shows that in hot and dry years, which are becoming increasingly frequent because of climate change, the TPE becomes more heterogeneous and more GxE is expressed.

3.8. Location classification based on G+GxE of all genotypes for each year type

Besides classifying locations based on the five highest-yielding genotypes, we also applied a clustering method on the first two environmental GGE scores calculated on the yield data. In this strategy, the response surfaces of all genotypes are considered simultaneously. For both year types, we observe that the classification pattern (Figure 11) is similar to the one realized when focusing on the five highest-yielding genotypes (Figure 10) and it coincides with the geographical distribution of temperature and water stress (Figure 4); in both there is a diagonal pattern running parallel to the longitude boundaries (Figure 11), and approximating the rainfall patterns which are higher in the East than West of this region (Figure 4). The variance components for genotypes, genotype by region and residual GxE indicated that the relative importance of GxE, compared to the main effect was larger in ‘hot and dry’ years than in ‘mild’ years (Table 6A). In ‘hot and dry’ years, the regions obtained after clustering on G+GxE of the P-spline surfaces also explained a...
smaller part of the G×E variation, indicating that the G×E has a more complex structure in ‘hot and dry’ years than in ‘mild’ years (Table 6A). When estimating the APSIM parameter contribution to G×E across the predicted response surfaces per year type, it was clear that the genotype main effect was largely influenced by thermal time requirements between floral initiation and flowering time (Earliness-per-se, Table 6B). However, the relative importance was larger for ‘hot and dry’ years than for ‘mild’ years, coinciding with the genotype response surfaces across ‘mild’ and ‘hot and dry’ years (Figure 6) and the observations made in the GGE biplot (Figure 7). Figure 6, indicates that this main effect of Eps was related to intermediate values generally leading to larger yield (although the exact position of this optimum somehow changes depending on the level of water and temperature stress). In both year types, Eps:Ppd:Region (interaction between earliness-per-se, photoperiod sensitivity and location group in the environment landscape) and Eps:Vrn:Region (interaction between earliness-per-se, vernalization requirement and location group in the environment landscape) made important contributions to G×E variation. This coincides with the results shown in Figure 6, which indicate that in regions that are less drought-prone (like region 3, Figures 6 and 11), larger values of Ppd and Vrn lead to higher yield. In contrast, for regions that are more drought-prone (like regions 1 and 2, Figures 6 and 11), intermediate values for Ppd and Vrn lead to larger yield. This is especially the case for hot years, and in combination with large values for Eps. The residual G×E variation was much smaller than the total variation captured by the two- and three-way interactions between flowering time parameters and regions (Table 6B), indicating that the clustering of the G+G×E present in the predicted response surfaces was an effective method to capture the most important sources of spatial variation for G×E within year type.

4. Discussion

4.1. Simulated yield landscapes allow testing of statistical methods over space and time

In this paper, we illustrate how to fit response surfaces for individual genotypes, and how these surfaces can be used to decompose the structure of repeatable G×E. We also describe the adaptation landscape encompassed by the phenotypic responses ascribed to traits regulating phenology. The P-spline methodology presented here could be directly applied to real yield data from METs. However, the use of bio-physical simulations of yield has the advantage of generating phenotypes for a long series of years. We illustrated how to identify year type scenarios, placing the predicted response surfaces and adaptation landscapes in a long-term context. The 39 years of data that we included in this study showed that the frequency of heat and drought stress is increasing over time, likely associated to climate change (Ababaei and Chenu, 2020; Fletcher et al., 2020).
When comparing the year types resulting from the clustering procedure and the ENSO events (Potgieter et al., 2005), it can be observed that most of the ‘El Niño’ events correspond to ‘hot and dry’ years, whereas ‘La Niña’ events correspond to ‘mild’ years. Years that could not be classified as either ‘El Niño’ or ‘La Niña’ were classified as ‘hot and dry’ or ‘mild’ in a comparable proportion (11 hot and 13 mild years; supplementary Table 1). Importantly, the approach that we have taken here could only improve the predictability of adaptation surfaces \textit{a posteriori}, i.e. when knowing which kind of year is encountered. Nevertheless, the scientific community is making large efforts to improve weather forecasting (for example by the European Centre for Medium-Range Weather Forecasts, \url{https://www.ecmwf.int/}). Such forecasts would be very useful to inform breeders and farmers about the level of heat and water stress to be encountered by the crop in the upcoming season. An alternative strategy could be to undertake bivariate modelling of both year types simultaneously, borrowing information across year types. In that way, predictions could be made at each location for the most likely year type, weighing year type information by their relative frequencies. The fact that the frequency of occurrence of ‘hot and dry’ years has dramatically increased indicates a shift in the TPE. This needs to be addressed by selection strategies for future varieties, focusing on phenologies and sowing dates that are better targeted for hot years (Collins and Chenu, 2021). In that sense, our response surfaces predicted for hot years might be more informative to select for varieties that are well-adapted to future growing conditions, than the surfaces predicted for ‘mild’ years.

The simulation setting used here shows how to arrive at useful insights about the environmental drivers of adaptation and repeatable G×E in the variable climates of the NE Australian wheat-belt, but remains limited to one sowing date and a fixed initial soil condition. The use of genotype specific yield response surfaces in combination with adaptation landscapes per year type can also be used to provide a better understanding of G×E for different management conditions, different regions and also future growing conditions, in the context of climate change. Another potential application of the predicted surfaces could be to inform the decisions about location sampling for METs. Given that breeders usually need to limit the number of trials that are conducted, the predicted surfaces could be used to identify those locations that represent well the G×E that is to be expected in the TPE in the long term (Chauhan et al., 2017). An option to ensure that the METs are informative about adaptation across the TPE could be to always include trial locations that do induce different adaptive responses across mild and hot years, and trial locations without much variation between mild and hot years (more consistent genotype discrimination across years).

Our spline approach is implicitly a spatial emulation of APSIM output, comparable to the approach taken by Stanfill et al. (2015) in the sense that the P-splines response surfaces produce yield as a statistical function of the same genotype and environment specific inputs that are fed
into APSIM to produce yield. In principle, the response surfaces could have been directly produced by running APSIM simulations on a very dense grid across latitude and longitude. However, that requires massive computation time, and a dense grid for soil and weather if we want to accurately simulate yield variation. One simplification applied by Chapman et al. (2000c) used a weather grid in the mapping of variation in environment types, and ran the model multiple times for different soil types with the implication that environment types could be spatially referenced by knowing the latitude, longitude and the soil type. Our combined approach of APSIM simulations and spline models provides a more feasible and computationally efficient framework for a detailed characterization of genotypic responses, also applicable to real-world data sets. The main difference between both approaches (P-splines vs. APSIM output) is that P-splines by definition assume smooth gradients across latitude or longitude. However, there might be abrupt changes in environmental quality, for example due to changes in the soil quality resulting from the geological history of soil landscape. While the weather tends to vary spatially in some smooth way (associated with geography and interactions with weather systems), the soil does not vary smoothly, especially in a continent as old as Australia. Geology has, at extremely long time scales, created a landscape which, for any given region may comprise both abrupt (over 10s of metres) and gradual (over 100s of metres) changes in soil characteristics like water holding capacity. Abrupt changes are better addressed by crop growth models like APSIM, which contain explicit functions of environmental variables. A potential way to compare both methodologies would be to re-fit our P-spline models on a very dense grid of APSIM output and quantify the prediction error introduced by the discontinuities and abrupt changes that cannot be captured with the P-spline model. In such a case, additional fixed effects could also be introduced in the P-spline model to account for soil covariables that are present in a spatially discontinuous fashion.

In this proof of concept, we used a simulated population that was segregating only for flowering time parameters. This had the advantage that we could examine an adaptation landscape represented in a simple space of variation determined only by three traits, making it conceptually straightforward to understand. The range of variation for APSIM parameters regulating flowering time was very broad, leading to very large GxE. This GxE was also partially driven by the fact that, for a given location, all genotypes were sown on the same date. However, slow-maturing genotypes are sown early, and fast-maturing genotypes are sown late. Overall, the flowering time variation observed in this data set can be considered as an upper bound and it could be made narrower, to match the range of flowering time variation (1 to 3 weeks) that is commonly encountered in Australian wheat germplasm targeted for a given region (Zheng et al., 2012, 2013, 2016). The way in which the simulations cover the parameter space can modify the GxE patterns for the final trait. Therefore, it would be informative (i) to extend this approach by restricting the range of flowering time variation to that observed for well-adapted local germplasm and sowing
dates to inform growers about better adapted lines depending on the year type or ENSO events (Zheng et al., 2018) and (ii) to consider other genetically-varying APSIM parameters to assist breeding progress (Bustos-Korts et al., 2019b; Chapman et al., 2003; Hammer et al., 2014). The decision about which parameters should show variation to achieve G×E patterns that directly relate to the target population of genotypes depends on the traits for which the population is segregating and on the range of environmental conditions that is explored. On the other hand, the sensitivity of APSIM to changes of specific parameters defines the impact of those traits on G×E in the TPE (Casadebaig et al., 2016). Furthermore, for a final proof of the utility of this approach, it would be important to implement it on real data, like the MET data coming from pre-registration and post-registration trials.

4.2. Environment classification to look for spatial adaptation per year type

In this paper, we investigated two levels of environment classification. First, we classified years according to explicit environmental quality and we identified two groups; one for years with mild levels of temperature and water (drought) stress and a second group that consisted of years with higher temperatures and stronger water and frost stress. These groups captured part of the G×E associated to weather patterns, which is in general not very repeatable. As a second step, we fitted yield response surfaces, conditional on year type. These response surfaces provided insight on G×L, allowing us to classify locations according to the genotypic ranking across latitude and longitude. Within year type, these location classes correspond more directly to geographical and soil characteristics, which are more repeatable than across year types, allowing breeders to exploit specific adaptation. We observed that hot years are becoming increasingly frequent, coinciding with results reported by Ababaei and Chenu (2020). Hence, the response surfaces and adaptation landscapes generated for hot years become especially useful to select varieties that are adapted to future growing conditions. Such an approach can also be directly applied for projected future climate scenarios (e.g. Collins and Chenu, 2021; Watson et al., 2017). We also showed that there is an increased spatial heterogeneity in environmental quality during ‘hot and dry’ years, reflecting that the TPE is becoming more heterogeneous because of climate change. In the long run, this would require breeders to revisit the number and location of their METs because and increased heterogeneous TPE would also require an increased number of testing locations.

We assessed two methods to classify locations into regions; grouping locations with a similar set of genotypes in the top five of the genotypic ranking (Figure 10), and clustering locations based on the GGE scores (Figure 11). The main conceptual difference between both methods is the degree of importance attached to the highest-yielding genotype, compared to the whole set of genotypes. Both strategies led to similar groupings, with South-Western locations generally belonging to a different region, than those more continental locations. However, from the plant
breeding perspective, it might be more appealing to focus the attention on the highest-yielding genotypes as those are the ones that will determine genetic gain (Yan and Kang, 2002).

4.3. Characterization of GxE across the adaptation landscape

In this study, a) we compared several methods to understand the structure of GxE across latitude and longitude within part of the north-eastern Australian wheat-belt; b) we calculated yield differences between hot and mild years for each genotype, c) we used the predicted response surfaces for each genotype to identify the highest-yielding genotype per virtual location (pixel), d) we clustered on G+GxE generated by all genotypes across latitude and longitude, creating regions for each year type, and e) we quantified the contribution of APSIM parameters regulating flowering time to GxE across those regions. In general, all strategies supported the conclusion that locations in the South-East of the studied region have higher yield (Figure 8), but also higher yield fluctuations between year types, making a strong contribution to GxE in this part of the north-eastern Australian wheat-belt. It also became apparent that the spatial heterogeneity is larger under ‘hot and dry’ years, which are becoming increasingly frequent because of climate change.

We used simulated data to illustrate the approach. However, it would be interesting to confirm these findings in real phenotypic data that are coming from METs. For example, to investigate the efficiency of variety testing in VCU trial networks (VCU= value for cultivation and use). Such an analysis could give useful insight about the changes in variance and genotypic ranking across the whole target production area. If large changes in genotypic ranking occur, it is convenient to subdivide the TPE into smaller areas (Atlin et al., 2000; Piepho and Möhring, 2005), and to potentially select or recommend varieties within those smaller regions. The high spatial resolution of our predictions contain valuable information to support variety recommendation at the farmer’s level because predictions can be made for the exact location of fields (Collins et al., 2021), provided that accurate soil information at the farm level is available. In this way, MET information is more directly leveraged to the level of specific locations.

As the relative contribution of traits underlying yield severely changes depending on the environment conditions (Bustos-Korts et al., 2019; Chenu et al., 2009, 2013; Collins et al., 2021; Slafer et al., 2021), another aspect that could be further explored is the contribution of underlying traits to genotype adaptation. Here, we focused on grain yield predictions using P-Splines. A follow-up research could be to use bi-variate models for yield and flowering time (or other traits of interest). Such a model would help to explicitly characterize the contribution of flowering time to yield across the TPE.
4.4. Including explicit environmental covariables

Where differences in adaptation are not related to latitude and longitude, or when G\(\times\)E variation is not ascribed to specific locations and years (not very repeatable), as often occurs in Australian environments (Chapman, 2008; Chapman et al., 2000a; Chenu et al., 2011, 2013), an alternative is to express G\(\times\)E as function of explicit environmental covariables (Nicotra and Davidson, 2010). These covariables can correspond to weather and soil characteristics, or to environmental indices (Bustos-Korts et al., 2019; Heslot et al., 2013; Jarquín et al., 2013; Millet et al., 2019). The use of environmental covariables could also open interesting opportunities to model the year to year variation more explicitly. We now used environmental indices related to water and temperature stress to classify years into the categories ‘mild’ and ‘hot’. Although these year types do explain substantial G\(\times\)E variation, they are still internally heterogeneous. Hence, part of the G\(\times\)E information is lost when predicting response surfaces for each year type. An alternative would be to model yield variation as explicit function of latitude, longitude and an environmental covariables (that could also be, for example, an index that condenses the information contained in water and temperature-related variables). In that way, the model would make a more explicit use of the continuous variation that is also contained in the year to year variation.

4.5. Future developments

In this paper, we fitted spline surfaces to subsets of the data (per genotype, per year type, etc). In future work, we plan to extend the model, fitting it to all genotypes simultaneously. Such a model extension would provide parameter values that have been estimated simultaneously, which allows to borrow strength between genotypes. A second advantage of estimating parameters in the same model fit is that it can provide more insight into G\(\times\)E than using the predicted response surfaces alone. An additional development of the spline approach would be to fit surfaces for different soil types, and then to determine how these could be combined and re-fitted if more local soil data can be provided. The effect of sowing dates is also an important aspect that would need to be considered in a more detailed study of wheat adaptation. This is especially relevant as large interactions between rainfall and soil type are to be expected because of the differences in soil water retention capacity (Chenu et al., 2013). Further developments (already introduced above) could be the use of bivariate models considering either yield and underlying traits or yield in mild and hot years. Such bivariate approaches would have the advantage of borrowing strength across traits/year types and would provide useful insights in the interplay between traits in modulating adaptive responses across environments.
Conclusions

Fitting yield response surfaces for individual genotypes was useful to understand the structure of G×E and to predict genotype adaptation across the whole latitude and longitude range encompassed by the simulated METs.

The long-term simulations indicated the presence of two types of years, with different levels of water and temperature stress. These year types contributed significantly to G×E variation and for that reason it is advisable to predict response surfaces per year type separately.

The frequency of years with increased temperature and water stress is increasing in the most recent years, indicating that, to select varieties that are well-adapted to future growing conditions in a context of climate change, it might be advisable for breeders to base their selections on the predicted response surfaces for ‘hot’ years.

We used simulated data to illustrate the approach, but the spline methodology presented here can also be applied to real data from METs in breeding programs and VCU networks.

From the three APSIM parameters regulating phenology, the thermal time requirement between floral initiation and flowering had a large yield main effect (genotypes with larger requirements had higher yield, especially in mild years). Yield G×E interactions were mostly driven by specific combinations of $Eps$ (thermal time requirement between floral initiation and flowering) and $Ppd$ values. Our approach combining crop simulations and statistical models was useful to interpret the adaptation landscape as combinations of the APSIM phenology parameters.

Acknowledgements

In addition to funding by the author institutions, this research is supported through a co-investment (UOQ2003-011RTX for INVITA – A technology and analytics platform for improving variety selection) from the Grains Research and Development Corporation (GRDC) of Australia and by the Horizon 2020 Framework Programme of the European Union project INVITE under grant agreement No 817970 (https://www.h2020-invite.eu/).
References

Ababaei, B., and Chenu, K. (2020). Heat shocks increasingly impede grain filling but have little effect on grain setting across the Australian wheatbelt. *Agric. For. Meteorol.* 284, 107889. doi:10.1016/j.agrformet.2019.107889.

Atlin, G. N., Baker, R. J., McRae, K. B., and Lu, X. (2000). Selection Response in Subdivided Target Regions. *Crop Sci.* 40, 7–13. doi:10.2135/cropsci2000.4017.

Boer, M. P., Piepho, H. P., and Williams, E. R. (2020). Linear Variance, P-splines and Neighbour Differences for Spatial Adjustment in Field Trials: How are they Related? *J. Agric. Biol. Environ. Stat.* 25, 676–698. doi:10.1007/s13253-020-00412-4.

Brancourt-Hulmel, M., Denis, J. B., and Lecomte, C. (2000). Determining environmental covariates which explain genotype environment interaction in winter wheat through probe genotypes and biadditive factorial regression. 100, 285–298. doi:10.1007/s001220050038.

Bustos-Korts, D., Malosetti, M., Chapman, S., and van Eeuwijk, F. A. van (2016). “Modelling of Genotype by Environment Interaction and Prediction of Complex Traits across Multiple Environments as a Synthesis of Crop Growth Modelling, Genetics and Statistics,” in *Crop Systems Biology - Narrowing the Gaps between Crop Modelling and Genetics*, eds. X. Yin and P. C. Struik (Springer), 55–82. doi:10.1007/978-3-319-20562-5_3.

Bustos-Korts, D., Malosetti, M., Chenu, K., Chapman, S., Boer, M. P., Zheng, B., et al. (2019). From QTLs to adaptation landscapes: using genotype-to-phenotype models to characterize GxE over time. *Front. Plant Sci.* 10, 1540. doi:10.3389/fpls.2019.01540.

Bustos-Korts, D., Romagosa, I., Borràs-Gelonch, G., Casas, A. M., Slafer, G. A., and van Eeuwijk, F. (2018). “Genotype by Environment Interaction and Adaptation - Encyclopedia of Sustainability Science and Technology,” in *Encyclopedia of Sustainability Science and Technology*, ed. R. A. Meyers (New York, NY: Springer New York), 1–44. doi:10.1007/978-1-4939-2493-6_199-3.

Butler, D., Cullis, B., Gilmour, A., Gogel, B., and Thompson, R. (2017). ASReml-R Reference Manual. 176.

Casadebaig, P., Zheng, B., Chapman, S., Huth, N., Faivre, R., and Chenu, K. (2016). Assessment of the potential impacts of plant traits across environments by combining global sensitivity analysis and dynamic modeling in wheat. *PLoS One* 11, 1–27. doi:10.1371/journal.pone.0146385.

Chapman, S. (2008). Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. *Euphytica* 161, 195–208.
Chapman, S. C., Cooper, M., Butler, D. G., and Henzell, R. G. (2000a). Genotype by environment interactions affecting grain sorghum. I. Characteristics that confound interpretation of hybrid yield. *Aust. J. Agric. Res.* 51, 197–208. doi:http://dx.doi.org/10.1071/AR99020.

Chapman, S. C., Cooper, M., Hammer, G. L., and Butler, D. G. (2000b). Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. *Aust. J. Agric. Res.* 51, 209–222. doi:http://dx.doi.org/10.1071/AR99021.

Chapman, S. C., Hammer, G. L., Butler, D. G., and Cooper, M. (2000c). Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population of environments. *Aust. J. Agric. Res.* 51, 223–234. doi:http://dx.doi.org/10.1071/AR99022.

Chapman, S. C., Hammer, G. L., Podlich, D. W., and Cooper, M. (2002). “Linking bio-physical and genetic models to integrate physiology, molecular biology and plant breeding,” in *Quantitative Genetics, Genomics, and Plant Breeding*, ed. M. Kang (Wallingford UK: CAB International), 167–187.

Chapman, S., Cooper, M., Podlich, D., and Hammer, G. (2003). Evaluating Plant Breeding Strategies by Simulating Gene Action and Dryland Environment Effects. *Agron. J.* 95, 99–113. doi:10.2134/agronj2003.9900.

Chauhan, Y., Allard, S., Williams, R., Williams, B., Mundree, S., Chenu, K., et al. (2017). Characterisation of chickpea cropping systems in Australia for major abiotic production constraints. *F. Crop. Res.* 204, 120–134. doi:10.1016/j.fcr.2017.01.008.

Chenu, K. (2015). “Characterizing the crop environment – nature, significance and applications,” in *Crop Physiology*, eds. V. O. Sadras and D. F. Calderini (London: Elsevier Inc.), 105–122. doi:10.1016/B978-0-12-417104-6/00013-3.

Chenu, K., Chapman, S. C., Tardieu, F., McLean, G., Welcker, C., and Hammer, G. L. (2009). Simulating the Yield Impacts of Organ-Level Quantitative Trait Loci Associated With Drought Response in Maize: A “Gene-to-Phenotype” Modeling Approach. *Genetics* 183, 1507–1523. doi:10.1534/genetics.109.105429.

Chenu, K., Cooper, M., Hammer, G. L., Mathews, K. L., Dreccer, M. F., and Chapman, S. C. (2011). Environment characterization as an aid to wheat improvement: interpreting genotype–environment interactions by modelling water-deficit patterns in North-Eastern Australia. *J. Exp.
Chenu, K., Dehimfard, R., and Chapman, S. C. (2013). Large-scale characterization of drought pattern: a continent-wide modelling approach applied to the Australian wheatbelt – spatial and temporal trends. *New Phytol.* 198, 801–820. doi:10.1111/nph.12192.

Chenu, K., Porter, J. R., Martre, P., Basso, B., Chapman, S. C., Ewert, F., et al. (2017). Contribution of Crop Models to Adaptation in Wheat. *Trends Plant Sci.* 22, 472–490. doi:10.1016/j.tplants.2017.02.003.

Collins, B., Chapman, S., Hammer, G., and Chenu, K. (2021). Limiting transpiration rate in high evaporative demand conditions to improve Australian wheat productivity. *in silico Plants* 3, 1–16. doi:10.1093/insilicoplants/diab006.

Collins, B., and Chenu, K. (2021). Improving productivity of Australian wheat by adapting sowing date and genotype phenology to future climate. *Clim. Risk Manag.* 32, 100300. doi:10.1016/j.crm.2021.100300.

Comstock, R. E., and Moll, R. H. (1963). “Genotype-Environment Interactions,” in *Statistical genetics and plant breeding: A Symposium and Workshop*, eds. W. D. Hanson and H. F. Robinson (National Academy of Sciences-National Research Council), 164–196.

Cooper, M., Messina, C. D., Podlich, D., Totir, L. R., Baumgarten, A., Hausmann, N. J., et al. (2014). Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop Pasture Sci.* 65, 311–336. doi:http://dx.doi.org/10.1071/CP14007.

Cullis, B. R., Thomson, F. M., Fisher, J. A., Gilmour, A. R., and Thompson, R. (1996). The analysis of the NSW wheat variety database. II. Variance component estimation. *Theor Appl Genet.* 92, 28–39. doi:10.1007/bf00222948.

Eilers, P. H. C., and Marx, B. D. (1996). Flexible Smoothing with B-splines and Penalties. *Stat. Sci.* 11, 89–102. doi:10.2307/2246049.

Eilers, P. H. C., Marx, B. D., and Durbán, M. (2015). Twenty years of P-splines. *SORT-Statistics Oper. Res. Trans.* 39, 149–186.

Fletcher, A. L., Chen, C., Ota, N., Lawes, R. A., and Oliver, Y. M. (2020). Has historic climate change affected the spatial distribution of water-limited wheat yield across Western Australia? *Clim. Change* 159, 347–364. doi:10.1007/s10584-020-02666-w.

Graffelman, J., and Van Eeuwijk, F. (2005). Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research. *Biometrical J.* 47,
Hammer, G. L., McLean, G., Chapman, S., Zheng, B., Doherty, A., Harrison, M. T., et al. (2014). Crop design for specific adaptation in variable dryland production environments. *Crop Pasture Sci.* 65, 614–626. Available at: http://dx.doi.org/10.1071/CP14088.

Hammer, G., Messina, C., Wu, A., and Cooper, M. (2019). Opinion Biological reality and parsimony in crop models — why we need both in crop improvement! 1, 1–21. doi:10.1093/insilicoplants/diz010.

Harrison, M. T., Tardieu, F., Dong, Z., Messina, C. D., and Hammer, G. L. (2014). Characterizing drought stress and trait influence on maize yield under current and future conditions. *Glob. Chang. Biol.* 20, 867–878. doi:10.1111/gcb.12381.

Heslot, N., Akdemir, D., Sorrells, M., and Jannink, J.-L. (2013). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 1, 1–18. doi:10.1007/s00122-013-2231-5.

Holzworth, D. P., Huth, N. I., DeVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., et al. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Env. Model. Soft.* 62, 327–350. doi:http://dx.doi.org/10.1016/j.envsoft.2014.07.009.

Jarquín, D., Crossa, J., Lacaze, X., Cheyron, P., Daucourt, J., Lorgeou, J., et al. (2013). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor Appl Genet.* 3, 1–13. doi:10.1007/s00122-013-2243-1.

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

Lobell, D. B., Hammer, G. L., Chenu, K., Zheng, B., Mclean, G., and Chapman, S. C. (2015). The shifting influence of drought and heat stress for crops in northeast Australia. *Glob. Chang. Biol.* 21, 4115–4127. doi:10.1111/gcb.13022.

Lowry, D. B., Lovell, J. T., Zhang, L., Bonnette, J., Fay, P. A., Mitchell, R. B., et al. (2019). QTL × environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. *Proc. Natl. Acad. Sci. U. S. A.* 116, 12933–12941. doi:10.1073/pnas.1821543116.

Malosetti, M., Ribaut, J.-M., and van Eeuwijk, F. A. (2013). The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* 4, 1–17. doi:10.3389/fphys.2013.00044.
Malosetti, M., Voltas, J., Romagosa, I., Ullrich, S. E., and van Eeuwijk, F. A. (2004). Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* 137, 139–145. doi:10.1023/B:EUPH.0000040511.46388.ef.

Messina, C. D., Jones, J. W., Boote, K. J., and Vallejos, C. E. (2006). A gene-based model to simulate soybean development and yield responses to environment. *Crop Sci.* 46, 456–466. doi:10.2135/cropsci2005.04-0372.

Messina, C. D., Podlich, D., Dong, Z., Samples, M., and Cooper, M. (2011). Yield–trait performance landscapes: from theory to application in breeding maize for drought tolerance. *J Exp. Bot.* 62, 855–868. doi:10.1093/jxb/erq329.

Millet, E. J., Kruijer, W., Coupel-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., et al. (2019). Genomic prediction of maize yield across European environmental conditions. *Nat. Genet.* 51. doi:10.1038/s41588-019-0414-y.

Millet, E., Welcker, C., Kruijer, W., Negro, S., Nicolas, S., Praud, S., et al. (2016). Genome-wide analysis of yield in Europe: allelic effects as functions of drought and heat scenarios. *Plant Physiol.* 172, pp.00621.2016. doi:10.1104/pp.16.00621.

Nicotra, A. B., and Davidson, A. (2010). Adaptive phenotypic plasticity and plant water use. *Funct. Plant Biol.* 37, 117–127. Available at: http://dx.doi.org/10.1071/FP09139.

Parent, B., Bonneau, J., Maphosa, L., Kovalchuk, A., Langridge, P., and Fleury, D. (2017). Quantifying Wheat Sensitivities to Environmental Constraints to Dissect Genotype × Environment Interactions in the Field. *Plant Physiol.* 174, 1669–1682. doi:10.1104/pp.17.00372.

Piepho, H.-P., Boer, M. P., and Williams, E. R. (2021). Tensor P-Spline Smoothing for Spatial Analysis of Plant Breeding Trials. *bioRxiv*, 2021.05.10.443463. doi:10.1101/2021.05.10.443463.

Piepho, H.-P., Laidig, F., Drobek, T., and Meyer, U. (2014). Dissecting genetic and non-genetic sources of long-term yield trend in German official variety trials. *Theor Appl Genet.* 127, 1009–1018. doi:10.1007/s00122-014-2275-1.

Piepho, H. P., and Möhring, J. (2005). Best Linear Unbiased Prediction of Cultivar Effects for Subdivided Target Regions. *Crop Sci.* 45, 1151–1159. doi:10.2135/cropsci2004.0398.

Potgieter, A. B., Hammer, G. L., Meinke, H., Stone, R. C., and Goddard, L. (2005). Three putative types of El Niño revealed by spatial variability in impact on Australian wheat yield. *J. Clim.* 18, 1566–1574. doi:10.1175/JCLI3349.1.

R Core Team (2019). R: A Language and Environment for Statistical Computing. Available at:
Rodríguez-Álvarez, M. X., Boer, M. P., van Eeuwijk, F. A., and Eilers, P. H. C. (2018). Correcting for spatial heterogeneity in plant breeding experiments with P-splines. Spat. Stat. 23, 52–71. doi:10.1016/j.spasta.2017.10.003.

Rodríguez-Álvarez, M. X., Lee, D. J., Kneib, T., Durbán, M., and Eilers, P. (2015). Fast smoothing parameter separation in multidimensional generalized P-splines: the SAP algorithm. Stat. Comput. 25, 941–957. doi:10.1007/s11222-014-9464-2.

Slafer, G. A., Savin, R., Pinochet, D., and Calderini, D. F. (2021). “Chapter 3 - Wheat,” in, eds., V. O. Sadras and D. F. B. T.-C. P. C. H. for M. C. Calderini (Academic Press), 98–163. doi:https://doi.org/10.1016/B978-0-12-819194-1.00003-7.

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

Stanfill, B., Mielenz, H., Clifford, D., and Thorburn, P. (2015). Simple approach to emulating complex computer models for global sensitivity analysis. Environ. Model. Softw. 74, 140–155. doi:10.1016/j.envsoft.2015.09.011.

van Eeuwijk, F. A., Bustos-Korts, D., Millet, E. J., Boer, M. P., Kruijer, W., Thompson, A., et al. (2019). Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. Plant Sci. 282, 23–39. doi:10.1016/j.plantsci.2018.06.018.

van Eeuwijk, F. A., Bustos-Korts, D. V, and Malosetti, M. (2016). What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype × Environment Interactions? Crop Sci. 56, 2119–2140. doi:10.2135/cropsci2015.06.0375.

van Eeuwijk, F. A., Malosetti, M., Yin, X., Struik, P. C., and Stam, P. (2005). Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. Aust. J. Agric. Res. 56, 883–894. doi:http://dx.doi.org/10.1071/AR05153.

Velazco, J. G., Rodríguez-Álvarez, M. X., Boer, M. P., Jordan, D. R., Eilers, P. H. C., Malosetti, M., et al. (2017). Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model. Theor. Appl. Genet. 130, 1375–1392. doi:10.1007/s00122-017-2894-4.

Voltas, J., van Eeuwijk, F. A., Araus, J. L., and Romagosa, I. (1999). Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley II: Grain growth. F. Crop. Res. 62, 75–84. doi:http://dx.doi.org/10.1016/S0378-4290(99)00007-6.
Wang, B., Liu, D. L., Asseng, S., Macadam, I., and Yu, Q. (2017). Modelling wheat yield change under CO2 increase, heat and water stress in relation to plant available water capacity in eastern Australia. *Eur. J. Agron.* 90, 152–161. doi:10.1016/j.eja.2017.08.005.

Watson, J., Zheng, B., Chapman, S., and Chenu, K. (2017). Projected impact of future climate on water-stress patterns across the Australian wheatbelt. *J. Exp. Bot.* 68, 5907–5921. doi:10.1093/jxb/erx368.

Williams, E. R. (1996). A Neighbour Model for Field Experiments. *Biometrika* 73, 279–287. Available at: https://www.jstor.org/stable/2335817.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition*. CRC Press Available at: https://books.google.be/books?id=HL-PDwAAQBAJ.

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

Yan, W., and Kang, M. S. (2002). *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists*. CRC Press Available at: https://books.google.nl/books?id=Bz2SpU-xgnkC.

Yan, W., and Rajcan, I. (2002). Biplot Analysis of Test Sites and Trait Relations of Soybean in Ontario. *Crop Sci.* 42, 11–20. doi:10.2135/cropsci2002.1100.

Zelterman, D. (2015). *Applied Multivariate Statistics with R*. Springer International Publishing Available at: https://books.google.nl/books?id=tsBOCgAAQBAJ.

Zheng, B., Biddulph, B., Li, D., Kuchel, H., and Chapman, S. (2013). Quantification of the effects of VRN1 and Ppd-D1 to predict spring wheat (Triticum aestivum) heading time across diverse environments. *J. Exp. Bot* 64, 3747–3761. doi:10.1093/jxb/ert209.

Zheng, B., Chapman, S. C., Christopher, J. T., Frederiks, T. M., and Chenu, K. (2015a). Frost trends and their estimated impact on yield in the Australian wheatbelt. *J. Exp. Bot.* 66, 3611–3623. doi:10.1093/jxb/erv163.

Zheng, B., Chapman, S., and Chenu, K. (2018). The Value of Tactical Adaptation to El Niño–Southern Oscillation for East Australian Wheat. *Climate* 6, 77. doi:10.3390/cli6030077.

Zheng, B., Chenu, K., and Chapman, S. C. (2016). Velocity of temperature and flowering time in wheat - assisting breeders to keep pace with climate change. *Glob. Chang. Biol.* 22, 921–933. doi:10.1111/gcb.13118.

Zheng, B., Chenu, K., Doherty, A., and Scott, C. (2015b). The APSIM-Wheat Module (7.5 R3008).
Zheng, B., Chenu, K., Fernanda Dreccer, M., and Chapman, S. C. (2012). Breeding for the future: what are the potential impacts of future frost and heat events on sowing and flowering time requirements for Australian bread wheat (Triticum aestivium) varieties? *Glob. Chang. Biol.* 18, 2899–2914. doi:10.1111/j.1365-2486.2012.02724.x.
Table 1. Location name, latitude, longitude and sowing date for the 13 locations considered in this study with their corresponding mean simulated yield and date to flowering together with their corresponding standard deviations across genotypes and years. Data simulated for 156 genotypes differing in phenology, over 1978-2016. QLD, Queensland; NSW, New South Wales.

| State | Region                  | Name       | Latitude | Longitude | Sowing date | Days to flowering | Yield (kg ha⁻¹) |
|-------|-------------------------|------------|----------|-----------|-------------|-------------------|----------------|
| QLD   | Eastern Darling Downs   | Dalby      | -27.18   | 151.26    | 29-May      | 114.1             | 1 593.3        |
| QLD   | Western Darling Downs  | Meandarra  | -27.32   | 149.88    | 7-May       | 115.7             | 1 578.6        |
| QLD   | Western Darling Downs  | Goondiwindi| -28.55   | 150.31    | 11-Apr      | 107.2             | 2 000.8        |
| NSW   | Northern NSW           | Moree      | -29.48   | 149.84    | 5-May       | 121.7             | 2 141.6        |
| NSW   | Northern NSW           | Walgett    | -30.04   | 148.12    | 5-May       | 121.5             | 1 205.7        |
| NSW   | Northern NSW           | Narrabri   | -30.32   | 149.78    | 21-Apr      | 121.0             | 2 005.5        |
| NSW   | Northern NSW           | Cootamundra| -30.98   | 148.38    | 21-Apr      | 124.3             | 1 838.7        |
| NSW   | Eastern NSW            | Gunnedah   | -30.98   | 150.25    | 5-May       | 127.7             | 2 504.1        |
| NSW   | Western NSW            | Nyngan     | -31.55   | 147.20    | 17-Apr      | 123.4             | 1 959.7        |
| NSW   | Western NSW            | Gilgandra  | -31.71   | 148.66    | 25-Apr      | 143.4             | 2 115.9        |
| NSW   | Western NSW            | Dubbo      | -32.24   | 148.61    | 29-Apr      | 140.9             | 2 073.4        |
| NSW   | Eastern NSW            | Wellington | -32.80   | 148.80    | 23-Apr      | 156.1             | 2 593.8        |
| NSW   | Western NSW            | Condobolin | -33.07   | 147.23    | 11-May      | 139.4             | 1 447.3        |
Table 2. Description of the environmental indices calculated with APSIM output and that were used to classify environments according to their levels of water and temperature stress. The four indices were calculated in a period from flowering + 100°C to flowering + 600°C.

| Name environmental index | Description                                                                 |
|--------------------------|-----------------------------------------------------------------------------|
| S2_sum.rain              | Sum of rainfall                                                             |
| S2_frost.sum             | Accumulated thermal time when minimum temperature is less than 0           |
| S2_vpd                   | Average of daily vapour pressure deficit, as calculated with APSIM         |
| S2_avg.maxt              | Average of daily maximum temperatures                                       |
Table 3. Variance components for the simulated yield data of 156 genotypes over 39 years at 13 locations (in total 503 year-location combinations because four of them were removed due to crop failure).

| Component     | Variance  | Std.error |
|---------------|-----------|-----------|
| Geno          | 86 559    | 10 380    |
| Geno:Loc      | 37 528    | 1 408     |
| Geno:Year     | 60 023    | 1 407     |
| Residual      | 208 106   | 1 115     |
Table 4. Variance components for the simulated yield data of 156 genotypes over 39 years at 13 locations (total of 503 year-location combinations) based on genotype and genotype interactions with location or year type.

| Component                  | Variance | Std.error |
|----------------------------|----------|-----------|
| Geno                       | 71 280   | 10 414    |
| Geno:Loc                   | 29 975   | 1 461     |
| Geno: Type                 | 28 852   | 3 764     |
| Geno:Year within Type      | 45 304   | 1 140     |
| Geno:Loc:Type              | 14 965   | 834       |
| Residual                   | 200 468  | 1 088     |
Table 5. Contribution of the APSIM traits regulating phenology to yield variation. Variance components were estimated by fitting the statistical model to 156 genotypes at 13 locations over 39 years.

| Term      | Component | Std error |
|-----------|-----------|-----------|
| Ppd       | 49 610    | 35 182    |
| Vrn       | 5 538     | 3 971     |
| Eps       | 31 056    | 12 783    |
| Ppd:env   | 62 232    | 2 287     |
| Vrn:env   | 25 992    | 1 178     |
| Eps:env   | 97 832    | 2 355     |
| Ppd:Vrn:env | 4 593 | 232       |
| Ppd:Eps:env | 69 719 | 788       |
| Vrn:Eps:env | 68 178 | 798       |
| Residual  | 22 221    | 232       |
Table 6. Variance components estimated for the APSIM yield in the 13 locations for 156 genotypes varying in APSIM parameter relative to photoperiod sensitivity (PPD), vernalization requirement (Vrn) and thermal time from floral initiation to flowering (earliness per se; Eps). The analysis was made separately for ‘mild’ and ‘hot and dry’ years. For each year type, locations were classified into regions by clustering on spline-predicted yields for the full set of genotypes (Figure 10).

| A | Component          | Mild years | Std error | Hot and dry years | Std error |
|---|--------------------|------------|-----------|-------------------|-----------|
|   | Cultivar           | 74 438     | 45 892    |                   |           |
|   | Cultivar:Region    | 102 972    | 14 619    |                   |           |
|   | Residual G×E       | 241 079    | 278 153   |                   |           |

| B | Component          | Mild years | Std error | Hot and dry years | Std error |
|---|--------------------|------------|-----------|-------------------|-----------|
|   | Ppd                | 52 472     | 55 300    | 9 549             | 8 508     |
|   | Vrn                | 793        | 5 762     | 4 196             | 3 504     |
|   | Eps                | 22 779     | 19 704    | 23 553            | 10 498    |
|   | Ppd:Region         | 41 274     | 31 722    | 5 090             | 3 520     |
|   | Vrn:Region         | 7 142      | 7 299     | 355               | 858       |
|   | Eps:Region         | 30 429     | 17 015    | 0                 | NA        |
|   | Vrn:Ppd:Region     | 1 669      | 1 054     | 462               | 329       |
|   | Eps:Ppd:Region     | 28 450     | 4 336     | 18 481            | 2 385     |
|   | Eps:Vrn:Region     | 25 787     | 3 942     | 11 607            | 1 621     |
|   | Residual G×E       | 240 870    | 1 712     | 277 491           | 2 031     |
Figure 1

- **Genotypes**
  - 156 APSIM parameter combinations
    - Ppd (0 to 1.2)
    - Vrn (0.9 to 1.7)
    - Eps (425 to 1025)

- **Environments**
  - 13 locations
  - 39 years (1978 to 2016)
  - Soil and weather variables

**APSIM output**

- **Yield**
  - Boxplots
  - Variance components models
  - GGE biplots

- **Environmental indices**
  - Year classification based on temperature and water-related indices: 2 Year Types (‘mild’ and ‘hot and dry’)

---

**Conditional on Year Type (mild / hot and dry)**

- Fit a latitude – longitude yield response surface and use that to create high resolution predictions for each genotype

- **Classify high resolution yield predictions into geographical adaptation zones**
  - (based on P-spline predicted surfaces)

  **Method 1**
  - Classify on the basis of identity of the five highest-yielding genotypes

  **Method 2**
  - Classify on the basis of performance of all genotypes
Figure 7

A) Ppd, mild years
GGE biplot for yield (genotype scaling)

B) Ppd, HD years
GGE biplot for yield (genotype scaling)

C) Vrn, mild years
GGE biplot for yield (genotype scaling)

D) Vrn, HD years
GGE biplot for yield (genotype scaling)

E) Earliness-per-se, mild years
GGE biplot for yield (genotype scaling)

F) Earliness-per-se, HD years
GGE biplot for yield (genotype scaling)
Figure 8

A Mean per pixel – general mean (Mild years)

B Mean per pixel – general mean (HD years)
Small requirement for thermal time between floral initiation and flowering (earliness-per-se)

A) g0_0.9_425, HD- mild year

B) g1.2_0.9_425, HD- mild year

C) g1.2_1.7_425, HD- mild year

Large requirement for thermal time between floral initiation and flowering (earliness-per-se)

D) g0_0.9_1025, HD- mild year

E) g1.2_0.9_1025, HD- mild year

F) g1.2_1.7_1025, HD- mild year

Medium requirement for thermal time between floral initiation and flowering (earliness-per-se)

G) gG0.3_1.1_625, HD- mild year

H) g0.6_1.1_675, HD- mild year

I) g1.2_1.7_675, HD- mild year
Figure 10

![Cluster Analysis Diagram](https://academic.oup.com/insilicoplants/advance-article/doi/10.1093/insilicoplants/diab018/6316219)
