Australian & New Zealand Journal of Statistics

Aust. N. Z. J. Stat. 2022 doi: 10.1111/anzs.12362

Visualising the pattern of long-term genotype performance by leveraging a genomic prediction model

Vivi N. Arief[1], Ian H. DeLacy[1], Thomas Payne[2] and Kaye E. Basford[1*]

The University of Queensland and International Maize and Wheat Improvement Center (CIMMYT)

Summary

Historical data from plant breeding programs provide valuable resources to study the response of genotypes to the changing environment (i.e. genotype-by-environment interaction). Such data have been used to evaluate the pattern of genotype performance across regions or locations, but its use to evaluate the long-term pattern of genotype performance across environments (i.e. locations-by-years) has been hampered by the lack of common genotypes across years. This lack of common genotypes is due to the structure of the breeding program, especially for annual crops, where only a proportion of selected genotypes are tested in subsequent years. This has resulted in a sparse prediction of the performance of genotypes across years (i.e. a genotype-by-year table). A genomic prediction method that fitted both a relationship matrix among genotypes and a relationship matrix among environments (i.e. years) could overcome this limitation and produce a dense genotype-by-year table, thereby enabling some evaluation of long-term genotype performance. In this paper, we applied the genomic prediction model to the yield data from CIMMYT's Elite Spring Wheat Yield Trials (ESWYT) to visualise the pattern of genotype performance over 25 years.

Key words: factor analytic; genotype-by-year; historical data; plant breeding; relationship matrix

1. Introduction

The primary goal of any plant breeding program is to identify elite genotypes by selecting the best performing genotypes from each breeding cycle. Selection is commonly made in stages. In each stage, genotypes are grown in multi-location field trials. However, these genotypes are usually only tested in a plant breeding program for a limited number of years.

Historical data from plant breeding programs have been used to evaluate the pattern of performance of the breeding germplasm over a range of observed environments (e.g. Cooper & DeLacy 1994; DeLacy, Ratnasiri & Mirzawan 1996; DeLacy *et al.* 2000). These data can and have been used to assess genotype performance across regions or locations

*Author to whom correspondence should be addressed.
[1]School of Agriculture and Food Sciences, The University of Queensland, Brisbane, QLD 4072, Australia
[2]International Maize and Wheat Improvement Center (CIMMYT), Apdo Postal 041, C.A.P Plaza Galerías, Col. Verónica Anzures, Ciudad de México, 11305, México. e-mail: k.e.basford@uq.edu.au

Acknowledgements. We thank the reviewer for helpful comments and suggestions. Opinions and attitudes expressed in this document, which are not explicitly designated as Journal policy, are those of the author and are *not* necessarily endorsed by the Journal, its editorial board, its publisher Wiley or by the Australian Statistical Publishing Association Inc.

(e.g. Cooper & DeLacy 1994; DeLacy, Ratnasiri & Mirzawan 1996; DeLacy *et al.* 2000). However, they often cannot be used to evaluate genotype performance across years due to the structure of the plant breeding program, where a limited number of common genotypes are tested across years.

When analysing data from a plant breeding program using a mixed model, a relationship matrix among genotypes can be fitted to obtain a better prediction of genotype performance (e.g. Crossa *et al.* 2006; Oakey *et al.* 2006, 2007; Burgueño *et al.* 2007). This approach also enables the prediction of a genotype's performance from its relatives (e.g. Oakey *et al.* 2007; Crossa *et al.* 2010). This provides a better connection across environments (which can be years, locations or years $\times$ locations), especially when the environments involve years. A relationship matrix among environments can also be fitted in the mixed model to obtain a better prediction of genotype performance (Smith *et al.* 2015). The genomic prediction model includes both a relationship matrix among genotypes and a relationship matrix among environments (Burgueño *et al.* 2012; Jarquín *et al.* 2014; Saint Pierre *et al.* 2016).

Fitting both a relationship matrix among genotypes and a relationship matrix among years (a component of the environments) allows an evaluation of the pattern of performance of the breeding germplasm over the range of observed environments in the past and in the present. Therefore, it extends the period for which the breeding material is evaluated and enables an assessment of long-term performance. As the characteristics of these past environments are usually also known, the performance of genotypes in these environments can be used to predict their performance in future environments that share characteristics with these past environments.

In this paper, the use of the genomic prediction model to evaluate the long-term genotype performance is demonstrated by analysing the data on the lines tested in the first 25 cycles/years (1979–2005) of the CIMMYT's Elite Spring Wheat Yield Trials (ESWYT). Our methodology takes advantage of the statistical prediction of responses in those cycles/years in which the genotypes were not grown and the visualisation of the results via pattern analysis—a combination of clustering and ordination techniques which are described in Manly (1994) and used in McLachlan & Basford (1988). Interpretation of the prediction of genotype performance should be undertaken prudently and used as a guide in any plant breeding program.

2. Methodology

2.1. Phenotypic data

Yield data were collected from the first 25 cycles of CIMMYT's Elite Spring Wheat Yield Trials (ESWYT). Each cycle of ESWYT represents 1 year of testing, and hereafter the terms year and cycle are used interchangeably. Each ESWYT cycle tested 30 to 50 elite wheat lines in a trial across regions of the world where spring wheat is grown. There are no or a limited number of common lines across these cycles. The term 'line' is used to refer to a homogenous genotype, and hereafter the more general term 'genotype' will be used instead of 'line'.

The data used here were obtained from trials established during the period 1979 to 2005 (no ESWYT was distributed in 1993). In total, they comprised yields (in tonne/ha)

collected from 25 cycles in which various subsets (of 30 to 50) of 686 genotypes were grown in 1,387 trials across diverse choices (of 17 to 91) from 382 worldwide locations. Further details on the ESWYT data are available in Arief *et al.* (2015).

2.2. Predicting genotype performance in each year

The analysis of the yields was conducted using REML (Restricted Maximum Likelihood, Patterson & Thompson 1971) as implemented in the ASReml software (Gilmour *et al.* 2021). Data validation and error checking were conducted prior to any analysis (Arief 2010). Due to the large amount of data, a two-stage analysis approach was used (e.g. Cullis *et al.* 1996; Smith, Cullis & Gilmour 2001). Two-stage analysis can handle considerably more data and is computationally less demanding than a one-stage analysis (Piepho *et al.* 2012; Cullis *et al.* 1996; Frensham, Cullis & Verbyla 1997; Smith, Cullis & Gilmour 2001), but it is less efficient than a one-stage analysis (Welham *et al.* 2010; Piepho *et al.* 2012). Some efficiency in a second-stage analysis can be recovered by using weights (e.g. Cullis *et al.* 1996; Mohring & Piepho 2009; Piepho *et al.* 2012).

The first-stage analysis was conducted for each trial by fitting the design factors as random and the genotype factor as fixed. The experimental design was a randomised complete block design for ESWYT 1 to 13, and an α-lattice design for all subsequent ESWYT cycles. The outputs from the first stage, the BLUE and weight for each genotype in each trial (Cullis *et al.* 1996), were used in the second-stage analysis to obtain the prediction of the genotype performances in each of the 25 years/cycles using the following model:

$$y_{ijk(q)} = \mu + t_q + (gy)_{ij} + (gyl)_{ijk} + \epsilon_{ijk(q)},$$

where $y_{ijk(q)}$ is the BLUE for genotype i in trial q in year j and location k; μ is the intercept; t_q is the effect of trial q ($q = 1, \ldots, Q$, here $Q = 1387$); $(gy)_{ij}$ is the effect of genotype i in year j ($i = 1, \ldots, G$, here $G = 686$; $j = 1, \ldots, Y$, here $Y = 25$) with $\boldsymbol{u}_1 \sim \mathrm{N}(0, \sigma^2_{gy}\boldsymbol{A}_{gy})$, where $\boldsymbol{u}_1$ is a vector of all $(gy)_{ij}$ terms, $\boldsymbol{A}_{gy}$ is a Kronecker product of a relationship matrix among the genotypes ($\boldsymbol{K}_{G\times G}$) and a relationship matrix among years ($\boldsymbol{W}_{Y\times Y}$); $(gyl)_{ijk}$ is the interaction effect of genotype i $\times$ year j $\times$ location k ($k = 1, \ldots, L$, here $L = 382$) with $\boldsymbol{u}_2 \sim \mathrm{N}(0, \sigma^2_{gyl}\boldsymbol{I})$, where $\boldsymbol{u}_2$ is a vector of all $(gyl)_{ijk}$ terms and $\boldsymbol{I}_{GYL\times GYL}$ is an identity matrix; and $\epsilon_{ijk(q)}$ is the effect of residual with $\boldsymbol{e} \sim \mathrm{N}(0, \sigma^2_{\epsilon}\boldsymbol{R})$, where $\boldsymbol{e}$ is a vector of all $\epsilon_{ijk(q)}$, and $\boldsymbol{R}_{N\times N}$ is a diagonal matrix of the weights for the $y_{ijk(q)}$ ($N = 54,850$). All terms, except for μ and t_q, were random.

The objective of this analysis was to obtain the genotype-by-year array by fitting $(gy)_{ij}$ However, because the exact weights were used in this second-stage analysis, the term $(gyl)_{ijk}$ was required to be included in the model to account for the extra variance. In a one-stage analysis, this extra term is not needed as the extra variance is usually captured by the residual variance.

The predictions of the performances of the genotypes in each year, $(gy)_{ij}$, can be obtained by fitting several forms of the $\boldsymbol{A}_{gy}$ matrix, each with a different level of complexity:

1. $\boldsymbol{A}_{gy} = \boldsymbol{I}_{G\times G} \otimes \boldsymbol{I}_{Y\times Y}$, where $\boldsymbol{I}_{G\times G}$ and $\boldsymbol{I}_{Y\times Y}$ are the identity matrices for genotypes and for years, respectively, and $\otimes$ is a Kronecker product. This is the simplest model, assuming no relationship among genotypes and no relationship among years. This model can only predict the performance of the genotypes in the year of testing; hence a sparse genotype-by-year table is produced.

2. $\boldsymbol{A}_{gy} = \boldsymbol{K}_{G\times G} \otimes \boldsymbol{I}_{Y\times Y}$, where $\boldsymbol{K}_{G\times G}$ is the relationship matrix among genotypes and $\boldsymbol{I}_{Y\times Y}$ is the identity matrix for years. This model assumes there is a relationship among genotypes (see 2.3) but no relationship among years. This model predicts the performance of genotypes in untested years, that is, years in which it was not tested, based on the performance of their relatives. This model can produce a denser genotype-by-year table than the first model, but it will still have missing predictions when a genotype has no relative tested in a particular year.
3. $\boldsymbol{A}_{gy} = \boldsymbol{K}_{G\times G} \otimes \boldsymbol{W}_{Y\times Y}$, where $\boldsymbol{K}_{G\times G}$ and $\boldsymbol{W}_{Y\times Y}$ are the relationship matrices among genotypes and among years respectively. This is the most complex model, assuming a relationship among genotypes and a relationship among years (see 2.3). This model will produce a complete genotype-by-year table.

2.3. Determining the $\boldsymbol{A}_{gy}$ matrix

The simplest form of the $\boldsymbol{A}_{gy}$ is determined by assuming no relationship among the genotypes and no relationship among the years. This gives an ID model.

The next level of complexity of the matrix is based on a relationship matrix $\boldsymbol{K}_{G\times G}$, among genotypes, and no relationship among the years. The genetic relationship among the genotypes is usually calculated based on either pedigree or marker information.

The coefficient of parentage (COP) matrix for the genotypes was calculated based on the pedigree information and selection history using the BROWSE module in IWIS3 (McLaren 2007). The BROWSE COP algorithm takes account of the inbreeding due to both common ancestors and the degree of selfing in the selection history of each line. This gave a (symmetric) COP-based $\boldsymbol{K}_{G\times G}$, which is used in what is referred to as the COP-based K model.

Marker data of 1,447 DArT markers (Jaccoud *et al.* 2001) were available for 599 of the 686 genotypes in the first 25 ESWYT cycles. Details of the marker data for these genotypes are available from Arief *et al.* (2013). These marker data were used to generate a genomic relationship matrix (GRM) among genotypes based on the simple matching coefficient (SMC, Reif, Melchinger & Frisch 2005). There are many ways to calculate a GRM (e.g. Van Raden 2008; Reif, Melchinger & Frisch 2005), but they produce similar predictions (Tier, Meyer & Ferdosi 2015). Compared to other GRM coefficients (e.g. Van Raden 2008), the SMC enables the inclusion of (almost) non-polymorphic markers in calculating the relationship among genotypes. The SMC is one of the similarity measures for binary data and it does not require standardisation. In a plant breeding population such as the ESWYT, these non-polymorphic markers are essential as most markers can be non-polymorphic due to the intense selection process over a long period of time. These non-polymorphic markers are usually removed when calculating GRM as it creates problems during standardisation. However, in a plant breeding population, their removal could underestimate the level of relatedness among genotypes. The full GRM-based $\boldsymbol{K}_{G\times G}$ matrix used the SMC for pairs from the 599 genotypes with marker data and an SMC set to zero for pairs from the 87 ungenotyped genotypes and for any pair from 1 of the 599 genotypes and 1 of the 87 genotypes. This gave a (symmetric) GRM-based $\boldsymbol{K}_{G\times G}$ which is used in what is referred to as the GRM-based K model.

The third (and highest) level of complexity involves a relationship matrix among the genotypes (just defined) and a relationship matrix among the years, $\boldsymbol{W}_{Y\times Y}$. While there are many forms of the relationship matrix among years, in this paper, we will use a covariance

Table 1. The five mixed models used to predict the performances of the 686 genotypes in each of the first 25 ESWYT cycles/years using a cross tabulation of relationship matrices among genotypes and among years.

		Among years ($\boldsymbol{W}_{Y\times Y}$ matrix)	
		Identity matrix	Covariance matrix†
Among genotypes ($\boldsymbol{K}_{G\times G}$ matrix)	Identity matrix	ID model	
	COP-based‡	COP-based K model	COP-based KFA model
	GRM-based§	GRM-based K model	GRM-based KFA model

†Estimated using the factor analytic (FA) model in ASReml.

‡COP = coefficient of parentage, calculated from pedigree information and selection history.

§GRM = genomic relationship matrix, calculated as a simple matching coefficient (SMC) from marker data.

matrix estimated from the yield data. A relationship among years could also be calculated using other data, such as weather data.

The factor analytic (FA) model has been used previously to model the covariance matrix among environments (e.g. Smith *et al.* 2015; Smith & Cullis 2018; Gogel, Smith & Cullis 2018). Here, the covariance matrix among years, $\boldsymbol{W}_{Y\times Y}$, was estimated by fitting a two-factor FA model, allowing for a two-dimensional visualisation. A higher order FA model could be used, but convergence will depend on the size of the data set and the level of sparseness. When the K model for genotypes is combined with the FA model for environments, we get the KFA model. In our case, we obtain what is referred to as a COP-based KFA model or a GRM-based KFA model.

Each model must converge before further analysis can be conducted. When the default number of iterations in the ASReml software (Gilmour *et al.* 2021) was reached before convergence, the default was increased until convergence occurred.

2.4. Prediction models

In summary, five mixed models were used to predict the performances of the 686 genotypes in each of the first 25 ESWYT cycles/years Table 1. These five models consisted of an ID model, where $\boldsymbol{A}_{gy} = \boldsymbol{I}_{G\times G} \otimes \boldsymbol{I}_{Y\times Y}$, two K models where $\boldsymbol{A}_{gy} = \boldsymbol{K}_{G\times G} \otimes \boldsymbol{I}_{Y\times Y}$ (by using COP-based $\boldsymbol{K}$ matrix and a GRM-based $\boldsymbol{K}$ matrix), and two KFA models where $\boldsymbol{A}_{gy} = \boldsymbol{K}_{G\times G} \otimes \boldsymbol{W}_{Y\times Y}$ (again by using COP-based $\boldsymbol{K}$ matrix and a GRM-based $\boldsymbol{K}$ matrix).

2.5. Visualisation of results

Heatmaps were used to visually display the two-way two-mode table of genotype performance for each ESWYT cycle (i.e. BLUP, Best Linear Unbiased Predictor), the genetic relationship matrix among genotypes (COP-based or GRM-based), and the relationship matrix among years presented as the correlation matrix version of the estimated covariance matrix calculated from the FA model. The genotype-by-year table was ordered to be consistent with the optimised dendrogram order (Gruvaeus & Wainer 1972; Arief *et al.* 2012) obtained from the complementary dissimilarity matrix among genotypes and Ward's clustering strategy (Ward 1963). The complementary dissimilarity matrix among genotypes was obtained using Gower's transformation (Gower 1966) for the COP-based $\boldsymbol{K}_{G\times G}$ and by subtracting the SMC from one for every term in the GRM-based $\boldsymbol{K}_{G\times G}$. For the ID model, the ordering of the

genotypes was based on the optimised dendrogram order of the GRM-based $\boldsymbol{K}_{G\times G}$, while the sequence of the years was always from 1 to 25.

Pattern analysis (Williams 1976; Cooper & DeLacy 1994), a combination of clustering and ordination methods, was used to visualise the pattern of responses of the genotypes across environments and the relationships among genotypes and among environments using the various models. Biplots (Gabriel 1971) with symmetrical scaling were used to display the ordination results. For both the COP-based and GRM-based K models, the ordination was calculated using the singular value decomposition (Eckart & Young 1936) on the column standardised two-way genotype-by-year table (DeLacy *et al.* 1996). For both the COP-based and GRM-based KFA models, the biplots were generated based on the results of the ASReml FA model. The names of the 35 released cultivars (Table 2) were displayed in these biplots. These cultivars were tested in the first 25 ESWYT cycles; most were tested only once, but some were tested in multiple ESWYT cycles.

Optimised dendrograms (Gruvaeus & Wainer 1972; Arief *et al.* 2012) with Ward's clustering strategy (Ward 1963) were used to display the results of the cluster analysis of the years. For COP-based and GRM-based K models, the clustering of the years was conducted using the average squared Euclidean distance calculated from the column standardised genotype-by-year table (Cooper & DeLacy 1994; DeLacy *et al.* 1996). For the COP-based and GRM-based KFA models, the correlation matrix among years was calculated from the covariance matrix obtained from the ASReml FA model. The complementary dissimilarity form of this correlation matrix was calculated via Gower's transformation (Gower 1966).

3. Results

When no relationship matrix was fitted (ID model), the predictions of the performances of the genotypes were only available for the ESWYT cycle (or year) in which they were tested. This resulted in a sparse genotype-by-year table (Figure 1a). When a genetic relationship matrix among genotypes was fitted (K model), the performances of the genotypes were available in most years (Figures 1c and g), provided they had relatives tested in that year (Figure 1b) or had marker data (Figure 1f). If a genotype had no pedigree information or no marker data, its performance was only available in the years in which it was tested (Figures 1c and g, respectively).

When the relationship among years was also fitted (KFA model), the predictions of the performances of the genotypes were available for all years. Hence a full genotype-by-year table became available (Figures 1d and h). The correlation matrix among years (calculated from the estimated covariance matrix from the FA model) showed that all ESWYT pairs, except for those including ESWYT 21, were highly correlated (Figures 1e and i).

The pattern of the genotype performance for the COP-based K model showed that the genotypes from the same family (e.g. Seri M82 (SR82), Glennson T81 (GL81), Genaro T81 (GN81) and Ures T81 (UR81) from the Veery cross, and Bacanora T88 (BC88) and Super Kauz (SK88) from the Kauz cross) tend to have similar responses (Figure 2a). In the COP-based K model, Bacanora T88 (BC88) and Super Kauz (SK88) showed very similar responses (Figure 2a); while in the GRM-based K model, their responses were slightly different (Figure 2b).

The genotypes tested in the ESWYT were known to either carry or not carry the 1B/1R translocation (Zeller 1973). This translocation can segregate within a cross (Warburton,

Table 2. Information about the 35 genotypes of spring bread wheat released as cultivars in Mexico and the Indian subcontinent and tested in the first 25 cycles of Elite Spring Wheat Yield Trials (ESWYT), including the status of their 1B/1R translocation.

Country†	Cultivar name‡	Code§	Cross name¶	ESWYT††	Cycles‡‡	1B/1R§§
Mexico	SieteCerros T66	SC66	II8156	04	1	Yes
Pakistan	Sonalika (66)	SL66	II18427	07	1	No
Mexico	Nacozari F76	NC76	Bluejay	01	3	No
Mexico	Pavon F76	PV76	Pavon	01	4	No
Mexico	Ciano T79	CN79	Buckbuck	03	3	No
Mexico	Tesia F79	TS79	Titmouse	03	1	No
Mexico	Genaro T81¶¶	GN81	Veery	01	5	Yes
Mexico	Glennson M81	GL81	Veery	04	4	Yes
Mexico	Guasave F81	GS81	CM33202	07	1	Yes
Mexico	Huasteco M81	HS81	Hork	07	1	No
Mexico	Sonoita F81	SN81	Mor	02	2	No
Mexico	Tonichi S81	TN81	SWM4610	01	3	No
Mexico	Ures T81	UR81	Veery	04	4	Yes
Mexico	Mexico M82	MX82	II46727	07	1	No
Mexico	Seri M82	SR82	Veery	05	12	Yes
Mexico	Opata M85	OP85	OPATA M-85-X2	04	4	No
Mexico	Cucurpe S86	CC86	Yaco	03	2	No
Mexico	Esmeralda M86¶¶	ES86	CM49641	06	1	No
Mexico	Oasis F 86	OS86	CMH77A.485	14	1	No
Mexico	Papago M86	PP86	CM52359	05	2	No
Mexico	Galvez S87	GL87	Junco	05	3	No
Mexico	Angostura F88 ¶¶	AN88	Bagula	10	2	No
Mexico	Bacanora T88	BC88	Kauz	09	9	Yes
Mexico	Cumpas T88	CM88	CM64624	09	3	Yes
Mexico	Mochis T88	MC88	CM64624	13	1	Yes
Mexico	Super Kauz (88)	SK88	Kauz	9	1	Yes
Mexico	Culiacan T89	CL89	TUI	12	1	Yes
Mexico	Rayon F89	RY89	CM90315	10	5	No
Mexico	Tepoca T89	TP89	Falke	05	5	Yes
Pakistan	Inqalab 91	IQ91	WL711/CROW	18	5	No
Mexico	Baviacora M92	BV92	Babax	14	4	No
Mexico	Borlaug M95	BR95	Weaver	14	4	Yes
India	PBW343 (95)	PB95	Attila	18	4	Yes
Mexico	Inifap M97	IN97	CMBW89Y01231	18	1	Yes
Mexico	Tobarito M97	TB97	CM103379	18	1	Yes

†Country where this cultivar was released.

‡The number in each cultivar name is the year the genotype was released in Mexico or the Indian subcontinent. For a cultivar name that does not contain the year of release, the year of release is given in parentheses.

§Abbreviated names used in this paper.

¶CIMMYT gives names to important crosses which must not be confused with the names of fixed lines derived from these crosses.

††ESWYT cycle in which the cultivar was first tested.

‡‡Total number of ESWYT cycles in which the cultivar was tested, but not necessarily in consecutive ESWYT cycles.

§§1B/1R translocation is a wheat-rye chromosome substitution and translocation. It is one of the major translocations in wheat. The genotypes tested in the ESWYT either carry or not carry this translocation (Zeller 1973).

¶¶Marker data were not available for these cultivars.

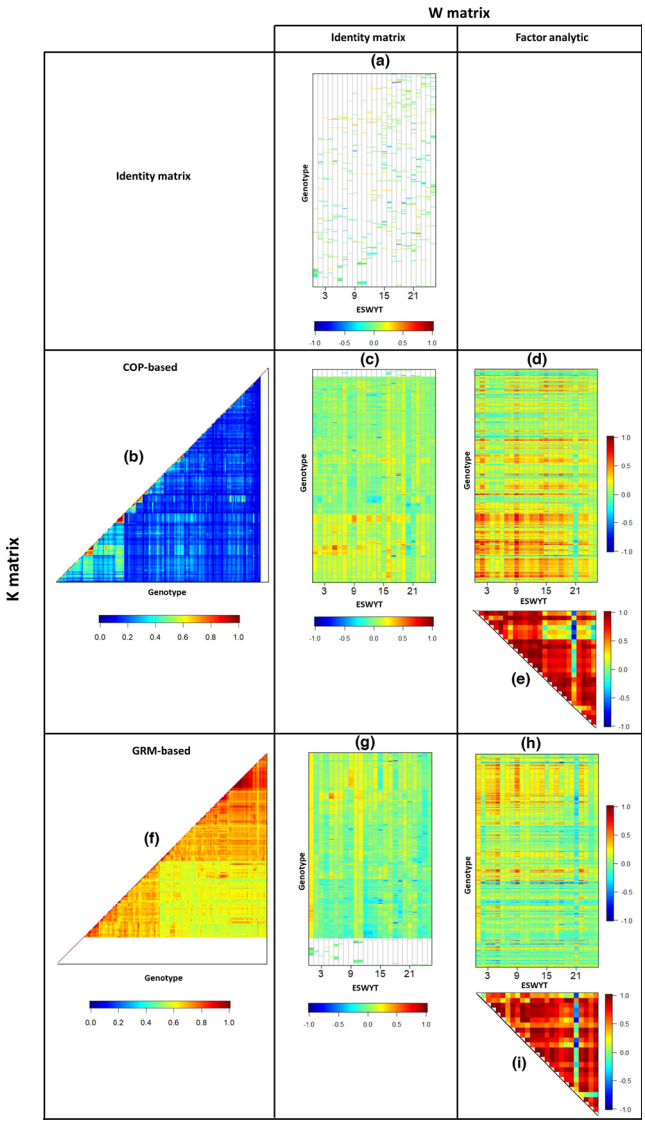

Figure 1. Heatmaps for the performances (Best Linear Unbiased Predictors, BLUPs) of the 686 genotypes in each ESWYT cycle/year from the five prediction models (Table 1) presented by cross-tabulating the genetic relationship matrices among genotypes ($\boldsymbol{K}_{G\times G}$ matrix) with the relationship matrix among years ($\boldsymbol{W}_{Y\times Y}$) presented as the correlation matrix version of the estimated covariance matrix calculated from the factor analytic (FA) model. (a) BLUPs from the ID model; (b) Lower-triangular $\boldsymbol{K}_{G\times G}$ calculated using the pedigree (COP-based $\boldsymbol{K}_{G\times G}$); (c) BLUPs for COP-based K model; (d) BLUPs from COP-based KFA model; (e) Upper-triangular $\boldsymbol{W}_{Y\times Y}$ from the COP-based KFA model; (f) Lower-triangular $\boldsymbol{K}_{G\times G}$ calculated using the marker data (GRM-based $\boldsymbol{K}_{G\times G}$); (g) BLUPs from GRM-based K model; (h) BLUPs from GRM-based KFA model; and (i) Upper-triangular$\boldsymbol{W}_{Y\times Y}$ from the GRM-based KFA model. The white in the heatmaps indicates either a missing BLUP for untested genotypes, zeros in the $\boldsymbol{K}_{G\times G}$ matrices, or diagonal values in the $\boldsymbol{W}_{Y\times Y}$ matrix (equal to 1). The ordering of the genotypes is consistent with the respective optimised dendrograms as explained in the text.

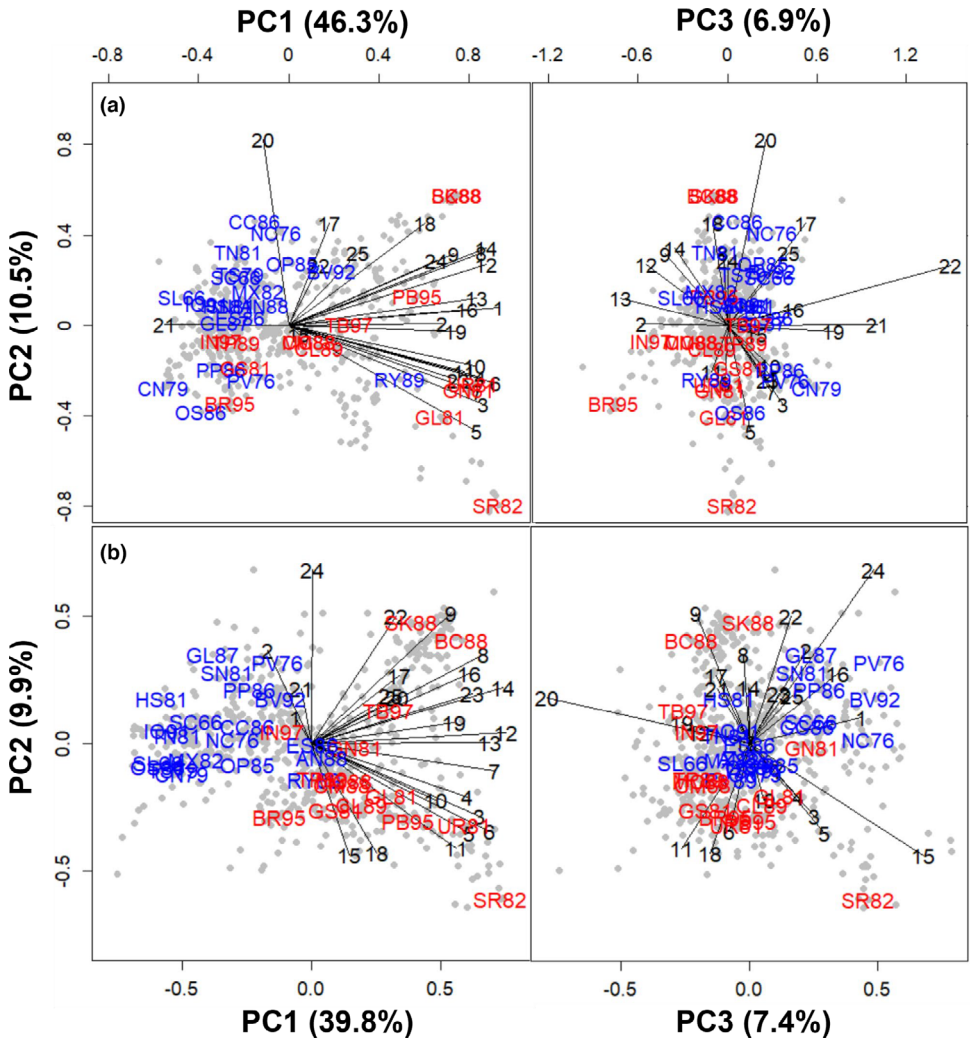

Figure 2. Biplots of the performances of the 686 genotypes in each ESWYT cycle from (a) the COP-based K model and (b) the GRM-based K model. The vectors indicate the ESWYT cycles/years (1 to 25). The names of the 35 released cultivars (Table 2) are included in these biplots. The cultivars in red have the 1B/1R translocation, while the cultivars in blue do not carry the 1B/1R translocation. Refer to Table 1 for a description of these models.

Skovmand & Mujeeb-Kazi 2002). The previous study (Arief *et al.* 2010) showed that the relationship matrix derived from the DArT markers clustered the genotypes in the first 25 ESWYT cycles at the two-group level based on the 1B/1R translocation. The pattern of the performance of genotypes for the GRM-based K model showed these two distinct groups of genotypes. The genotypes that carry 1B/1R translocation (Figure 2b; red) tend to perform more similarly to one another than genotypes without the 1B/1R translocation (Figure 2b; blue).

The grouping of the years reflects how these years discriminate among the genotypes (Cooper & DeLacy 1994). While the groupings of the years based on the genotype-by-year tables from the COP-based and the GRM-based K models were different, some years were

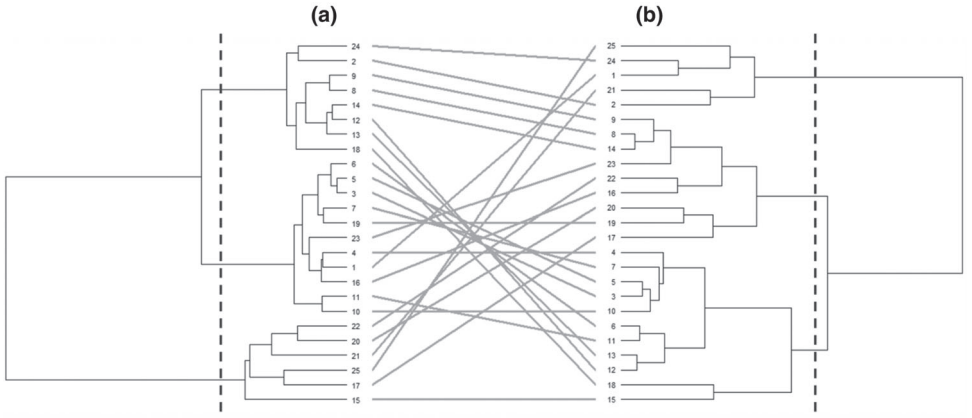

Figure 3. Comparison of the optimised dendrograms of the first 25 ESWYT cycles/years based on the genotype-by-year performance tables from (a) the COP-based K model and (b) the GRM-based K model. The vertical dotted lines indicate the three-group levels. Refer to Table 1 for a description of these models.

grouped similarly (Figure 3). For example, in both cases, ESWYT 9, 8 and 14 were in the same group at the three-group level.

The COP-based KFA model showed that PBW343 (PB95) performed relatively better than Seri M82 (SR82) in ESWYT 4 and 5 (Figure 4a), while the GRM-based KFA model showed that Seri M82 performed slightly better than PBW343 in these two ESWYTs. In the COP-based KFA model, the performance of Seri M82 was more like its sister genotypes (Figure 4a; Glennson T81 (GL81), Genaro T81 (GN81) and Ures T81 (UR81)); while in the GRM-based KFA model, the performance of Seri M82 was closer to the performance of PBW343 (Figure 4b). Note that PBW343 and Seri M82 shared some common parents. Seri M82 was a high-yielding cultivar released in 1982. It was first tested in ESWYT 5 (1983) and tested in 12 of these 25 ESWYT cycles (Table 2). Both the COP-based and GRM-based KFA models showed that Seri M82 performed relatively well in the early ESWYT cycles but poorly in the later ESWYT cycles (Figure 4).

The distinction of the genotypes due to the status of the 1B/1R translocation from the GRM-based K model (Figure 2b) was not observed in the GRM-based KFA model (Figure 4b).

As with the K models, the grouping of the years from the genotype-by-year tables from the COP-based and GRM-based KFA models were different, but some years grouped similarly (Figure 5). In both hierarchies, ESWYT 21 seems to be the most different ESWYT as it does not group with any other ESWYT at the four-group level. In the COP-based KFA model, ESWYT 21 is most similar to ESWYT 22, while in the GRM-based KFA model, ESWYT 21 is most similar to ESWYT 17 (Figure 5).

4. Discussion

Evaluation of long-term genotype performance is often difficult due to the lack of common genotypes across the years. In annual crops, such as wheat, most test genotypes are grown in a single-year trial, and only some elite genotypes are grown in multiple years. A genomic prediction model, which leverages the genetic information (relationship among genotypes) and environmental information (relationship among environments), can be used

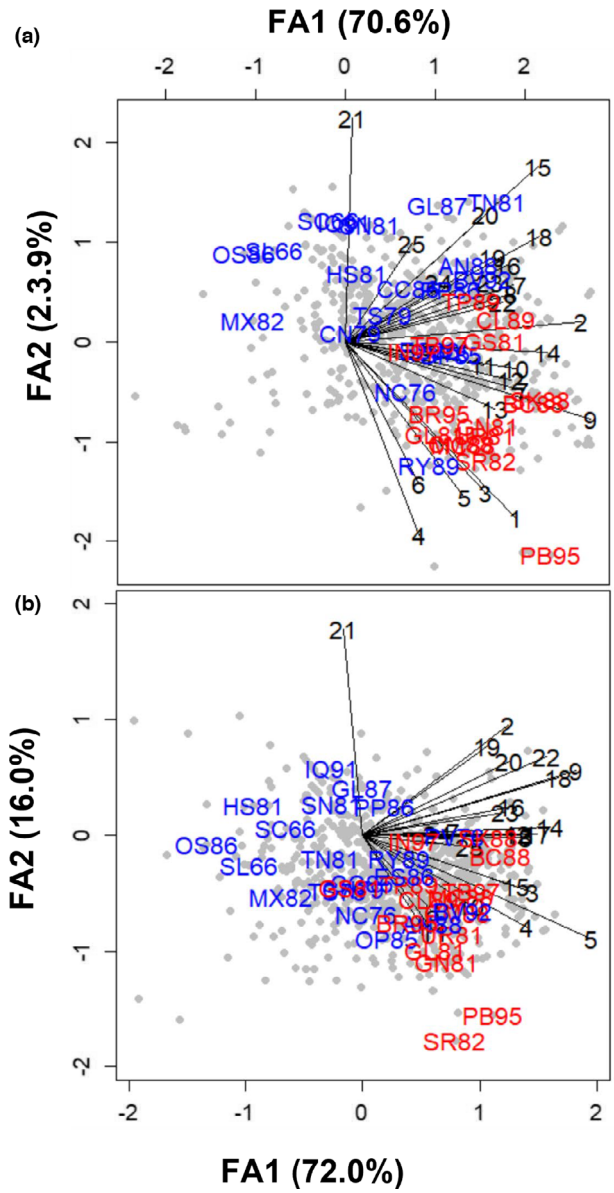

Figure 4. Biplots of the performances of the 686 genotypes in each ESWYT cycle from (a) the COP-based KFA model and (b) the GRM-based KFA model. The vectors indicate the ESWYT cycles/years (1 to 25). The name of the 35 released cultivars (Table 2) are included in these biplots. The cultivars in red have the 1B/1R translocation, while the cultivars in blue do not carry the 1B/1R translocation. Refer to Table 1 for a description of these models.

to predict genotype performance in environments in which they were not grown. Therefore, a complete two-way two-mode table of genotype-by-year estimates (Figure 1) can be created by leveraging the relationship among genotypes and the relationship among years. Such a table can be used to evaluate long-term genotype performance (Figures 2 and 4). It can also

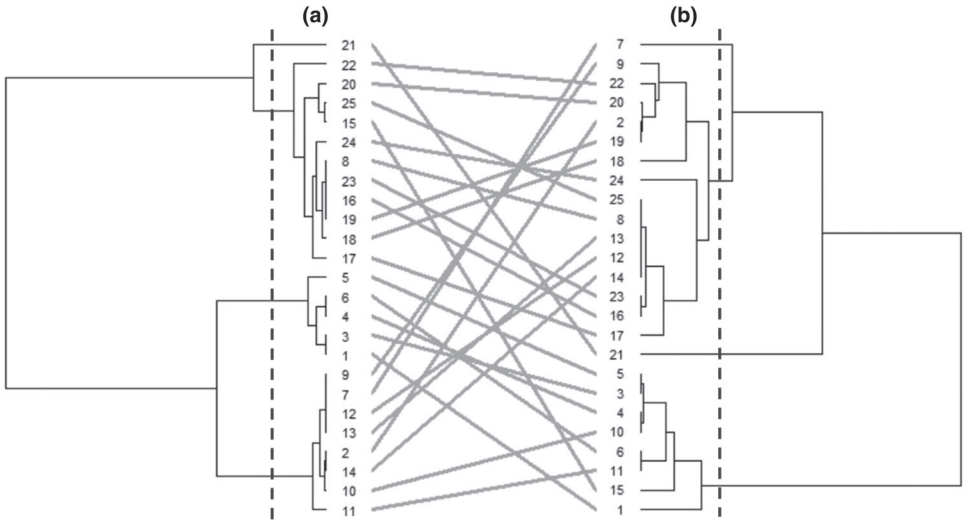

Figure 5. Comparison of the optimised dendrograms of the first 25 ESWYT cycles/years based on the genotype-by-year performance tables from (a) the COP-based KFA model and (b) the GRM-based KFA model. The vertical dotted lines indicate the four-group levels. Refer to Table 1 for a description of these models.

be used to classify the years (Figures 3 and 5). As the characteristics of these past years are usually known, the grouping of years could be useful to forecast the performance of genotypes in years with similar characteristics to those past years. Thus, the application of a genomic prediction model to historical data from plant breeding multi-environment trials can enable the evaluation of the genotypes over a relatively long-term period. Naturally, there is more uncertainty when data are not observed, and the prediction accuracy would depend on the quality of the relationship matrices and the sparseness of the data. However, visualisation of a full genotype-by-year table could provide a general interpretation of the genotype responses across years. The ability to predict genotype performance in past environments can be useful in a cross-validation analysis. With this approach, we can use the past data to predict the future as well as use current data to predict the past. This would increase the number of cross-validation analyses available to calculate prediction accuracy.

The long-term performance in the first 25 ESWYT cycles (1979–2005) shows that plant breeding is important to cope with the changes within and across environments. The performance of genotypes in the environments in which they were selected and released is known as the relative peak performance (De La Vega, DeLacy & Chapman 2007). For example, Seri M82 was one of CIMMYT's highest-yielding genotypes that were released in 1982. The prediction of its performance across the 25 ESWYT cycles showed that Seri M82 performed relatively well in the early ESWYT cycles, especially in the ESWYT in which it was first tested (ESWYT 5 – 1983), but it performed poorly in later ESWYT cycles (ESWYT 20 onwards – Figures 2 and 4). In the later ESWYTs, as the environment changed, more recently released cultivars had a better yield performance than Seri M82.

The response pattern depends on the model used to analyse the data. In this paper, two types of genetic relationships among genotypes ($\boldsymbol{K}_{G\times G}$ matrices) were used: one using a

coefficient of parentage (COP) matrix from the pedigree (COP-based $\boldsymbol{K}_{G\times G}$) and the other using marker information from the genomic relationship matrix (GRM-based $\boldsymbol{K}_{G\times G}$).

In the ESWYT example, the COP-based and GRM-based models produced different patterns, especially for the K model (Figure 2). The performance of genotypes from the COP-based K model showed the effect of families (Figure 2a); while the performance of genotypes from the GRM-based K model showed the effect of the 1B/1R translocation (Figure 2b). These results are expected as the COP-based relationship grouped these genotypes in the first 25 ESWYT cycles based on their family structure; while the GRM-based relationship grouped the genotypes mostly based on the absence or presence of this translocation (Arief *et al.* 2010). The GRM-based model provides a realised relationship matrix that can account for common ancestry not considered in the pedigree and for departures from Mendelian segregation (Burgueño *et al.* 2012). The GRM-based model, derived from the dense molecular markers, has been shown previously to be more predictive than the COP-based model (Crossa *et al.* 2010).

The genotype performance across years can be used to classify the years (Figures 3 and 5). These groupings of years reflect how the years discriminate among the genotypes. As the characteristics of these years are likely to be known, they can be used to calculate the probability of occurrences of these groups of years. Then the average genotype performance can be calculated for each of these year groups (Chapman *et al.* 2000) to provide a forecast of the performance of genotypes in future years. The long-term performance of genotypes can be predicted by calculating the weighted mean using the average frequency occurrences across the groups (Chapman *et al.* 2000).

When the genetic relationship matrix among genotypes is incomplete due to missing pedigree or marker information, the addition of the relationship matrix among environments in the KFA model enables the prediction of all genotypes (Figure 1). In this example, the relationship matrix among environments (i.e. years/cycles) was estimated using an FA model with two factors. An FA model provides an intrinsic relationship matrix (i.e. relationship matrix calculated from the data). A relationship among environments can also be calculated from environmental covariates, such as soil types and rainfall (Jarquín *et al.* 2014; Saint Pierre *et al.* 2016), in which case it would be an extrinsic relationship matrix.

A complete genetic relationship matrix among genotypes can also be obtained by combining the pedigree- and marker-based relationships, known as the H matrix (Legarra, Aguilar & Misztal 2009; Misztal, Legarra & Aguilar 2009; Christensen & Lund 2010). The inclusion of a complete genetic relationship matrix among genotypes and an appropriate relationship matrix among environments would be expected to provide a better prediction of the performances of genotypes in the past (and present) environments (and inferred for similar future environments). A dense genotype-by-environment table (especially when the environments include years) would be invaluable as a guide for the long-term evaluation of a plant breeding program. Moreover, this approach incurs no extra cost to the plant breeding program as the data required (i.e. phenotypic, pedigree, marker and environment characteristics) are likely to be available from a routine breeding program.

5. Conclusion

The evaluation of the long-term performance of genotypes is often tricky due to a limited number of test years and a limited number of common genotypes across years. A genomic

prediction model that uses appropriate relationship matrices among genotypes and among environments (which include years) can be used to generate a complete genotype-by-year array when only a sparse one was measured. While the accuracy of the predicted genotype performance depends on the sparseness of the data and the quality of the relationship matrices, the visualisation of this complete genotype-by-year array can be useful to provide a guide on the long-term performance of genotypes in a plant breeding program.

Appendix

An ASReml parameter file, R script for visualisation of results with some examples and electronic version of all figures are available from https://github.com/varief/VizGenByYear.

References

ARIEF, V.N. (2010). Methodology development to integrate phenotypic and genotypic data from a plant breeding program through association analysis with a case study using data from the CIMMYT International Spring Wheat Field Trials. Ph.D. thesis, School of Agriculture and Food Sciences.

ARIEF, V.N., DELACY, I.H., BASFORD, K. & DIETERS, M. (2012). Optimized dendrogram: extracting population information using seriation to enhance hierarchical clustering. In *4th International Conference in Quantitative Genetics - Programme & book of abstract.* p. 72.

ARIEF, V.N., DELACY, I.H., CROSSA, J., DIETERS, M.J. & BASFORD, K.E. (2010). The use of pedigree, molecular marker and phenotypic data to investigate population structures in 25 years of the CIMMYT global wheat breeding program. In *8th International Wheat Conference - Abstracts of oral and poster presentations. p. 425.*

ARIEF, V.N., DELACY, I.H., CROSSA, J., *et al.* (2015). Evaluating testing strategies for plant breeding field trials: redesigning a CIMMYT international wheat nursery to provide extra genotype connection accross cycles. *Crop Science* **55**, 164–177.

ARIEF, V.N., DELACY, I.H., WENZL, P., *et al.* (2013). Using molecular marker order to compare genetic structure in plant populations undergoing selection. *Journal of Environmental Statistics* **4**, 1.

BURGUEÑO, J., CAMPOS, G.D.L., WEIGE, K. & CROSSA, J. (2012). Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. *Crop Science* **52**, 707–719.

BURGUEÑO, J., CROSSA, J., CORNELIUS, P.L., TRETHOWAN, R., MCLAREN, G. & KRISHNAMACHARI, A. (2007). Modeling additive × environment and additive × additive × environment using genetic covariances of relatives of wheat genotypes. *Crop Science* **47**, 311–320.

CHAPMAN, S.C., HAMMER, G.L., BUTLER, D.G. & COOPER, M. (2000). Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population environments. *Australian Journal of Agricultural Research* **51**, 223–233.

CHRISTENSEN, O.F. & LUND, M.S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* **42**, 2.

COOPER, M. & DELACY, I.H. (1994). Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theoretical and Applied Genetics* **88**, 561–572.

CROSSA, J., BURGUEÑO, J., CORNELIUS, P.L., MCLAREN, G., TRETHOWAN, R. & KRISHNAMACHARI, A. (2006). Modeling genotype × environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. *Crop Science* **46**, 1722–1733.

CROSSA, J., DE LOS CAMPOS, G., GIANOLA, D., *et al.* (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* **186**, 713–724.

CULLIS, B.R., THOMSON, F.M., FISHER, J.A., GILMOUR, A.R. & THOMPSON, R. (1996). The analysis of the NSWwheat variety database. II. Variance component estimation. *Theoretical and Applied Genetics* **92**, 28–39.

DE LA VEGA, A.J., DELACY, I.H. & CHAPMAN, S.C. (2007). Progress over 20 years of sunflower breeding in central Argentina. *Field Crops Research* **100**, 61–72.

DeLacy, I.H., Basford, K.E., Cooper, M. & Fox, P.N. (1996). Restropective analysis of historical data sets from multi-environmental trials – theoretical development. In *Plant Adaptation and Crop Improvement*, eds. M. Cooper and G.L. Hammer, pp. 243–267. Wallingford, UK: CAB International.

DeLacy, I.H., Ratnasiri, W.G.A. & Mirzawan, P.D.N. (1996). Retrospective analysis of historical data sets from multi-environmental trials – case studies. In *Plant Adaptation and Crop Improvement*, eds. M. Cooper and G.L. Hammer, pp. 269–290. Wallingford, UK: CAB International.

DeLacy, I.H., Redden, R.J., Butler, D.G. & Usher, T. (2000). Analysis of line × environment interactions for yield in navy beans. 3. Pattern analysis of environments over years. *Australian Journal of Agricultural Research* **51**, 619–628.

Eckart, C. & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika* **1**, 211–218.

Frensham, A., Cullis, B. & Verbyla, A. (1997). Genotype by environment variance heterogeneity in a two-stage analysis. *Biometrics* **53**, 1373–1383.

Gabriel, K.R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika* **58**, 453–467.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. & Thompson, R. (2021). *ASReml User Guide Release 4.2 Functional Specification*. VSN International Ltd., Hemel Hempstead, HP2 4TP, UK. Available from URL: hrrp://www.vsni.co.uk

Gogel, B., Smith, A. & Cullis, B. (2018). Comparison of a one- and two-stage mixed model analysis of Australia's National Variety Trial Southern Region wheat data. *Euphytica* **214**, 44–64.

Gower, J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325–338.

Gruvaeus, G. & Wainer, H. (1972). Two additions to hierarchical cluster analysis. *British Journal of Mathematical and Statistical Psychology* **25**, 200–206.

Jaccoud, D., Peng, K., Feinstein, D. & Kilian, A. (2001). Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* **29**, e25.

Jarquín, D., Crossa, J., Lacaze, X., *et al.* (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics* **127**, 595–607.

Legarra, A., Aguilar, I. & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science* **92**, 4656–4663.

Manly, B.F.J. (1994). *Multivariate Statistical Methods: A primer*. London, UK: Chapman & Hall.

McLachlan, G.J. & Basford, K.E. (1988). *Mixture Models: Inference and Applications to Clustering*. New York: Marcel Dekker.

McLaren, C.G. (2007). TDM GMS Browse. Available from URL: https://cropforge.github.io/iciswiki/articles/t/d/m/TDM_GMS_Browse_53fa.html [Last accessed 23 Feb 2018].

Misztal, I., Legarra, A. & Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science* **92**, 4648–4655.

Mohring, J. & Piepho, H.P. (2009). Comparison of weighting in two-stage analysis in plant breeding trials. *Crop Science* **49**, 1977–1988.

Oakey, H., Verbyla, A., Pitchford, W., Cullis, B. & Kuchel, H. (2006). Joint modeling of additive and non-additive genetic line effects in single field trials. *Theoretical and Applied Genetics* **113**, 809–819.

Oakey, H., Verbyla, A.P., Cullis, B.R., Wei, X. & Pitchford, W.S. (2007). Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. *Theoretical and Applied Genetics* **114**, 1319–1332.

Patterson, H.D. & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554.

Piepho, H.P., Möhring, J., Schulz-Streeck, T. & Ogutu, J.O. (2012). A stage-wise approach for the analysis of multi-environment trials. *Biometrical Journal* **54**, 844–860.

Reif, J.C., Melchinger, A.E. & Frisch, M. (2005). Genetical and mathematical properties of similarity and dissimilarity of coefficients applied in plant breeding and seed bank management. *Crop Science* **45**, 1–7.

Saint Pierre, C., Burgueño, J., Crossa, J., *et al.* (2016). Genomic prediction models for grain yield of spring bread wheat in diverse agro-ecological zones. *Scientific Reports* **6**, 27312.

Smith, A., Cullis, B. & Gilmour, A. (2001). The analysis of crop variety evaluation data in Australia. *Australian & New Zealand Journal of Statistics* **43**, 129–145.

Smith, A.B. & Cullis, B.R. (2018). Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica* **214**, 143–161.

SMITH, A.B., GANESALINGAM, A., KUCHEL, H. & CULLIS, B.R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics* **128**, 55–72.

TIER, B., MEYER, K. & FERDOSI, M.H. (2015). Which genomic relationship matrix? *Proceedings of Association for the Advancement of Animal Breeding and Genetics* **21**, 461–464.

VAN RADEN, P.M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–4423.

WARBURTON, M.L., SKOVMAND, B. & MUJEEB-KAZI, A. (2002). Themolecular genetics characterization of the 'Bobwhite' bread wheat family using AFLPs and the effect of the T1BL.1RS translocation. *Theoretical and Applied Genetics* **104**, 868–873.

WARD, J.H. (1963). Hierarchical grouping to optimise an objective function. *Journal of the American Statistical Association* **58**, 236–244.

WELHAM, S.J., GOGEL, B.J., SMITH, A.B., THOMPSON, R. & CULLIS, B.R. (2010). A comparison of analysis methods for late-stage variety evaluation trials. *Australian & New Zealand Journal of Statistics* **52**, 125–149.

WILLIAMS, W.T. (1976). *Pattern Analysis in Agricultural Science*. Amsterdam: Elsevier Scientific Publishing Company.

ZELLER, F.J. (1973). 1B/1R wheat-rye chromosome substitutions and translocation. In *4th International Wheat Genetics Symposium*, eds. E.R. Sears and L.M.S. Sears, pp. 209–221. Columbia, MO: University of Misourri.