




# A singular value decomposition Bayesian multiple-trait and multiple-environment genomic model

Osva A. Montesinos-López<sup>1</sup> · Abelardo Montesinos-López<sup>2</sup> · José Crossa<sup>3</sup>  · Kismiantini<sup>4</sup> · Juan Manuel Ramírez-Alcaraz<sup>1</sup> · Ravi Singh<sup>3</sup> · S. Mondal<sup>3</sup> · P. Juliana<sup>3</sup>

Received: 3 February 2018 / Revised: 22 May 2018 / Accepted: 12 June 2018 / Published online: 17 August 2018  
© The Author(s) 2018. This article is published with open access

## Abstract

Today, breeders perform genomic-assisted breeding to improve more than one trait. However, frequently there are several traits under study at one time, and the implementation of current genomic multiple-trait and multiple-environment models is challenging. Consequently, we propose a four-stage analysis for multiple-trait data in this paper. In the first stage, we perform singular value decomposition (SVD) on the resulting matrix of trait responses; in the second stage, we perform multiple trait analysis on transformed responses. In stages three and four, we collect and transform the traits back to their original state and obtain the parameter estimates and the predictions on these scale variables prior to transformation. The results of the proposed method are compared, in terms of parameter estimation and prediction accuracy, with the results of the Bayesian multiple-trait and multiple-environment model (BMTME) previously described in the literature. We found that the proposed method based on SVD produced similar results, in terms of parameter estimation and prediction accuracy, to those obtained with the BMTME model. Moreover, the proposed multiple-trait method is attractive because it can be implemented using current single-trait genomic prediction software, which yields a more efficient algorithm in terms of computation.

## Introduction

Breeders often want to improve more than one trait simultaneously in their breeding programs and thus conduct

various experiments. For example, Teixeira et al. (2016) reported that a breeding program in Brazil measured 41 pig traits obtained by crossing 345 F2 pig populations of Brazilian Piau × commercial pigs. To analyze this type of experiment, breeders implement one of the following two approaches: (i) they perform a univariate analysis (one trait at a time), therefore ignoring the correlation between traits that does not allow to improve either parameter estimates or prediction accuracy, or (ii) they perform a multiple-trait analysis, which may not only take into account the correlation between traits but may also significantly increase the computing intensity.

Implementing first approach is valid when the correlation between traits is low or close to zero, but it is less desirable when the correlation between traits is moderate to strong. However, implementing the second approach is sometimes challenging—for example, when there are a large number of traits, and under these circumstances, breeders opt for the first approach. Furthermore, implementing the second approach is also challenging because early breeding programs start with at least 1000 lines that are evaluated in multiple environments, which complicates the analysis, as including genotype × environment ( $G \times E$ ) increases the

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1038/s41437-018-0109-7>) contains supplementary material, which is available to authorized users.

---

✉ Osva A. Montesinos-López  
oamontes2@hotmail.com

✉ José Crossa  
j.crossa@cgiar.org

<sup>1</sup> Facultad de Telemática, Universidad de Colima, 28040 Colima, Mexico

<sup>2</sup> Departamento de Matemáticas, Centro Universitario de Ciencias Exactas e Ingenierías (CUCEI), Universidad de Guadalajara, 44430 Guadalajara, Jalisco, Mexico

<sup>3</sup> International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Ciudad de México, Mexico

<sup>4</sup> Department of Mathematics, Yogyakarta State University, Yogyakarta, Indonesia

dimensionality of the data considerably (Chiquet et al. 2013).

As mentioned above, multiple-trait analysis improves parameter estimates and prediction accuracy. With regard to parameter estimates, Schulthess et al. (2017) found that multiple-trait analysis improves parameter estimates, while Calus and Veerkamp (2011) found modest improvement using multiple-trait analysis compared to separate-trait analysis for prediction accuracy; these authors also showed that the performance of multiple-trait analysis depends considerably on whether only some traits are missing in only some individuals or in all individuals. Jia and Jannink (2012) also showed evidence favoring multiple-trait analysis compared to single-trait analysis, finding that the genetic correlation between traits is the basis for the benefit of multiple-trait analysis. Jiang et al. (2015) later arrived at the same conclusion in favor of multiple-trait analysis. Montesinos-López et al. (2016) also found modest improvement in the prediction accuracy of multiple-trait analysis when comparing correlated traits to the analysis that assumes null correlation between traits. Along these lines, He et al. (2016) concluded that modeling multiple traits could improve the prediction accuracy for correlated traits in comparison to univariate-trait analysis. Schulthess et al. (2017) also found that multiple-trait analysis is better in terms of prediction accuracy than separate-trait analysis, pointing out that the multiple-trait model is better when the degree of relatedness between genotypes is weaker.

There is evidence suggesting that even when traits are correlated, genomic-enabled prediction accuracy is not improved, as stated by Montesinos-López et al. (2017a) and as shown by Märtens et al. (2016), who compared multiple-trait analysis to single-trait analysis and found no difference in terms of prediction accuracy. This issue was also documented by Oliveira and Teixeira-Pinto (2015) who proved this result by stating that, in the multivariate linear regression case (when the covariates in each equation are the same), even if the errors are strongly correlated, the multivariate model gives the same result (both point estimates and standard errors) as fitting individual regressions with ordinal least squares for each outcome, despite the level of correlation between the errors.

As a breeding tool, genomic selection (GS) uses all available molecular markers (Meuwissen et al. 2001) to design genomic-assisted breeding programs and develop new marker-based models for genetic evaluation. GS provides opportunities to obtain higher rates of genetic gain than traditional phenotypic selection in less time and at a reasonable cost. For example, in animal breeding, GS allows animal scientists to select young animals early in life that do not have records, greatly reducing evaluation costs and generation intervals when compared to the traditional

progeny test schemes (Schaeffer 2006; Boichard et al. 2016). In general, for traits that have a long generation time or are difficult to evaluate (i.e., insect resistance, bread-making quality, and others), GS is cheaper and/or easier than traditional phenotypic selection because more candidates can be characterized for a given cost, thus enabling increased selection intensity. Hence, GS has a number of merits over traditional selection because it reduces selection duration and increases selection accuracy, intensity, efficiency, and gains per unit of time. In addition, it saves time and financial investment, along with producing reliable results (Rutkoski et al. 2011; Desta and Ortiz 2014). This enables faster development of improved crop varieties to cope with the challenges of climate change and the decrease in arable land (Bhat et al. 2016).

Therefore, we propose an alternative method for analyzing multiple-trait and multi-environment data—one that takes into account the correlation between traits. This model will be useful for analyzing multiple-trait and multi-environment data because the linear predictor may include the following interaction terms: environment  $\times$  trait, genotype  $\times$  trait, and three-way interaction (environment  $\times$  genotype  $\times$  trait), assuming an unstructured variance–covariance matrix in the genetic and residual covariance matrices of traits and an identity matrix for the correlation matrix between environments.

The proposed method consists of four steps. In the first step, we transform the original matrix of response variables into a matrix of response variables of the same dimension but between uncorrelated transformed traits using singular value decomposition (SVD), which is equivalent to using principal component analyses (PCA). This first step is performed ignoring all the information of the design effect and other covariates. In the second step, given that the traits are not correlated, we apply a single-trait analysis where we can take into account the design effect, along with the effects of genotypes, environments, genotype  $\times$  environment interaction, and other covariates, if they are available. In the third step, we collect and put together all the parameter estimates of the single analysis, which are transformed in the fourth step to obtain the parameter estimates and/or predictions for the traits in the original scale of the multiple-trait and multiple-environment data. Our approach has the same goal as the canonical transformation method proposed by Thompson (1977), which involves using special matrices to transform the observations on several correlated traits into new variables that are uncorrelated to each other, which means that these new variables can be analyzed as single-trait analysis, but the results (predictions) are transformed back to the original scale of the observations (Mrode 2014). However, our method is different to the Thompson (1977) method since our approach directly

decorrelates the matrix of response variables with the SVD, while the Thompson (1977) method transforms the variance of the response  $\text{var}(\mathbf{Y}) = \mathbf{\Sigma}_t + \mathbf{R}$ , such that  $\mathbf{QRQ}^T = \mathbf{I}$  and  $\mathbf{Q}\mathbf{\Sigma}_t\mathbf{Q}^T = \mathbf{W}$ , where  $\mathbf{I}$  is an identity matrix and  $\mathbf{W}$  is a diagonal matrix, of course assuming that  $\mathbf{\Sigma}_t$  and  $\mathbf{R}$  are positive definite matrices and that there is a matrix  $\mathbf{Q}$ . More details of this method can be found in chapter 6 of the book by Mrode (2014). A detailed example of the Thompson (1977) method is available in Appendix E, section E.1, in the book by Mrode (2014). The Thompson (1977) method has the inconvenience that if the (co)variances  $\mathbf{R}$  and  $\mathbf{\Sigma}_t$  are unknown, the transformations mentioned above need to be applied at each iteration of the Henderson mixed model equations. For this reason, its implementation is not straightforward using current univariate software.

The advantage of the proposed alternative method is that the analysis can be performed directly using the current software for univariate genomic selection and prediction. However, it is important to point out that when the distribution of the traits has considerably departed from normality, the proposed method does not guarantee independence between the transformed traits, a key assumption for the successful implementation of the proposed model, since for non-normal traits lack of correlation does not imply independence. Also, the proposed method can be implemented to perform the analysis even when there are some missing traits per individual and some individuals are missing in some environments, as long as the level of unbalance is not strong.

## Materials and methods

### Statistical models

#### Multiple-trait multiple-environment model

Since genotype  $\times$  environment interaction is of paramount importance in plant breeding, the following univariate linear mixed model is usually used for each trait:

$$y_{ij} = E_i + g_j + gE_{ij} + e_{ij} \quad (1)$$

where  $y_{ij}$  represents the normal response from the  $j$ th line in the  $i$ th environment ( $i = 1, 2, \dots, I, j = 1, 2, \dots, J$ ).  $E_i$  represents the effect of the  $i$ th environment and is assumed as a fixed effect,  $g_j$  represents the random effect of the genomic effect of the  $j$ th line, with  $\mathbf{g} = (g_1, \dots, g_J)^T \sim N(\mathbf{0}, \sigma_1^2 \mathbf{G}_g)$ ,  $\sigma_1^2$  denotes the genomic variance and  $\mathbf{G}_g$  is of order  $J \times J$  and represents the genomic relationship matrix (GRM) and is calculated using the Van Raden (2008) method as  $\mathbf{G}_g = \frac{\mathbf{Z}\mathbf{Z}^T}{p}$ , where  $p$  denotes the number of markers and  $\mathbf{Z}$  the matrix of markers of order  $J \times p$ . The  $\mathbf{G}_g$  covariance matrix is constructed using the observed similarity at the genomic

level between lines, rather than the expected similarity based on pedigree.  $gE_{ij}$  is the random interaction term between the genomic effect of the  $j$ th line and the  $i$ th environment where  $\mathbf{gE} = (gE_{11}, \dots, gE_{IJ})^T \sim N(\mathbf{0}, \sigma_2^2 \mathbf{I}_I \otimes \mathbf{G})$ ,  $\sigma_2^2$  denotes the variance of the interaction term of genotype by environment, and  $e_{ij}$  is a random error term associated with the  $j$ th line in the  $i$ th environment distributed as  $N(0, \sigma^2)$ , with  $\sigma^2$  denoting the residual variance. This model is usually used for each of the  $l = 1, \dots, L$  traits, where  $L$  denotes the number of traits under study. Next we will present the multivariate version of model (1); for this reason, first we provide the notation for the matrix variate normal distribution, which is a generalization of the multivariate normal distribution. In particular, let the  $(n \times p)$  random matrix,  $\mathbf{M}$ , be distributed as matrix variate normal distribution denoted as  $\mathbf{M} \sim \text{NM}_{n \times p}(\mathbf{H}, \mathbf{\Omega}, \mathbf{\Sigma})$ , if and only if, the  $(np \times 1)$  random vector  $\text{vec}(\mathbf{M})$  is distributed as multivariate normal denoted as  $N_{np}(\text{vec}(\mathbf{H}), \mathbf{\Sigma} \otimes \mathbf{\Omega})$ ; therefore,  $\text{NM}_{n \times p}$  denotes the  $(n \times p)$  dimensional matrix variate normal distribution,  $\mathbf{H}$  is a  $(n \times p)$  location matrix,  $\mathbf{\Sigma}$  is a  $(p \times p)$  first covariance matrix, and  $\mathbf{\Omega}$  is a  $(n \times n)$  second covariance matrix (Srivastava and Khatri 1979).  $\text{vec}(\cdot)$  and  $\otimes$  are the standard vector operator and Kronecker product, respectively.

To account for the correlation between traits, all of the  $L$  traits given in Eq. (1) are jointly modeled in a whole multiple-trait, multiple-environment mixed model as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{b}_1 + \mathbf{Z}_2\mathbf{b}_2 + \mathbf{e} \quad (2)$$

where  $\mathbf{Y}$  is of order  $n \times L$ ,  $\mathbf{X}$  is of order  $n \times I$ ,  $\boldsymbol{\beta}$  is of order  $I \times L$ ,  $\mathbf{Z}_1$  is of order  $n \times J$ ,  $\mathbf{b}_1$  is of order  $J \times L$  and contains the first interaction term genotype  $\times$  trait,  $\mathbf{Z}_2$  is of order  $n \times IJ$ ,  $\mathbf{b}_2$  is of order  $IJ \times L$  and contains the second interaction term genotype  $\times$  environment  $\times$  trait, and  $\mathbf{e}$  is of order  $n \times L$ , with  $\mathbf{b}_1$  distributed under matrix variate normal distribution as  $\text{NM}_{J \times L}(\mathbf{0}, \mathbf{G}_g, \mathbf{\Sigma}_t)$ , where  $\mathbf{\Sigma}_t$  is the unstructured genetic (co)variance matrix of traits of order  $L \times L$ ,  $\mathbf{b}_2 \sim \text{NM}_{IJ \times L}(\mathbf{0}, \mathbf{\Sigma}_E \otimes \mathbf{G}_g, \mathbf{\Sigma}_t)$ , where  $\mathbf{\Sigma}_E$  is an unstructured (co)variance matrix of order  $I \times I$ , and  $\mathbf{e} \sim \text{NM}_{n \times L}(\mathbf{0}, \mathbf{I}_n, \mathbf{R}_e)$ , where  $\mathbf{R}_e$  is the unstructured residual (co)variance matrix of traits of order  $L \times L$ , and  $\mathbf{G}_g$  is the GRM described above. The Bayesian multiple-trait and multiple-environment (BMTME) model resulting from Eq. (2) was implemented by Montesinos-López et al. (2016).

First, we provided a modified version of the original BMTME model proposed by Montesinos-López et al. (2016), and in the next section, we will provide the modified Gibbs sampler for this modified BMTME model.

#### Gibbs sampler for the BMTME model

Outlined below is the Gibbs sampler for estimating the parameter of interest in the BMTME model. While the order is somewhat arbitrary, we suggest the following:

Step 1. Simulate  $\beta$  according to the normal distribution given in Supplementary material (A.1).

Step 2. Simulate  $b_1$  according to the normal distribution given in Supplementary material (A.2).

Step 3. Simulate  $b_2$  according to the normal distribution given in Supplementary material (A.3).

Step 4. Simulate  $\Sigma_t$  according to the inverse Wishart (IW) distribution given in Supplementary material (A.4).

Step 5. Simulate  $\Sigma_E$  according to the IW distribution given in Supplementary material A (A.5).

Step 6. Simulate  $R_e$  according to the IW distribution given in Supplementary material A (A.6).

Step 7. Return to step 1 or terminate when chain length is adequate to meet convergence diagnostics.

The main differences between this Gibbs sampler and that given by Montesinos-López et al. (2016) are: (i) that this modified Gibbs sampler assumes an unstructured variance-covariance matrix for environments, while the original BMTME model assumes a diagonal variance-covariance matrix for environments, and (ii) that the original BMTME model used non-informative priors based on the Half- $t$  distribution of each standard deviation term and uniform priors on each correlation of the covariance matrices of traits (genetic and residual). The modified BMTME model presented here assumes weak informative priors not based on the Half- $t$  distribution of each standard deviation (details of the hyperparameters of the BMTME model are given in Supplementary material). The hyperparameters for the BMTME model were set similar to those used in the BGLR software (Pérez-Rodríguez and de los Campos 2014). The modified full conditional of this modified BMTME model, which supports the Gibbs sampler given above, is provided in Supplementary material. Next, we will describe the parameterization of the model given in Eq. (2) to develop the proposed alternative method that is based on SVD.

**BMTME\_Thompson version**

Following Thompson (1977) and Ducrocq and Chapuis (1993), we define  $Q$  as a matrix of order  $L \times L$  such that  $Q\Sigma_t Q^T = D_t$ , where  $D_t$  is a diagonal matrix also of order  $L \times L$  and  $QR_e Q^T = I_t$ . The  $Q$  matrix always exists and can be calculated as  $Q = L^T P$ , where  $P = U_e B_e^{-0.5} U_e^T$ , where  $U_e$  and  $B_e$  are obtained by applying the SVD to  $R_e = U_e B_e U_e^T$ , while  $L^T$  is obtained also by applying the SVD to  $P\Sigma_t P^T = LD_t L^T$ . Then, by applying a linear transformation to Eq. (2), we obtain:

$$\begin{aligned} YQ^T &= X\beta Q^T + Z_1 b_1 Q^T + Z_2 b_2 Q^T + eQ^T \\ Y^{\&} &= X\beta^{\&} + Z_1 b_1^{\&} + Z_2 b_2^{\&} + Qe^{\&} \end{aligned} \tag{3}$$

Then note that:

$$\begin{aligned} \text{Var}(\text{vec}(b_1^{\&})) &= (Q \otimes I_J) \text{Var}(\text{vec}(b_1)) (Q^T \otimes I_J) \\ &= Q\Sigma_t Q^T \otimes G_g = D_t \otimes G_g \\ \text{Var}(\text{vec}(b_2^{\&})) &= (Q \otimes I_{Jl}) \text{Var}(\text{vec}(b_2)) (Q^T \otimes I_{Jl}) \\ &= Q\Sigma_t Q^T \otimes \Sigma_E \otimes G_g = \\ &= D_t \otimes \Sigma_E \otimes G_g \\ \text{Var}(\text{vec}(e^{\&})) &= (Q \otimes I_n) \text{Var}(\text{vec}(e)) (Q^T \otimes I_n) \\ &= QR_e Q^T \otimes I_n = I_t \otimes I_n \end{aligned}$$

Since  $D_t$  and  $I_t$  are diagonal matrices of order  $t \times t$ , the full conditional distributions for the transformed random effects  $b_1^{\&}$  and  $b_2^{\&}$  of the BMTME model are:

Full conditional for  $\text{vec}(b_1^{\&})$

$$P(\text{vec}(b_1^{\&} | ELSE) \propto N(\text{vec}(b_1^{\&}), \tilde{\Sigma}_{b_1^{\&}}) \tag{4}$$

where  $\tilde{\Sigma}_{b_1^{\&}} = (D_t^{-1} \otimes G_g^{-1} + I_L \otimes Z_1^T Z_1)^{-1}$  and  $\text{vec}(\tilde{b}_1^{\&}) = \tilde{\Sigma}_{b_1^{\&}} (I_L \otimes Z_1^T) [\text{vec}(Y^{\&}) - \text{vec}(X\beta^{\&}) - \text{vec}(Z_2 b_2^{\&})]$ .

Full conditional for  $\text{vec}(b_2^{\&})$

$$P(\text{vec}(b_2^{\&} | ELSE) \propto N(\text{vec}(\tilde{b}_2^{\&}), \tilde{\Sigma}_{b_2^{\&}}) \tag{5}$$

where  $\tilde{\Sigma}_{b_2^{\&}} = (D_t^{-1} \otimes \Sigma_E^{-1} \otimes G_g^{-1} + I_L \otimes Z_2^T Z_2)^{-1}$  and  $\text{vec}(\tilde{b}_2^{\&}) = \tilde{\Sigma}_{b_2^{\&}} (I_L \otimes Z_2^T) \{ \text{vec}(Y^{\&}) - \text{vec}(X\beta^{\&}) - \text{vec}(Z_1 b_1^{\&}) \}$ .

It is important to point out that the full conditionals of  $b_1^{\&}$  and  $b_2^{\&}$  are diagonals for traits; for this reason, these full conditionals can be sampled independently for each trait, as:

Full conditional for  $b_1^{\&(l)}$  for  $l = 1, 2, \dots, L$

$$P(b_1^{\&(l)} | ELSE) \propto N(\tilde{b}_1^{\&(l)}, \tilde{\Sigma}_{b_1^{\&(l)}}) \tag{6}$$

where  $\tilde{\Sigma}_{b_1^{\&(l)}} = (d_l^{-1} \otimes G_g^{-1} + Z_1^T Z_1)^{-1}$  and  $\tilde{b}_1^{\&(l)} = \tilde{\Sigma}_{b_1^{\&(l)}} (Z_1^T) [Y^{\&(l)} - X\beta^{\&(l)} - Z_2 b_2^{\&(l)}]$ .

Full conditional for  $b_2^{\&(l)}$  for  $l = 1, 2, \dots, L$

$$P(b_2^{\&(l)} | ELSE) \propto N(\tilde{b}_2^{\&(l)}, \tilde{\Sigma}_{b_2^{\&(l)}}) \tag{7}$$

where  $\tilde{\Sigma}_{b_2^{\&(l)}} = (d_l^{-1} \otimes \Sigma_E^{-1} \otimes G_g^{-1} + Z_2^T Z_2)^{-1}$  and  $\tilde{b}_2^{\&(l)} = \tilde{\Sigma}_{b_2^{\&(l)}} (Z_2^T) \{ Y^{\&(l)} - X\beta^{\&(l)} - Z_1 b_1^{\&(l)} \}$ .

Therefore, the Gibbs sampler for the BMTME Thompson version should be:

Step 1. Simulate  $\beta$  according to the normal distribution given in Supplementary material (A.1).

Step 2. Simulate  $\mathbf{b}_1^{k(l)}$  for  $l = 1, 2, \dots, L$  according to the normal distribution given in Eq. (6).

Step 3. Simulate  $\mathbf{b}_2^{k(l)}$  for  $l = 1, 2, \dots, L$  according to the normal distribution given in Eq. (7).

Step 4. Then transform back  $\mathbf{Y} = \mathbf{Y}^* \mathbf{Q}^{(-1)T}$ ,  $\boldsymbol{\beta} = \boldsymbol{\beta}^* \mathbf{Q}^{(-1)T}$ ,  $\mathbf{b}_1 = \mathbf{b}_1^* \mathbf{Q}^{(-1)T}$  and  $\mathbf{b}_2 = \mathbf{b}_2^* \mathbf{Q}^{(-1)T}$ .

Step 5. Simulate  $\boldsymbol{\Sigma}_t$  according to the IW distribution given in Supplementary material (A.4).

Step 6. Simulate  $\boldsymbol{\Sigma}_E$  according to the IW distribution given in Supplementary material (A.5).

Step 7. Simulate  $\mathbf{R}_e$  according to the IW distribution given in Supplementary material (A.6).

Step 8. Return to step 1 or terminate when chain length is adequate to meet convergence diagnostics.

The Gibbs sampler for the BMTME Thompson version is similar to the original modified Gibbs sampler except that steps 2 and 3 were replaced for univariate sampling for the transformed random effects ( $\mathbf{b}_1$  and  $\mathbf{b}_2$ ), which are back transformed in Step 4. The advantage of the BMTME Thompson version is that it allows sampling the random effects of  $\mathbf{b}_1$  and  $\mathbf{b}_2$  independently for each trait, which allows improving the speed of the Gibbs sampler because it can be parallelized. However, the remaining parameters ( $\boldsymbol{\beta}$ ,  $\boldsymbol{\Sigma}_t$ ,  $\boldsymbol{\Sigma}_E$ , and  $\mathbf{R}_e$ ) are sampled exactly as the original Gibbs sampler.

### BMTME\_Approx model with SVD

An alternative but only approximate method is to transform the matrix of response variables with SVD as  $\mathbf{Y} = \mathbf{UDV}^T$ , where  $\mathbf{U}$  and  $\mathbf{V}$  are orthogonal matrices called the left and right singular vectors, respectively, of dimensions  $n \times n$  and  $L \times L$ , while  $\mathbf{D}$  is a rectangular diagonal matrix of the singular values of order  $n \times L$  (where only the first  $L$  values of  $\mathbf{D}$  are positive while the rest are zeros). Therefore, the reparametrized model given in Eq. (2) can be rewritten as:

$$\begin{aligned} \mathbf{UDV}^T &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{b}_1 + \mathbf{Z}_2\mathbf{b}_2 + \mathbf{e} \\ \mathbf{Y}^* &= \mathbf{X}\boldsymbol{\beta}^* + \mathbf{Z}_1\mathbf{b}_1^* + \mathbf{Z}_2\mathbf{b}_2^* + \mathbf{e}^* \end{aligned} \tag{8}$$

where  $\mathbf{Y}^* = \mathbf{UD} = \mathbf{YV}$ ,  $\boldsymbol{\beta}^* = \boldsymbol{\betaV}$ ,  $\mathbf{b}_1^* = \mathbf{b}_1\mathbf{V}$ ,  $\mathbf{b}_2^* = \mathbf{b}_2\mathbf{V}$ ,  $\mathbf{e}^* = \mathbf{eV}$ .  $\mathbf{Y}^*$  is of order  $n \times L$ ,  $\boldsymbol{\beta}^*$  is of order  $I \times L$ ,  $\mathbf{b}_1^*$  is of order  $J \times L$ ,  $\mathbf{b}_2^*$  is of order  $IJ \times L$ , and  $\mathbf{e}^*$  is of order  $n \times L$ . Note that the parametrized model does not have the same conceptual definition as the original model (2), first, because  $\boldsymbol{\beta}^*$  is a random matrix and not an unknown constant matrix as the matrix of fixed effects,  $\boldsymbol{\beta}$ , and second, because  $\mathbf{b}_1^*$ ,  $\mathbf{b}_2^*$ , and  $\mathbf{e}^*$  are no longer independent. However, if we fix  $\mathbf{V}$  as part of the subjacent data structure, and we suppose that  $\boldsymbol{\Sigma}_t = \mathbf{VD}_{t1}\mathbf{V}^T$  and  $\mathbf{R}_e = \mathbf{VD}_{te}\mathbf{V}^T$  (they have a restricted parameter space, Lin and Smith (1990)),  $\mathbf{b}_1^*$  is distributed as a matrix variate normal distribution as  $\text{NM}_{J \times L}(\mathbf{0}, \mathbf{G}_g, \mathbf{D}_{t1})$ ,  $\mathbf{b}_2^* \sim \text{NM}_{IJ \times L}(\mathbf{0}, \boldsymbol{\Sigma}_E \otimes \mathbf{G}_g, \mathbf{D}_{t1})$ , and  $\mathbf{e}^* \sim \text{NM}_{n \times L}(\mathbf{0}, \mathbf{I}_n, \mathbf{D}_{te})$ , where

$\mathbf{D}_{t1}$  and  $\mathbf{D}_{te}$  are diagonal variance–covariance matrices of dimension  $L \times L$ . It is important to point out that, under this approximate model, the (co)variance matrix for environment is assumed an identity matrix,  $\boldsymbol{\Sigma}_E = \mathbf{I}_I$ .

To use the existing software, we replaced the distribution of transformed random effects  $\mathbf{b}_2^*$  with  $\mathbf{b}_2^* \sim \text{NM}_{IJ \times L}(\mathbf{0}, \mathbf{I}_I \otimes \mathbf{G}_g, \mathbf{D}_{t2})$ , where  $\mathbf{D}_{t2}$  is another diagonal variance–covariance matrix of traits of dimension  $L \times L$ . With this, the resulting model (8) can be estimated with the R package BGLR, which is appropriate for univariate analysis. From Eq. (8), it is clear that the parameter estimates and predicted values of the original model (Eq. (2)) without transformation can be approximated as:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}^* \mathbf{V}^T \tag{9}$$

$$\hat{\mathbf{b}}_1 = \hat{\mathbf{b}}_1^* \mathbf{V}^T \tag{10}$$

$$\hat{\mathbf{b}}_2 = \hat{\mathbf{b}}_2^* \mathbf{V}^T \tag{11}$$

$$\hat{\mathbf{Y}}^* = \hat{\mathbf{Y}}^* \mathbf{V}^T \quad ; \quad \text{with} \quad \hat{\mathbf{Y}}^* = \mathbf{X}\hat{\boldsymbol{\beta}}^* + \mathbf{Z}_1\hat{\mathbf{b}}_1^* + \mathbf{Z}_2\hat{\mathbf{b}}_2^* \tag{12}$$

$$\hat{\boldsymbol{\Sigma}}_{t1} = \mathbf{V}\hat{\mathbf{D}}_{t1}\mathbf{V}^T \tag{13}$$

$$\hat{\boldsymbol{\Sigma}}_{t2} = \mathbf{V}\hat{\mathbf{D}}_{t2}\mathbf{V}^T \tag{14}$$

$$\hat{\boldsymbol{\Sigma}}_{te} = \mathbf{V}\hat{\mathbf{D}}_{te}\mathbf{V}^T \tag{15}$$

### Steps for implementing the proposed BMTME\_Approx model

Step 1: De-correlate the original traits with the SVD as  $\mathbf{Y} = \mathbf{UDV}^T$  and use it as response variable  $\mathbf{Y}^* = \mathbf{UD} = \mathbf{YV}$ .

Step 2: Implement model (1) but using one column at a time of the uncorrelated matrix  $\mathbf{Y}^*$  as the response variable. This means that a total of  $L$  single analyses are done with the model in Eq. (1).

Step 3: With the output of Step 2, the predicted values can be calculated in terms of the transformed values with  $\hat{\mathbf{Y}}^* = \mathbf{X}\hat{\boldsymbol{\beta}}^* + \mathbf{Z}_1\hat{\mathbf{b}}_1^* + \mathbf{Z}_2\hat{\mathbf{b}}_2^*$ . With this new output, we construct the diagonal matrices  $\hat{\mathbf{D}}_{t1}$ ,  $\hat{\mathbf{D}}_{t2}$ , and  $\hat{\mathbf{D}}_{te}$ .

Step 4: Finally, with Eq. (12), we obtain the predicted values in terms of the original traits. Furthermore, with Eqs. (9–11, 13–15), we obtain the parameter estimates of the beta coefficients,  $\boldsymbol{\beta}$ , random effects  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , as well as the variance–covariance matrices of traits corresponding to traits in the first interaction term,  $\boldsymbol{\Sigma}_{t1}$ , for traits in the second interaction term,  $\boldsymbol{\Sigma}_{t2}$ , and the residual variance–covariance matrix of traits,  $\mathbf{R}_e$ . The R code for implementing this proposed BMTME\_Approx model in BGLR is given in Supplementary material.



For implementing the proposed BMTME\_Approx model with random cross-validation, we will follow the four-step procedure exactly as described above, except for the first step, which is modified. Now, in Step 1, we will decorrelate the traits in the training data set ( $Y_t$ ) with the SVD as  $Y_t = U_t D_t V_t^T$ , where the subscript  $t$  denotes that these matrices were estimated with the training data set. Then we transform the response variable into  $Y^* = U_t D_t = Y_t U_t$ , and expand  $Y^*$  with the number of rows that are the same size as the testing data set; the expanded rows should all be replaced with NA, to represent the missing values. The positions of the expanded rows with missing values will correspond to the testing data set. Once this is done, the remaining steps must be followed exactly as in the above procedure that is described for the full data set. Note that the dimension of the response variable matrix with the training data set has fewer rows than the full data set of response variables.

It is important to point out that the BMTME model was built to estimate only one variance-covariance matrix of traits,  $\Sigma_t$ , involved in the two interaction terms, genotype  $\times$  trait and genotype  $\times$  environment  $\times$  trait. However, the BMTME\_Approx model allows estimating a variance-covariance matrix of traits for each interaction term in which the traits are involved. The BMTME model is also able to estimate an unstructured variance-covariance matrix for the environments,  $\Sigma_E$ ; this is not reported for the BMTME\_Approx model because an identity matrix is assumed.

### Hyperparameters

First we provide the hyperparameters for the BMTME model:  $\beta \sim MN_{I \times L}(\beta_0, I_t, S_{\beta t})$ ,  $b_1 | \Sigma_t \sim MN_{J \times L}(\mathbf{0}, G_g, \Sigma_t)$ ,  $b_2 | \Sigma_t, \Sigma_E \sim MN_{I \times L}(\mathbf{0}, \Sigma_E \otimes G_g, \Sigma_t)$ ,  $\Sigma_E \sim IW(\nu_E = 5, S_E = S_E)$ ,  $\Sigma_t \sim IW(\nu_t = 5, S_t = S_t)$ , and  $R_e \sim IW(\nu_e = 5, S_e = S_e)$ ;  $\beta_0$  was obtained as the least square of each trait. The remaining hyperparameters  $S_{\beta t}$ ,  $S_t$ ,  $S_E$ , and  $S_e$  are given in Supplementary material. The hyperparameters for the BMTME\_Approx were exactly the same to those of the BMTME, but for the univariate analysis, it were as those used in the BGLR software (Pérez-Rodríguez and de los Campos 2014). The proposed Gibbs sampler was implemented in the R-software (R Core Team 2018). A total of 60,000 iterations were performed with a burn-in of 20,000, so that 40,000 samples were used for inference. To eliminate potential problems due to the autocorrelation function (ACF), we considered a thinning of 5. The convergence of the MCMC chains was monitored using trace plots, ACF and Gelman-Rubin diagnostics. We provide weakly informative priors to implement the proposed models. It is important to point out that the proposed BMTME\_Approx model only works for normally distributed traits. However,

when there is considerable departure from normality, we suggest using independent component analysis (ICA) instead of SVD for transforming the matrix of response variables ( $Y$ ) (see Supplementary material for its implementation).

### Simulated data set 1 and data set 2

To test the proposed models and methods, we simulated multiple-trait and multiple-environment data using the model in Eq. (2). For this first data set, we used the following parameters: 3 environments, 3 traits, 200 genotypes, and 1 replication of the environment-trait-genotype combination. We assumed that  $\beta^T = [13, 10, 5, 12, 8, 7, 11, 9, 6]$ , where the first three beta coefficients belong to traits 1, 2, and 3 in environment 1, the second three values belong to the three traits in environment 2, and the last three belong to environment 3. We assumed that the GRM is known and equal to  $G_g = 0.3I_{200} + 0.7J_{200}$ , where  $I_{200}$  is an identity matrix of order 200 and  $J_{200}$  is a matrix of order  $200 \times 200$  of ones. The parameters used for building the GRM were chosen to provide a high-level relationship between lines (Montesinos-López et al. 2016).

Therefore, the total number of observations is  $3 \times 200 \times 3 \times 1 = 1800$ , that is, 600 for each trait. Since a covariance matrix can be expressed in terms of a correlation matrix ( $R_r$ ) and a standard deviation matrix ( $D_r^{1/2}$ ) as:  $\Sigma_r = D_r^{1/2} R_r D_r^{1/2}$ , with  $r = t, E, e$ , where  $r = t$  represents the genetic covariance between traits,  $r = E$  represents the genetic covariance matrix between environments, and  $r = e$  represents the residual covariance matrix between traits. For the three covariance matrices ( $r = t, E, e$ ), we used  $R_r = 0.75I_3 + 0.25J_3$ , where  $J_3$  is a matrix of order  $3 \times 3$  of ones, and  $D_t^{1/2} = \text{diag}(0.9, 0.8, 0.9)$ ,  $D_E^{1/2} = \text{diag}(0.5, 0.65, 0.75)$  and  $D_e^{1/2} = \text{diag}(0.6, 0.42, 0.33)$ . For the second data set, the parameters used in the simulation were:  $\beta^T = [13, 12.5, 12, 11.5, 11, 10.5, 10, 12, 11.5, 11, 10.5, 10.5, 10, 10, 11, 11.5, 12, 12, 11, 10, 10.5]$ , where the first seven beta coefficients belong to traits 1–7 in environment 1, the second seven values to the 7 traits in environment 2, and the last seven belong to environment 3.

The matrix of the relationship between lines was generated as  $G_g = 0.3I_{200} + 0.7J_{200}$  and was equal to the first simulation data set. Here the total number of observations is  $3 \times 200 \times 7 \times 1 = 4200$ , that is, 600 for each trait. For two of the three covariance matrices ( $f = t, e$ ), we used

$$R_f = \begin{bmatrix} 1.000 & 0.970 & 0.944 & 0.917 & 0.890 & 0.862 & 0.833 \\ - & 1.000 & 0.925 & 0.899 & 0.872 & 0.844 & 0.816 \\ - & - & 1.000 & 0.875 & 0.849 & 0.822 & 0.794 \\ - & - & - & 1.000 & 0.825 & 0.798 & 0.772 \\ - & - & - & - & 1.000 & 0.775 & 0.748 \\ - & - & - & - & - & 1.000 & 0.725 \\ - & - & - & - & - & - & 1.000 \end{bmatrix} \text{ and}$$

$D_f^{1/2} = \text{diag}(1, 1.0003, 1.0003, 1, 0.9996, 1, 1.0003)$ , while  $R_E = D_E^{1/2} = \text{diag}(1, 1, 1)$ , that is, we assumed independence between the environments. It is important to point out that this second data set has a high genetic and environmental correlation between traits.

## Experimental data sets

### Maize data set

The first real data set used for implementing the proposed model is composed of 309 double-haploid maize lines. Traits available in this data set include grain yield (GY), anthesis-silking interval (ASI), and plant height (PH); each of these traits was evaluated in three optimum rain-fed environments (EBU, KAT, and KTI). After editing, information from 158,281 markers was used. This data set was also used by Montesinos-López et al. (2016) and includes best linear unbiased estimates (BLUEs) obtained based on a mixed model analysis of individual trials of a first analysis.

### Wheat data set

Here we present information on the second real data set used for implementing the proposed model. This real data set is composed of 250 wheat lines that were extracted from a large set of 39 yield trials grown during the 2013–2014 crop season in Ciudad Obregon, Sonora, Mexico (Rutkoski et al. 2016). The traits under study were days to heading (DH), GY, PH, and the green normalized difference vegetation index (NDVI). Each of these traits was evaluated in three environments (Bed2IR, Bed5IR, and Drip). The marker information used after editing was from 12,083 markers, this data set also used by Montesinos-López et al. (2016); those interested in obtaining more details about this data set can consult this publication. The phenotypes of each trait are BLUEs obtained after a first analysis where they were adjusted by the experimental field design.

### High-throughput (HTP) data set

This data set belongs to an experiment that used HTP wheat plant phenotyping conducted at Ciudad Obregon, Sonora, México. The data set is comprised of 976 wheat lines that were extracted from a large set of 1170 lines from the CIMMYT Global Wheat Program. The following traits were under study: GY, DH, red normalized difference vegetation index (RNDVI), green normalized difference vegetation index (GNDVI), simple ratio (SRa), ratio analysis of reflectance spectra chlorophyll a (RARSa), ratio analysis of reflectance spectra chlorophyll b (RARSb), ratio analysis of reflectance spectra chlorophyll c (RARSb), normalized pheophytinization index (NPQI), and

photochemical reflectance index (PR). Each of these traits was evaluated in three environments (drought, irrigated, and reduced irrigation). After marker editing, information from 1448 markers was used. This data set was also used by Montesinos-López et al. (2017b, c). The phenotypes of each trait are BLUEs obtained after a first analysis where they were adjusted by the experimental field design.

### Large EYT set

This data set belongs to CIMMYT's three elite yield trial (EYT) nurseries, consisting of 2505 lines of wheat, genotyped by genotyping-by-sequencing (GBS). These were evaluated for GY, DH, and PH in five environments (BED\_5IR, FLAT\_5IR, BED\_2IR, FLAT\_DRIP, and LHT) evaluated in Ciudad Obregon, Mexico, under bed and flat planting systems. The EYT nurseries were sown in 39 trials, each containing 28 lines and two checks that were arranged in an alpha lattice design with three replications and six blocks. The nurseries were evaluated for the three traits under study on a plot basis during 2014 (EYT 13–14), 2015 (EYT 14–15), and 2016 (EYT 15–16). We used BLUEs as observed values of the breeding lines resulting from adjusting for the corresponding experimental design. All the 2505 lines were genotyped using GBS (Elshire et al. 2011; Poland et al. 2012) at Kansas State University, with an Illumina HiSeq2500 for obtaining genome-wide markers. Markers with missing data >60% (minor allele frequency <5% and percentage of heterozygosity >10%) were removed, and we obtained 2038 markers, which were used for the analysis. Also, the traits used are BLUEs obtained after a first analysis where they were adjusted by the experimental field design in each trial.

### Random cross-validation scheme

For testing the prediction ability of the proposed models, the BMTME\_Approx model and the BMTME model, we implemented a type of cross-validation where all the traits are missing in some individuals and the information of some individuals is missing in some environments (that is, their lines and traits are missing), but in at least one environment there is information available on those individuals. We implemented a 20 random cross-validation scheme for all the data sets under study, with the exception of the HTP and the large EYT data sets, in which we implemented 10 random cross-validations. For the 20 random cross-validation scheme, in each partition we assigned 20% of the data to the testing set and the remaining 80% of the data to the training set, while for the 10 random cross-validation scheme, we assigned 30% of the data to the testing set and 70% to the training set. The models were fitted with the information in the training data sets, and

the prediction accuracy was evaluated with the testing data sets.

The metrics used for reporting the prediction accuracy were the Pearson's correlation (Cor) and the mean square error of prediction (MSEP) obtained by averaging the information of the 20 (or 10) random partitions resulting from the testing data sets. The models were implemented in the R package (R Core Team 2018) and the proposed BMTME\_Approx model can be implemented in the BGLR package of de los Campos and Pérez-Rodríguez (2014). On the other hand, the BMTME model was implemented in R with the model proposed by Montesinos-López et al. (2016).

### Data repository

The phenotypic and genotypic information of the experimental wheat and maize data sets included in this study can be downloaded from the link <http://hdl.handle.net/11529/10646> (Montesinos-López et al. 2016). This link includes phenotypic data on maize (Data.maize) and wheat (Data.trigo), as well as genomic data on maize (G.maize) and wheat (G.trigo).

The HTP data and materials used in this study can be downloaded from the link given in Montesinos-Lopez et al. (2017c): <http://hdl.handle.net/11529/10693> that contains a file corresponding to the phenotypic and band data for each environment, Drought.Phe and Bands.RData, EarlyHeat.Phe, and Bands.RData, Irrigated.Phe and Bands, RData, Irrigated.Phe and Bands.RData. The large EYT data (Data.EYT.2018) can be downloaded from the link <http://hdl.handle.net/11529/10547920>.

## Results

The results are described in two main sections: The first section presents the results of the simulated data sets, while the second presents the results of the real data sets. It is important to point out that we do not present results for the BMTME Thompson version because it is only a different reparametrization of the original BMTME model.

### Simulated data sets

#### Data set 1

First, we present the parameter estimates of the BMTME and the BMTME\_Approx models. Table 1 shows that the beta coefficients of both models are similar. In general, the beta coefficients of the BMTME model are larger than those of the BMTME\_Approx model, the smallest difference is

3% observed in environment 1 and trait 1, and the largest difference is 15.7% observed in environment 3 and trait 3. The variance-covariance matrix of traits ( $\Sigma_t$ ) of the BMTME model is quite similar to the variance-covariance matrices of the BMTME\_Approx model ( $\Sigma_{t1}$ ,  $\Sigma_{t2}$ ). The variance-covariance components of the residual of both models are quite similar, with the smallest difference (1.6%) observed in trait 2 and trait 1 and the largest difference (33.7%) observed in trait 3 and trait 2. It is worth pointing out again that the BMTME and BMTME\_Approx models are different; consequently, we cannot expect the exact same parameter estimates. Finally, when comparing the observed versus the predicted values for each trait using Pearson's correlation and MSEP, we observed that the BMTME\_Approx model produced predicted values that are very similar to those of the BMTME model (see Table 1).

With regard to prediction accuracy, Table 2 shows that the BMTME model was the best: In six out of the nine trait-environment combinations, it was superior to the BMTME\_Approx model in terms of Pearson's correlation and MSEP. On average, the BMTME model was superior to the BMTME\_Approx model by 0.94% and 0.88% in terms of Pearson's correlation and MSEP, respectively. From the results of these simulated data, it is evident that the BMTME\_Approx model is very similar to the BMTME model in terms of prediction accuracy (Table 2).

#### Data set 2

First, we present the parameter estimates of the BMTME and BMTME\_Approx models. Table 3 shows that the beta coefficients of both models are similar, and in general, the beta coefficients of the BMTME model are larger than those of the BMTME\_Approx model. The smallest difference is 5.15%, which is observed in environment 3 and trait 6, while the largest difference is 12.37%, observed in environment 1 and trait 5. The variance-covariance matrix of traits ( $\Sigma_t$ ) in the BMTME model is very similar to the variance-covariance matrices of the BMTME\_Approx model ( $\Sigma_{t1}$ ,  $\Sigma_{t2}$ ). The variance-covariance components of the residual of both models are quite similar, with the smallest difference (0.48%) observed in trait 7 and trait 1 and the largest difference (27.38%) observed in the variance of trait 5. It is worth reiterating that, with the BMTME and BMTME\_Approx, we should not expect exactly the same parameter estimates, as the models are different. Finally, when comparing the observed versus the predicted values for each trait using Pearson's correlation and MSEP, the BMTME\_Approx model produced predicted values that were slightly better than those of the BMTME model (see Table 3). In terms of Pearson's correlation, the smallest



**Table 1** Parameter estimates (posterior means) of the BMTME model and the BMTME\_Approx model for simulated data set 1

	True values			BMTME			BMTME_Approx		
	$\beta$			$\hat{\beta}$			$\hat{\beta}$		
	Trait1	Trait2	Trait3	Trait1	Trait2	Trait3	Trait1	Trait2	Trait3
Env1	13	12	11	14.173	11.516	11.331	13.747	10.982	10.387
Env2	10	8	9	9.514	7.993	9.767	9.173	7.438	8.915
Env3	5	7	6	4.180	6.286	6.540	3.906	5.915	5.512
	$\Sigma_t$			$\hat{\Sigma}_t$			$\hat{\Sigma}_{t1}$		
	Trait1	Trait2	Trait3	Trait1	Trait2	Trait3	Trait1	Trait2	Trait3
	Trait1	0.900	0.212	0.225	0.979	0.297	0.382	0.911	0.266
Trait2	0.212	0.800	0.212	0.297	1.096	0.288	0.266	0.970	0.128
Trait3	0.225	0.212	0.900	0.382	0.288	0.803	0.331	0.128	0.934
	$\Sigma_E$			$\hat{\Sigma}_E$			$\hat{\Sigma}_{t2}$		
	Env1	Env2	Env3	Env1	Env2	Env3	Trait1	Trait2	Trait3
	Env1	0.500	0.143	0.153	0.374	0.184	0.052	0.571	0.197
Env2	0.143	0.650	0.175	0.184	0.521	0.201	0.197	0.438	0.216
Env3	0.153	0.175	0.750	0.052	0.201	0.412	0.193	0.216	0.487
	$R_e$			$\hat{R}_e$			$\hat{R}_e$		
	Trait1	Trait2	Trait3	Trait1	Trait2	Trait3	Trait1	Trait2	Trait3
	Trait1	0.600	0.125	0.111	0.775	0.180	0.202	0.603	0.177
Trait2	0.125	0.420	0.093	0.180	0.436	0.150	0.177	0.475	0.201
Trait3	0.111	0.093	0.330	0.202	0.150	0.474	0.171	0.201	0.522
				Prediction accuracy			Prediction accuracy		
				Trait1	Trait2	Trait3	Trait1	Trait2	Trait3
	Cor	—	—	—	0.986	0.978	0.966	0.989	0.978
MSEP	—	—	—	0.479	0.247	0.307	0.479	0.247	0.307

Cor and MSEP denote Pearson’s correlation and mean square error of prediction, respectively, between the observed and predicted values. True values denotes the true parameter values used for simulating the data  $\beta$  denotes the beta coefficients,  $\Sigma_t$  denotes the genetic (co)variance matrix of traits,  $\Sigma_E$  denotes the genetic (co)variance matrix of environments,  $R_e$  denotes the residual (co)variance matrix of traits, and the symbol hat (^) denotes estimates of the corresponding parameters

**Table 2** Average Pearson’s correlation (Cor) and average mean square error of prediction (MSEP) for each trait–environment combination for simulated data set 1 resulting from the testing set of the 20 random partitions

Trait_Env	BMTE				BMTME_Approx			
	Cor	SE	MSEP	SE	Cor	SE	MSEP	SE
y1_Env1	0.326	0.032	0.714	0.026	<b>0.338</b>	0.036	<b>0.705</b>	0.026
y2_Env1	<b>0.515</b>	0.029	<b>0.554</b>	0.021	0.508	0.031	0.561	0.022
y3_Env1	<b>0.472</b>	0.026	<b>0.597</b>	0.027	0.461	0.025	0.606	0.027
y1_Env2	0.344	0.019	1.062	0.057	<b>0.351</b>	0.019	<b>1.061</b>	0.057
y2_Env2	<b>0.546</b>	0.028	<b>0.579</b>	0.027	0.527	0.029	0.608	0.028
y3_Env2	<b>0.418</b>	0.032	<b>0.733</b>	0.041	0.408	0.033	0.734	0.043
y1_Env3	0.411	0.026	0.978	0.041	<b>0.418</b>	0.023	<b>0.976</b>	0.040
y2_Env3	<b>0.529</b>	0.023	<b>0.700</b>	0.029	0.523	0.023	0.714	0.028
y3_Env3	<b>0.456</b>	0.019	<b>0.556</b>	0.025	0.444	0.019	0.564	0.026
Average	<b>0.446</b>	0.026	<b>0.719</b>	0.033	0.442	0.026	0.726	0.033

The best predictions for each trait–environment combination are in bold

**Table 3** Parameter estimates (posterior means) of the BMTME model and the BMTME\_Approx model for simulated data set 2

True values		BMTME														BMTME_Approx																		
$\beta$		$\hat{\beta}$														$\hat{\beta}$																		
Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7							
Env1	13	12.5	12	11.5	11	10.5	10	13.245	11.618	10.645	11.121	9.486	11.027	11.333	11.899	10.328	9.483	9.979	8.312	9.884	10.400	10.394	9.056	9.958	8.546	7.551	10.175	8.148						
Env2	12	11.5	11	10.5	10	10	10	11.335	9.981	10.750	9.283	8.544	10.888	8.666	10.394	9.056	9.958	8.546	7.551	10.175	8.148	10.344	8.714	10.271	9.265	9.167	10.399	9.101						
Env3	11	11.5	12	12	11	10	10.5	11.018	9.346	10.925	9.811	10.056	10.963	9.676	10.344	8.714	10.271	9.265	9.167	10.399	9.101	10.344	8.714	10.271	9.265	9.167	10.399	9.101						
$\Sigma_r$		$\hat{\Sigma}_r$														$\hat{\Sigma}_r$																		
Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7							
1.000	0.970	0.944	0.917	0.890	0.862	0.833	0.663	0.620	0.575	0.565	0.577	0.548	0.550	0.488	0.448	0.448	0.474	0.532	0.472	0.580	0.630	0.686	0.649	0.633	0.549	0.663	0.630							
0.970	1.001	0.925	0.899	0.872	0.844	0.816	0.620	0.610	0.547	0.524	0.548	0.546	0.416	0.686	0.636	0.551	0.534	0.474	0.580	0.630	0.686	0.636	0.551	0.534	0.474	0.560	0.528							
0.944	0.925	1.001	0.875	0.848	0.822	0.794	0.575	0.547	0.578	0.495	0.550	0.528	0.391	0.649	0.551	0.718	0.559	0.532	0.635	0.523	0.649	0.551	0.718	0.559	0.532	0.635	0.523							
0.917	0.899	0.875	1.000	0.824	0.799	0.772	0.565	0.524	0.495	0.607	0.488	0.488	0.448	0.633	0.534	0.559	0.635	0.472	0.550	0.528	0.633	0.534	0.559	0.635	0.472	0.550	0.528							
0.890	0.872	0.848	0.824	0.999	0.774	0.748	0.577	0.548	0.550	0.488	0.691	0.539	0.369	0.549	0.474	0.532	0.472	0.580	0.524	0.460	0.549	0.474	0.532	0.472	0.580	0.524	0.460							
0.862	0.844	0.822	0.799	0.774	1.000	0.725	0.570	0.546	0.528	0.488	0.539	0.634	0.338	0.663	0.560	0.635	0.550	0.524	0.755	0.550	0.663	0.560	0.635	0.550	0.524	0.755	0.550							
0.833	0.816	0.794	0.772	0.748	0.725	1.001	0.477	0.416	0.391	0.448	0.369	0.338	0.631	0.630	0.528	0.523	0.528	0.460	0.550	0.665	0.630	0.528	0.523	0.528	0.460	0.550	0.665							
$\Sigma_E$		$\hat{\Sigma}_E$														$\hat{\Sigma}_E$																		
Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3	Env1	Env2	Env3					
1	0	0	—	—	—	0.649	0.261	-0.026	—	—	—	—	—	—	0.715	0.620	0.582	0.564	0.481	0.592	0.620	0.582	0.564	0.481	0.592	0.546	0.546							
0	1	0	—	—	—	0.261	0.681	0.033	—	—	—	—	—	—	0.620	0.582	0.497	0.478	0.413	0.447	0.582	0.497	0.478	0.413	0.497	0.447	0.447							
0	0	1	—	—	—	-0.026	0.033	0.449	—	—	—	—	—	—	0.582	0.497	0.648	0.493	0.465	0.466	0.497	0.648	0.493	0.465	0.560	0.466	0.466							
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.564	0.478	0.493	0.573	0.430	0.462	0.478	0.493	0.573	0.430	0.485	0.462	0.462							
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.481	0.413	0.465	0.430	0.555	0.423	0.413	0.465	0.430	0.555	0.439	0.423	0.423							
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.592	0.497	0.560	0.485	0.439	0.495	0.497	0.560	0.485	0.439	0.701	0.495	0.495							
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.546	0.447	0.466	0.462	0.423	0.626	0.447	0.466	0.462	0.423	0.495	0.626	0.626							
$R_e$		$\hat{R}_e$														$\hat{R}_e$																		
Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7							
1.000	0.970	0.944	0.917	0.890	0.862	0.833	1.364	1.340	1.297	1.211	1.270	1.164	1.162	1.492	1.274	1.262	1.196	1.063	1.280	1.167	1.364	1.340	1.297	1.211	1.270	1.164	1.162	1.492	1.274	1.262	1.196	1.063	1.280	1.167
0.970	1.001	0.925	0.899	0.872	0.844	0.816	1.340	1.389	1.281	1.193	1.255	1.150	1.147	1.274	1.177	1.084	1.018	0.919	1.084	0.974	1.340	1.389	1.281	1.193	1.255	1.150	1.147	1.274	1.177	1.084	1.018	0.919	1.084	0.974
0.944	0.925	1.001	0.875	0.848	0.822	0.794	1.297	1.281	1.368	1.169	1.216	1.112	1.107	1.262	1.084	1.320	1.055	0.975	1.158	1.046	1.297	1.281	1.368	1.169	1.216	1.112	1.107	1.262	1.084	1.320	1.055	0.975	1.158	1.046
0.917	0.899	0.875	1.000	0.824	0.799	0.772	1.211	1.193	1.169	1.267	1.140	1.049	1.027	1.196	1.018	1.055	1.170	0.927	1.071	0.980	1.211	1.193	1.169	1.267	1.140	1.049	1.027	1.196	1.018	1.055	1.170	0.927	1.071	0.980
0.890	0.872	0.848	0.824	0.999	0.774	0.748	1.270	1.255	1.216	1.140	1.450	1.082	1.108	1.063	0.919	0.975	0.927	1.053	0.961	0.906	1.270	1.255	1.216	1.140	1.450	1.082	1.108	1.063	0.919	0.975	0.927	1.053	0.961	0.906

Table 3 (continued)

$R_e$		$\hat{R}_e$							$\hat{R}_e$												
		Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7						
Trait6	0.862	0.844	0.822	0.799	0.774	1.000	0.725	1.164	1.150	1.112	1.049	1.082	1.340	0.981	1.280	1.084	1.158	1.071	0.961	1.394	1.086
Trait7	0.833	0.816	0.794	0.772	0.748	0.725	1.001	1.162	1.147	1.107	1.027	1.108	0.981	1.323	1.167	0.974	1.046	0.980	0.906	1.086	1.208

		Prediction accuracy							Prediction accuracy						
		Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7	Trait1	Trait2	Trait3	Trait4	Trait5	Trait6	Trait7
Cor	—	0.697	0.672	0.704	0.709	0.727	0.756	0.773	0.770	0.783	0.771	0.784	0.835	0.811	
MSEP	—	1.183	1.170	1.062	1.221	1.131	1.112	1.006	1.009	0.987	0.923	1.023	0.969	0.933	

Cor and MSEP denote Pearson's correlation and mean square error of prediction, respectively, between the observed and predicted values. True values denotes the true parameter values used for simulating the data.  $\beta$  denotes the beta coefficients,  $\Sigma_i$  denotes the genetic (co)variance matrix of traits,  $\Sigma_E$  denotes the genetic (co)variance matrix of environments,  $R_e$  denotes the residual (co)variance matrix of traits, and the symbol hat (^) denotes estimates of the corresponding parameters

difference in favor of the BMTME\_Approx model was 6.73% observed in trait 6, while the largest difference, also in favor of the BMTME\_Approx model, was 14.16% observed in trait 3. In terms of MSEP, the smallest difference was 13.81% in trait 1 and the largest difference was 19.34% in trait 5, both in favor of the BMTME\_Approx model.

Table 4 shows that the proposed approximate model, BMTME\_Approx, was better than the BMTME model in terms of Pearson's correlation, as shown in 14 out of the 21 trait-environment combinations. However, in terms of MSEP, the BMTME model was superior to the the BMTME\_Approx model, as exemplified by 12 out of the 21 trait-environment combinations. On average, in terms of Pearson's correlation, the BMTME\_Approx model was better than the BMTME model by 7.24%, while in terms of MSEP, both models were, on average, almost identical.

### Experimental data sets

#### Maize data set

First, we compared the parameter estimates of the two models (BMTME and BMTME\_Approx). Table 5 shows that the beta coefficients of the proposed BMTME\_Approx model are all similar to those of the BMTME model. The variance-covariance matrix of traits ( $\Sigma_i$ ) in the BMTME model is quite similar to the variance-covariance matrices of the BMTME\_Approx model ( $\Sigma_{i1}, \Sigma_{i2}$ ). When comparing the variance-covariance components of the residual of both models, we observe that 6 out of the 9 terms are not significantly different; however, the remaining 3 terms are quite different, as they belong to the covariance of trait PH with the other traits. Finally, when comparing the observed versus the predicted values for each trait using Pearson's correlation and MSEP, we see that the BMTME\_Approx model is slightly better, since in Pearson's correlation, it outperformed the BMTME model in 2 out of the 3 traits and, on average, the BMTME\_Approx model was 2.1% (for trait GY) and 1.4% (for trait ASI) better than the BMTME model. In terms of MSEP, the BMTME\_Approx model was better than the BMTME model in 2 out of the 3 traits, and, on average, the BMTME\_Approx model was 7% better (for trait GY) and 4.11% better (for trait ASI) than the BMTME model (see Table 5). In general, the parameter estimates and predictions are relatively similar in this data set.

Next, in Table 6, we see that the proposed approximate model, BMTME\_Approx, was better than the BMTME model in terms of Pearson's correlation, as shown in 5 out of the 9 trait-environment combinations. However, in terms of MSEP, the BMTME model was superior to the BMTME\_Approx model in 7 out of the 9 trait-environment combinations. On average, the BMTME\_Approx model

**Table 4** Average Pearson’s correlation (Cor) and average mean square error of prediction (MSEP) for each trait–environment combination for simulated data set 2 resulting from the testing set of the 20 random partitions

Trait_Env	BMTME				BMTME_Approx			
	Cor	SE	MSEP	SE	Cor	Se	MSEP	SE
y1_1	0.279	0.035	<b>1.461</b>	0.049	<b>0.323</b>	0.036	1.469	0.047
y2_1	0.251	0.035	<b>1.473</b>	0.047	<b>0.287</b>	0.036	1.477	0.043
y3_1	0.278	0.032	1.486	0.066	<b>0.324</b>	0.029	<b>1.472</b>	0.060
y4_1	0.251	0.032	<b>1.408</b>	0.051	<b>0.288</b>	0.031	1.410	0.050
y5_1	0.272	0.038	<b>1.721</b>	0.074	<b>0.298</b>	0.040	1.720	0.072
y6_1	0.269	0.038	<b>1.311</b>	0.065	<b>0.275</b>	0.037	1.313	0.063
y7_1	0.227	0.034	1.419	0.061	<b>0.261</b>	0.033	<b>1.411</b>	0.059
y1_2	0.221	0.024	1.666	0.095	<b>0.265</b>	0.024	<b>1.648</b>	0.091
y2_2	0.203	0.026	1.608	0.089	<b>0.258</b>	0.024	<b>1.584</b>	0.085
y3_2	0.156	0.027	1.637	0.095	<b>0.158</b>	0.028	<b>1.628</b>	0.093
y4_2	0.237	0.024	1.515	0.069	<b>0.255</b>	0.023	<b>1.513</b>	0.066
y5_2	<b>0.226</b>	0.031	1.681	0.084	0.221	0.034	<b>1.680</b>	0.079
y6_2	0.166	0.032	1.904	0.069	<b>0.188</b>	0.034	<b>1.882</b>	0.065
y7_2	<b>0.314</b>	0.029	<b>1.572</b>	0.100	0.306	0.025	1.606	0.101
y1_3	<b>0.076</b>	0.036	<b>1.753</b>	0.077	0.068	0.034	1.763	0.077
y2_3	<b>0.057</b>	0.035	<b>1.818</b>	0.074	0.051	0.033	1.824	0.075
y3_3	<b>0.080</b>	0.031	<b>1.661</b>	0.077	0.067	0.031	1.677	0.077
y4_3	<b>0.107</b>	0.043	<b>1.636</b>	0.066	0.100	0.042	1.648	0.064
y5_3	0.103	0.032	<b>1.780</b>	0.068	<b>0.109</b>	0.035	1.784	0.067
y6_3	0.065	0.041	<b>1.761</b>	0.091	<b>0.070</b>	0.036	1.762	0.090
y7_3	<b>0.119</b>	0.034	1.680	0.050	0.096	0.032	<b>1.700</b>	0.052
Average	0.189	0.033	<b>1.617</b>	0.072	<b>0.203</b>	0.032	1.618	0.070

The best predictions for each trait–environment combination are in bold

was better than the BMTME model by 11.72% in terms of Pearson’s correlation, while conversely, the BMTME model was better than the BMTME\_Approx by 37.94% (Table 6) in terms of MSEP.

**Wheat data set**

First, we compared the parameter estimates of the two models (BMTME and BMTME\_Approx) and then their prediction accuracy. Table 7 shows that the beta coefficients of the proposed BMTME\_Approx model are similar to those of the BMTME in 9 out of the 12 parameters; however, there are substantial differences in three of the beta coefficients corresponding to trait NDVI. The variance–covariance matrix of traits ( $\Sigma_t$ ) in the BMTME model is significantly different from the variance–covariance matrices of the BMTME\_Approx model ( $\Sigma_{t1}$ ,  $\Sigma_{t2}$ ).

When comparing the variance–covariance components of the residual of both models, it is evident that 9 out of the 16 terms are not notably different, while the remaining 7 terms are. These very different terms belong to the

covariance of trait NDVI with the other traits. Finally, when we compared the models in terms of the observed versus the predicted values for each trait using Pearson’s correlation and MSEP, we saw that the BMTME\_Approx model was marginally better, and it was superior to the BMTME model for Pearson’s correlation in 3 out of the 4 traits and, on average, 0.99% better for trait DH, 22.04% better for trait GY, and 17.16% better for trait PH. Furthermore, for MSEP, the BMTME\_Approx model was better than the BMTME model in 3 out of the 4 traits and, on average, 18.88% better for trait DH, 48.85% better for trait GY, and 57.5% better for trait PH (see Table 7).

When comparing both models in terms of prediction accuracy with only the testing set of the 20 random partitions implemented, Table 8 shows that the proposed alternative model, BMTME\_Approx, was better than the BMTME model in 8 out of the 12 trait–environment combinations in terms of Pearson’s correlation and in 7 out of the 12 trait–environment combinations in terms of MSEP. On average, the BMTME\_Approx model was also better than the BMTME model by 2.1% in terms of Pearson’s correlation and by 6.87% in terms of MSEP (Table 8).

**Table 5** Parameter estimates (posterior means) of the BMTME model and the BMTME\_Approx model for the experimental maize data set

Env	BMTME			BMTME_Approx		
	$\hat{\beta}$			$\hat{\beta}$		
	GY	ASI	PH	GY	ASI	PH
EBU	6.554	2.051	2.444	6.419	1.903	2.342
KAK	5.107	1.288	2.136	4.942	1.203	2.053
KTI	6.213	2.487	2.415	6.068	2.337	2.317

Tr-ait <sup>a</sup>	$\hat{\Sigma}_t$			$\hat{\Sigma}_{t1}$		
	GY	ASI	PH	GY	ASI	PH
GY	0.801	-0.167	0.044	1.205	-0.203	0.365
ASI	-0.167	0.843	-0.031	-0.203	1.890	0.078
PH	0.044	-0.031	0.022	0.365	0.078	0.248

Env	$\hat{\Sigma}_E$			$\hat{\Sigma}_{r2}$		
	EBU	KAK	KTI	GY	ASI	PH
EBU	1.424	0.518	0.604	0.814	-0.041	0.236
KAK	0.518	2.079	1.812	-0.041	0.962	0.056
KTI	0.604	1.812	2.687	0.236	0.056	0.217

Tr-ait	$\hat{R}_e$			$\hat{R}_e$		
	GY	ASI	PH	GY	ASI	PH
GY	0.540	-0.080	0.022	0.403	-0.041	0.119
ASI	-0.080	0.526	-0.011	-0.041	0.546	0.027
PH	0.022	-0.011	0.012	0.119	0.027	0.098

	Prediction accuracy			Prediction accuracy		
	GY	ASI	PH	GY	ASI	PH
Cor	0.803	0.790	0.880	0.820	0.802	0.877
MSEP	0.435	0.419	0.009	0.405	0.402	0.010

<sup>a</sup>Traits: grain yield (GY), anthesis silking interval (ASI), and plant height (PH)

Cor and MSEP denote Pearson's correlation and mean square error of prediction, respectively, between the observed and predicted values.  $\beta$  denotes the beta coefficients,  $\Sigma_t$  denotes the genetic (co)variance matrix of traits,  $\Sigma_E$  denotes the genetic (co)variance matrix of environments,  $R_e$  denotes the residual (co)variance matrix of traits, and the symbol hat (^) denotes estimates of the corresponding parameters

**Table 6** Average Pearson's correlation (Cor) and average mean square error of prediction (MSEP) for each trait-environment combination for the experimental maize data set resulting from the testing set of the 20 random partitions

Env-Trait <sup>a</sup>	BMTME				BMTME_Approx			
	Cor	SE	MSEP	SE	Cor	SE	MSEP	SE
EBU_GY	0.327	0.019	0.787	0.019	<b>0.338</b>	0.036	<b>0.705</b>	0.026
EBU_ASI	0.508	0.016	<b>0.392</b>	0.012	0.508	0.031	0.561	0.022
EBU_PH	0.311	0.023	<b>0.014</b>	0.003	<b>0.461</b>	0.025	0.606	0.027
KAK_GY	<b>0.406</b>	0.024	<b>0.442</b>	0.021	0.351	0.019	1.061	0.057
KAK_ASI	0.397	0.015	0.937	0.044	<b>0.527</b>	0.029	<b>0.608</b>	0.028
KAK_PH	<b>0.476</b>	0.027	<b>0.011</b>	0.001	0.408	0.033	0.734	0.043
KTI_GY	0.292	0.019	<b>0.841</b>	0.023	<b>0.418</b>	0.023	0.976	0.040
KTI_ASI	0.295	0.017	<b>0.610</b>	0.018	<b>0.523</b>	0.023	0.714	0.028
KTI_PH	<b>0.502</b>	0.018	<b>0.018</b>	0.001	0.444	0.019	0.564	0.026
Average	0.390	0.020	<b>0.450</b>	0.016	<b>0.442</b>	0.026	0.726	0.033

The best predictions for each trait-environment combination are in bold

<sup>a</sup>Traits: grain yield (GY), anthesis silking interval (ASI), and plant height (PH)



**Table 7** Parameter estimates (posterior means) of the BMTME model and the BMTME\_Approx model for the experimental wheat data set

Env	BMTME				BMTME_Approx			
	$\hat{\beta}$				$\hat{\beta}$			
	DH	NDVI	GY	PH	DH	NDVI	GY	PH
Bed2IR	-3.164	0.060	-0.075	-4.496	-3.202	-0.004	-0.138	-4.647
Bed5IR	-4.026	0.053	-0.284	-7.535	-4.023	-0.011	-0.340	-7.436
Drip	-0.272	0.070	-0.342	-0.511	-0.314	0.006	-0.409	-0.692
Trait <sup>a</sup>	$\hat{\Sigma}_t$				$\hat{\Sigma}_{t1}$			
	DH	NDVI	GY	PH	DH	NDVI	GY	PH
DH	3.440	0.005	-0.098	-1.125	24.552	0.042	-0.566	-11.455
NDVI	0.005	0.000	0.000	-0.002	0.042	0.000	-0.001	-0.019
GY	-0.098	0.000	0.011	0.047	-0.566	-0.001	0.059	0.377
PH	-1.125	-0.002	0.047	1.872	-11.455	-0.019	0.377	8.532
Env	$\hat{\Sigma}_E$				$\hat{\Sigma}_{t2}$			
	Bed2IR	Bed5IR	Drip		DH	NDVI	GY	PH
Bed2IR	7.426	7.265	6.505	—	4.247	0.007	0.059	2.376
Bed5IR	7.265	7.619	6.361	—	0.007	0.000	0.000	0.004
Drip	6.505	6.361	6.388	—	0.059	0.000	0.048	0.256
	—	—	—	—	2.376	0.004	0.256	7.556
Trait	$\hat{R}_e$				$\hat{R}_e$			
	DH	NDVI	GY	PH	DH	NDVI	GY	PH
DH	4.762	0.002	0.131	1.778	6.304	0.011	0.042	2.257
NDVI	0.002	0.000	0.000	0.002	0.011	0.000	0.000	0.004
GY	0.131	0.000	0.109	0.531	0.042	0.000	0.083	0.323
PH	1.778	0.002	0.531	18.064	2.257	0.004	0.323	9.443
	Prediction accuracy				Prediction accuracy			
	DH	NDVI	GY	PH	DH	NDVI	GY	PH
Cor	0.9604	0.9953	0.7467	0.7945	0.9699	0.9748	0.9112	0.9308
MSEP	2.7257	0.0000	0.0866	14.1430	2.2085	0.0000	0.0442	6.0107

Cor and MSEP denote Pearson's correlation and mean square error of prediction, respectively, between the observed and predicted values.  $\beta$  denotes the beta coefficients,  $\Sigma_t$  denotes the genetic (co)variance matrix of traits,  $\Sigma_E$  denotes the genetic (co)variance matrix of environments,  $R_e$  denotes the residual (co)variance matrix of traits, and the symbol hat (^) denotes estimates of the corresponding parameters

<sup>a</sup>Traits: days to heading (DH), grain yield (GY), plant height (PH), and the green normalized difference vegetation index (NDVI). Each of these traits was evaluated in three environments (Bed2IR, Bed5IR, and Drip)

## HTP data set

In Table 9 we provide the parameter estimates of the HTP data set, which has 10 traits evaluated in 3 environments, with 976 lines evaluated in each environment. The estimates of the beta coefficients are reasonable since they are consistent with the sample average for each trait in each environment, which is equivalent to the least square estimates.

Rather than reporting the variance–covariance of traits, we report the correlation between traits, which can be more useful, as it gives a better idea of the level of correlation between traits in the whole data set. Our proposed BMTME\_Approx model estimates two variance–covariance matrices for the genetic part of the traits and one for the residual part of the traits; in this vein, we report two correlation matrices for the genetic part of the traits and one for the residual correlation of the traits. In the section “genetic

**Table 8** Average Pearson's correlation (Cor) and average mean square error of prediction (MSEP) for each trait-environment combination for the experimental wheat data set resulting from the testing set of the 20 random partitions

Trait-Env <sup>a</sup>	BMTME				BMTME_Approx			
	Cor	SE	MSEP	SE	Cor	SE	MSEP	SE
DH_Bed2IR	<b>0.889</b>	0.009	<b>7.450</b>	0.717	0.875	0.011	8.877	0.822
NDVI_Bed2IR	0.843	0.008	0.000	0.000	<b>0.849</b>	0.006	0.000	0.000
GY_Bed2IR	0.640	0.013	0.056	0.002	<b>0.677</b>	0.013	<b>0.051</b>	0.002
PH_Bed2IR	0.641	0.013	23.252	0.804	<b>0.693</b>	0.014	<b>20.911</b>	0.814
DH_Bed5IR	<b>0.874</b>	0.006	<b>12.446</b>	0.598	0.853	0.008	13.762	0.664
NDVI_Bed5IR	<b>0.800</b>	0.010	0.000	0.000	0.779	0.009	0.000	0.000
GY_Bed5IR	0.178	0.021	0.253	0.008	<b>0.200</b>	0.022	<b>0.238</b>	0.007
PH_Bed5IR	0.086	0.015	24.200	0.599	<b>0.170</b>	0.015	<b>18.721</b>	0.539
DH_Drip	<b>0.925</b>	0.005	4.575	0.286	0.905	0.005	<b>4.532</b>	0.188
NDVI_Drip	0.710	0.012	0.000	0.000	<b>0.711</b>	0.014	0.000	0.000
GY_Drip	0.649	0.012	0.127	0.005	<b>0.678</b>	0.014	<b>0.119</b>	0.004
PH_Drip	0.655	0.018	21.457	0.543	<b>0.669</b>	0.019	<b>20.157</b>	0.570
Average	0.657	0.012	7.818	0.297	<b>0.672</b>	0.012	<b>7.281</b>	0.301

<sup>a</sup>Traits: days to heading (DH), grain yield (GY), plant height (PH), and the green normalized difference vegetation index (NDVI). Each of these traits was evaluated in three environments (Bed2IR, Bed5IR, and Drip)

The best predictions for each trait-environment combination are in bold

correlation between traits" in Table 9, the upper diagonal part gives the correlations between traits corresponding to the term genotype  $\times$  trait, where we can observe that 33 (73.33%) of the 45 possible correlations were, in absolute values,  $>0.5$ . On the other hand, in the lower diagonal part of the genetic correlation section are given the genetic correlations that correspond to the three-way interaction term environment  $\times$  genotype  $\times$  trait. Here we can observe that only 16 (35.56%) of the 45 possible correlations are  $>0.5$ .

In general, there is evidence of a reasonably high genetic correlation between traits in this HTP data set. In Table 9, the section on the residual correlation between traits shows that only 15 (33.33%) out of the 45 possible residual correlations are  $>0.5$ . This information indicates that the phenotypic correlation between the ten traits is more influenced by the genetic part than by the residual part. Finally, the "prediction accuracy" section of Table 9 shows the Pearson's correlation and MSEP obtained for the whole data set between the observed and predicted values for each trait. In general, all the observed Pearson's correlations are  $>0.86$ , with the exception of the RNDVI, GNDVI, and SRa traits, which had a Pearson's correlation  $<0.8$ .

The prediction accuracies for the HTP data set are given in Table 10. The predictions reported were the average of the ten random partitions of the testing data set. Here we only report the prediction accuracy for the BMTME\_Approx model since it is extremely difficult in terms of implementation time to run the BMTME\_model due to the large data set. Table 10 shows that the best predictions in

terms of Pearson's correlation were observed for trait DH, while the predictions for the rest of the traits in general were low. Also, in terms of Pearson's correlation, the best predictions were observed in the irrigated environment. On the other hand, in terms of MSEP for traits GY and DH, the best predictions were observed in the drought environment; however, there were no significant differences in terms of prediction accuracy between the remaining traits, since in all traits, the MSE was almost zero.

### Large EYT data set

In Table 11, we provide the parameter estimates of the EYT data set that has 3 traits evaluated in 5 environments, with 2505 lines evaluated in each environment. We found that the estimates of the beta coefficients are reasonable since they are consistent with the least square estimates.

Here we also report the correlation between traits to have a better idea of the level of correlation between traits in the whole data set. Furthermore, we report two matrices of correlation for the genetic part of traits and one for the residual correlation of traits. In the section "genetic correlation between traits" in Table 11, the upper diagonal part gives the correlations between traits corresponding to the term genotype  $\times$  trait, where we can observe that the correlation between PH and DH was the only one  $>0.5$ . On the other hand, in the lower diagonal part of the genetic correlation section, between traits are given the genetic correlations that correspond to the three-way interaction term

**Table 9** Parameter estimates (posterior means of beta coefficients, genetic and residual correlation) of the BMTME\_Approx model for the experimental HTP data set

Env	$\hat{\beta}$									
	GY	DH	RNDVI	GNDVI	SRa	RARSa	RARSb	RARSc	NPQI	PR
Drought	2.172	77.408	-0.071	-0.085	-0.092	-0.163	-0.184	-0.200	-0.140	-0.198
Irrigated	6.522	85.728	-0.093	-0.085	-0.083	-0.106	-0.125	-0.132	-0.106	-0.170
Red_Irrig	3.735	81.946	-0.084	-0.088	-0.092	-0.136	-0.163	-0.186	-0.127	-0.195

Genetic correlation between traits (upper diagonal corresponds to $\hat{\Sigma}_{t1}$ , while lower diagonal to $\hat{\Sigma}_{t2}$ )										
Trait <sup>a</sup>	GY	DH	RNDVI	GNDVI	SRa	RARSa	RARSb	RARSc	NPQI	PR
GY	1.000	<b>0.781</b>	<b>-0.612</b>	-0.421	-0.340	-0.251	-0.275	-0.265	-0.302	-0.437
DH	0.154	1.000	<b>-0.708</b>	<b>-0.576</b>	-0.495	-0.467	-0.489	-0.483	-0.499	<b>-0.658</b>
RNDVI	-0.271	-0.248	1.000	<b>0.860</b>	<b>0.827</b>	<b>0.776</b>	<b>0.767</b>	<b>0.707</b>	<b>0.811</b>	<b>0.799</b>
GNDVI	0.123	-0.260	0.276	1.000	<b>0.972</b>	<b>0.934</b>	<b>0.929</b>	<b>0.886</b>	<b>0.949</b>	<b>0.914</b>
SRa	0.256	-0.263	0.429	0.463	1.000	<b>0.944</b>	<b>0.944</b>	<b>0.898</b>	<b>0.970</b>	<b>0.899</b>
RARSa	<b>0.605</b>	-0.231	-0.170	0.493	0.441	1.000	<b>0.992</b>	<b>0.971</b>	<b>0.983</b>	<b>0.949</b>
RARSb	<b>0.667</b>	-0.284	-0.040	0.350	<b>0.547</b>	<b>0.817</b>	1.000	<b>0.985</b>	<b>0.988</b>	<b>0.967</b>
RARSc	<b>0.648</b>	-0.262	-0.050	0.260	0.417	<b>0.681</b>	<b>0.846</b>	1.000	<b>0.957</b>	<b>0.970</b>
NPQI	<b>0.512</b>	-0.274	0.196	0.260	<b>0.659</b>	<b>0.578</b>	<b>0.825</b>	<b>0.658</b>	1.000	<b>0.949</b>
PR	0.485	-0.402	0.070	0.352	0.411	<b>0.595</b>	<b>0.786</b>	<b>0.830</b>	<b>0.629</b>	1.000

Residual correlation between traits										
Trait	GY	DH	RNDVI	GNDVI	SRa	RARSa	RARSb	RARSc	NPQI	PR
GY	1.000	0.164	-0.210	0.086	0.187	<b>0.508</b>	<b>0.548</b>	<b>0.539</b>	0.402	0.382
DH	—	1.000	-0.199	-0.201	-0.207	-0.206	-0.248	-0.231	-0.229	-0.339
RNDVI	—	—	1.000	0.281	0.361	-0.051	0.026	0.023	0.184	0.108
GNDVI	—	—	—	1.000	0.406	0.430	0.315	0.264	0.226	0.348
SRa	—	—	—	—	1.000	0.425	<b>0.500</b>	0.391	<b>0.572</b>	0.354
RARSa	—	—	—	—	—	1.000	<b>0.765</b>	<b>0.642</b>	<b>0.548</b>	<b>0.546</b>
Env										
RARSb	—	—	—	—	—	—	1.000	<b>0.777</b>	<b>0.765</b>	<b>0.707</b>
RARSc	—	—	—	—	—	—	—	1.000	<b>0.581</b>	<b>0.749</b>
NPQI	—	—	—	—	—	—	—	—	1.000	<b>0.541</b>
PR	—	—	—	—	—	—	—	—	—	1.000

Prediction accuracy										
	GY	DH	RNDVI	GNDVI	SRa	RARSa	RARSb	RARSc	NPQI	PR
Cor	0.970	0.958	0.797	0.612	0.709	0.928	0.958	0.961	0.867	0.872
MSEP	0.206	2.108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

<sup>a</sup>Traits are: grain yield (GY), days to heading (DH), red normalized difference vegetation index (RNDVI), green normalized difference vegetation index (GNDVI), simple ratio (SRa), ratio analysis of reflectance spectra chlorophyll a (RARSa), ratio analysis of reflectance spectra chlorophyll b (RARSb), ratio analysis of reflectance spectra chlorophyll c (RARSc), normalized pheophytinization index (NPQI), and photochemical reflectance index (PR)

Cor and MSEP denote Pearson’s correlation and mean square error of prediction, respectively, between the observed and predicted values. Environments are drought, irrigated, and reduced irrigation (Red\_Irrig). In bold are the correlations >0.5.  $\beta$  denotes the beta coefficients,  $\Sigma_t$  denotes the genetic (co)variance matrix of traits,  $\Sigma_E$  denotes the genetic (co)variance matrix of environments,  $R_e$  denotes the residual (co)variance matrix of traits, and the symbol hat (^) denotes estimates of the corresponding parameters

**Table 10** Average Pearson's correlation (Cor) and average mean square error of prediction (MSEP) for each trait–environment combination for the experimental HTP data set resulting from the testing set of the 10 random partitions

Trait <sup>a</sup>	Drought		Irrigated		Reduced Irrigated	
	Cor	SE	Cor	SE	Cor	SE
GY	0.080	0.009	−0.054	0.007	<b>0.152</b>	0.006
DH	<b>0.812</b>	0.005	0.734	0.005	0.771	0.004
RNDVI	0.085	0.009	<b>0.160</b>	0.008	0.027	0.005
GNDVI	0.150	0.012	<b>0.290</b>	0.012	0.047	0.011
SRa	0.179	0.007	<b>0.276</b>	0.011	0.202	0.012
RARSa	0.167	0.007	<b>0.304</b>	0.015	0.150	0.010
RARSb	0.224	0.012	<b>0.326</b>	0.017	0.238	0.015
RARSc	0.195	0.015	<b>0.391</b>	0.013	0.078	0.009
NPQI	0.135	0.010	<b>0.237</b>	0.007	0.141	0.011
PR	−0.062	0.010	<b>0.342</b>	0.020	0.046	0.007

Trait	Drought		Irrigated		Reduced Irrigated	
	MSEP	SE	MSEP	SE	MSEP	SE
GY	<b>0.298</b>	0.005	0.464	0.005	0.143	0.002
DH	<b>3.695</b>	0.169	11.726	0.163	4.941	0.072
RNDVI	0.000	0.000	0.000	0.000	0.000	0.000
GNDVI	0.000	0.000	0.000	0.000	0.000	0.000
SRa	0.000	0.000	0.000	0.000	0.000	0.000
RARSa	0.000	0.000	0.000	0.000	0.000	0.000
RARSb	0.000	0.000	0.000	0.000	0.000	0.000
RARSc	0.000	0.000	0.000	0.000	0.000	0.000
NPQI	0.000	0.000	0.000	0.000	0.000	0.000
PR	0.000	0.000	0.000	0.000	0.000	0.000

The best predictions for each trait in the three environments are in bold

<sup>a</sup>Traits are: grain yield (GY), days to heading (DH), red normalized difference vegetation index (RNDVI), green normalized difference vegetation index (GNDVI), simple ratio (SRa), ratio analysis of reflectance spectra chlorophyll a (RARSa), ratio analysis of reflectance spectra chlorophyll b (RARSb), ratio analysis of reflectance spectra chlorophyll c (RARSc), normalized pheophytinization index (NPQI), and photochemical reflectance index (PR)

environment  $\times$  genotype  $\times$  trait. Here we can observe that only the correlation between GY and PH was  $>0.5$  (Table 11).

In Table 11, in the section on the residual correlation between traits, we can observe that only the correlation between GY and PH was  $>0.5$ . Finally, in the “prediction accuracy” section of Table 11, we can observe the Pearson's correlation and MSEP obtained for the whole data set between the observed and predicted values for each trait. In general, all the observed Pearson's correlations are  $>0.89$ .

The prediction accuracies for this large wheat data set are given in Table 12. The predictions reported are the average of the ten random partitions of the testing data set. Here we only report the prediction accuracy for the BMTME\_Aprox model; owing to the size of the data set, it almost impossible to run the BMTME\_model. Table 12 shows that

the best predictions in terms of Pearson's correlation were for trait GY, while the worst were for trait DH. At the same time, the best predictions for GY, DH, and PH in terms of Pearson's correlation were observed in environments BED\_5IR, FLAT\_5IR, and BED\_2IR, respectively. On the other hand, in terms of MSEP for traits GY, DH, and PH, the best predictions were observed in BED\_2IR, LHT, and FLAT\_5IR, respectively (Table 12).

## Discussion

The amounts of data that the breeding programs around the world are generating continue to increase; consequently, there is a growing need to extract more knowledge from the data being produced. To this end, multiple-trait models are

**Table 11** Parameter estimates (posterior means of beta coefficients, genetic and residual correlation) of the BMTME\_Approx model for the experimental large EYT data set

Tr-ait	$\hat{\beta}$					
	BED_5IR	FLAT_5IR	BED_2IR	FLAT_DRIP	LHT	—
GY	6.349	6.407	3.792	2.130	3.302	—
DH	81.810	79.538	80.628	75.410	58.412	—
PH	102.856	102.766	84.986	73.347	68.810	—
Tr-ait	<sup>a</sup> Genetic correlation between traits			Residual correlation between traits		
	GY	DH	Height	GY	DH	Height
GY	1.000	-0.487	0.599	1.000	-0.393	0.580
DH	-0.086	1.000	-0.421	—	1.000	-0.178
PH	0.631	0.403	1.000	—	—	1.000
Prediction accuracy	Cor			MSEP		
	GY	DH	Height	GY	DH	Height
Prediction accuracy	0.928	0.894	0.907	0.489	19.688	46.073

Cor and MSEP denote Pearson's correlation and mean square error of prediction, respectively, between the observed and predicted values

<sup>a</sup>Genetic correlation between traits (the upper diagonal corresponds to  $\hat{\Sigma}_{11}$ , while the lower diagonal is for  $\hat{\Sigma}_{22}$ ). Traits are grain yield (GY), days to heading (DH), and plant height (PH)

**Table 12** Average Pearson's correlation (Cor) and average mean square error of prediction (MSEP) for each trait–environment combination for the experimental large EYT data set resulting from the testing set of the 10 random partitions

Env	GY		DH		PH	
	Cor	SE	Cor	SE	Cor	SE
BED_5IR	<b>0.417</b>	0.012	0.292	0.014	0.110	0.013
FLAT_5IR	0.274	0.012	<b>0.330</b>	0.017	0.276	0.009
BED_2IR	0.247	0.015	0.230	0.015	<b>0.408</b>	0.014
FLAT_DRIP	0.259	0.013	0.259	0.012	0.371	0.017
LHT	0.341	0.011	0.278	0.006	0.213	0.009
Average	0.308	0.013	0.278	0.013	0.276	0.012
Env	MSEP	SE	MSEP	SE	MSEP	SE
BED_5IR	0.528	0.008	31.793	0.513	24.620	0.321
FLAT_5IR	0.514	0.008	18.144	0.272	<b>23.882</b>	0.435
BED_2IR	<b>0.390</b>	0.004	19.859	0.448	63.251	1.168
FLAT_DRIP	0.545	0.010	19.800	0.327	102.330	1.217
LHT Env	0.654	0.010	<b>15.662</b>	0.363	33.657	0.416
Average	0.526	0.008	21.051	0.385	49.548	0.712

The best predictions for each trait in the five environments are in bold; the comparisons are made by column (trait). Traits are grain yield (GY), days to heading (DH), and plant height (PH)

commonly used to take advantage of correlated traits to improve parameter estimation and prediction accuracy. However, when there is a large number of traits, implementing these types of models is challenging. Therefore, it is necessary to develop efficient multiple-trait and multiple-environment models for whole-genome selection in order to

take advantage of multiple correlated traits. In this paper, we propose an alternative method for analyzing multi-trait data that could be useful for whole-genome selection in the context of an abundance of traits. Some advantages of the proposed method are: (i) it can be implemented in current genomic selection software that was built for univariate



analysis (e.g., BGLR, ASREML, package Sommer of R); (ii) it can be implemented with a large number of traits and, in general, with a large data set since this model is implemented in four steps; and (iii) this method can be more efficient in terms of implementation time because univariate analyses for each trait (step 2 of the procedure) are required for its implementation, which allows an implementation in parallel and with low dimensions compared to the BMTME model that models all the traits simultaneously and also (iv) decreases the probability of having collinearity and convergence problems, since when more correlated traits are added to the multivariate analysis, both problems increase (Schulthess et al. 2016).

In Supplementary material, we provide the R code for implementing the proposed method. When the matrix of response variable ( $\mathbf{Y}$ ) shows a considerable departure from normality, the proposed BMTME\_Approx model is expected to be inefficient due to the fact that uncorrelated traits do not imply independence for non-normal data. For this reason, under these circumstances, in Supplementary material we also provide a solution to this situation based on the ICA that transforms the original unnecessary Gaussian matrix of response variables ( $\mathbf{Y}$ ) into a matrix of independent variables ( $\mathbf{Y}^*$ ).

Also, it is important to point out that the BMTME\_Thompson version is more efficient computationally than the original BMTME model since it allows sampling independently for each trait the random effects of  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , which improves computational efficiency because it avoids sampling from huge multivariate normal distributions. More specifically, the dimensions of the full conditionals are  $L$  times smaller than the original BMTME model. Another advantage is that the BMTME\_Thompson model is not an approximation to the BMTME model, since it is only a reparametrization of the original model, which allows obtaining exactly the same parameter estimates and prediction accuracy at a lower cost in terms of implementation time, since the sampling process of the full conditionals of the random effects of  $\mathbf{b}_1$  and  $\mathbf{b}_2$  is done individually for each trait.

The proposed BMTME\_Approx model takes advantage of the fact that optimal point estimates of any linear combination of the means and variances of the various separate analyses for each response variable can be obtained. Therefore, our proposed method uses a linear transformation of the separate parameter estimates to provide reasonable estimates of the beta coefficients ( $\boldsymbol{\beta}$ ), random effects ( $\mathbf{b}_1$  and  $\mathbf{b}_2$ ), and variance–covariance matrices ( $\boldsymbol{\Sigma}_{t1}$ ,  $\boldsymbol{\Sigma}_{t2}$ , and  $\mathbf{R}_e$ ) for a multivariate model. However, although the proposed model is able to provide reasonable approximate parameter estimates for the multivariate model by doing a separate analysis for each trait, it is unable to provide estimates of

the standard error for off diagonal elements in the variance–covariance estimates.

Based on the results of the simulated and real data sets, we have reason to argue that the proposed BMTME\_Approx model produces competitive predictions compared to those produced by the BMTME model, even though the parameter estimates resulting from the proposed BMTME\_Approx model are quite different from those resulting from the BMTME model (mainly in the wheat real data set). However, the differences in parameter estimates between the BMTME and the BMTME\_Approx models can be attributed to the fact that the BMTME\_Approx model estimates two genetic variance–covariance matrices for traits (one for the interaction term genotype  $\times$  trait and the other for the three-way interaction term environment  $\times$  genotype  $\times$  trait), while the BMTME only estimates one variance–covariance for both interaction terms. Also, in terms of prediction accuracies for simulated data set 1 (simulated with environments and correlated traits), the BMTME was better than the BMTME\_Approx model, but in the second simulated data set, we observed that the BMTME\_Approx (that assumes independence between environments and correlated traits) was better than the BMTME (that assumes correlated traits between environments and traits), which can be attributed to the fact that this second data set was simulated assuming independence between environments and correlated traits.

It is also important to point out that the proposed BMTME\_Approx model can be implemented when there are many traits. For example, in the HTP and Large EYT data sets where there are a large number of traits, the application of the BMTME model becomes almost impossible due to the fact that samples are extracted from a very large number of multivariate normal distributions, and the modeling process is performed jointly for all the traits and not separately for each trait, as in the proposed BMTME\_Approx model. This is a key element for achieving an efficient estimation process in terms of implementation time. Additionally, the proposed BMTME\_Approx model can be implemented simultaneously since the procedure for estimating the required parameters requires a separate analysis for each trait.

Another point that we would like to highlight is that our proposed model is multiple-trait and multiple-environment but with the restriction that an identity matrix is assumed for the variance–covariance matrix of environments. However, even with this restrictive assumption in the variance–covariance matrix of environments, the model has the advantage of taking into account the interaction terms environment  $\times$  trait, genotype  $\times$  trait, and the three-way interaction environment  $\times$  genotype  $\times$  trait. Furthermore, it takes into account the correlated traits and can be

implemented using conventional software for whole-genome prediction.

## Conclusions

The results of the simulated and real data sets show that the proposed alternative method produced results that are similar to those of the conventional multiple-trait analysis. For this reason, the proposed method is an attractive alternative for analyzing multiple-trait data in the context of a large number of traits. However, it is important to point out that the significant differences found in parameter estimates between the proposed BMTME\_Approx model and the BMTME model can be attributed mainly to the fact that the BMTME\_Approx model allows the estimation of two genetic variance–covariance matrices for traits, one for the interaction term genotype  $\times$  trait, and the other for the term environment  $\times$  genotype  $\times$  trait, while the BMTME model only estimates one genetic matrix of variance–covariance for traits.

**Acknowledgements** We thank all CIMMYT scientists who were involved in conducting the extensive multi-environment maize and wheat trials and generated the phenotypic and genotypic data used in this study. We acknowledge the financial support provided by USAID and the Cornell-CIMMYT Genomic Selection projects in maize and wheat breeding financed by the Bill and Melinda Gates Foundation.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bhat JA, Ali S, Salgotra RK, Mir ZA, Dutta S, Jadon V, Prabhu KV (2016) Genomic selection in the era of next generation sequencing for complex traits in plant breeding. *Front Genet* 7:221. <https://doi.org/10.3389/fgene.2016.00221>
- Boichard D, Ducrocq V, Croiseau P, Fritz S (2016) Genomic selection in domestic animals: principles, applications and perspectives. *C R Biol* 339:274–277
- Calus MP, Veerkamp RF (2011) Accuracy of multi-trait genomic selection using different methods. *Genet Sel Evol* 43:26. <https://doi.org/10.1186/1297-9686-43-26>
- Chiquet J, Tristan M, Robin S (2013). Multi-trait genomic selection via multivariate regression with structured regularization. MLCB NIPS 2013, Oct 2013, South Lake Tahoe, United States.
- de los Campos G, Pérez-Rodríguez P (2014). Bayesian Generalized Linear Regression. R package version 1.0.4. <http://CRAN.R-project.org/package=BGLR>.
- Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci* 19:592–601. <https://doi.org/10.1016/j.tplants.2014.05.006>.
- Ducrocq V, Besbes B (1993) Solution of multiple trait animal models with missing data on some traits. *J Anim Breed Genet.* 110:81–92
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. <https://doi.org/10.1371/journal.pone.0019379>
- He D, Kuhn D, Parida L (2016) Novel applications of multitask learning and multiple output regression to multiple genetic trait prediction. *Bioinformatics* 32:i37–i43. <https://doi.org/10.1093/bioinformatics/btw249>
- Hyvärinen A, Karhunen J, Oja E (2001) Independent component analysis, 1st edn. John Wiley & Sons, New York, ISBN 0-471-22131-7
- Jia Y, Jannink J-L (2012) Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* 192:1513–1522. <https://doi.org/10.1534/genetics.112.144246>
- Jiang J, Zhang Q, Ma L, Li J, Wang Z, Liu JF (2015) Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity* 115:29–36
- Lin CY, Smith SP (1990) Transformation of multitrait to mixed model analysis of data with multiple random effects. *J Dairy Sci* 73:2494–2502
- Märtens K, Hallin J, Warringer J, Liti G, Parts L (2016) Predicting quantitative traits from genome and phenome with near perfect accuracy. *Nat Commun* 7:11512. <https://doi.org/10.1038/ncomms11512>
- Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829
- Montesinos-López OA, Montesinos-López A, Crossa J, Toledo F, Pérez-Hernández O, Eskridge KM, Rutkoski J (2016) A genomic Bayesian multi-trait and multi-environment model. *G3* 6:2725–2744
- Montesinos-López OA, Montesinos-López A, Crossa J, Toledo FH, Montesinos-López JC, Singh P, Salinas-Ruiz J (2017a) A Bayesian Poisson-lognormal model for count data for multiple-trait multiple-environment genomic-enabled prediction. *G3* 7:1595–1606. <https://doi.org/10.1534/g3.117.039974>
- Montesinos-López A, Montesinos-López OA, Cuevas J, Mata-López WA, Burgueño J, Mondal S, Huerta J, Singh R, Autrique E, González-Pérez L, Crossa J (2017b) Genomic Bayesian functional regression models with interactions for predicting wheat grain yield using hyper-spectral image data. *Plant Methods* 13:1–29
- Montesinos-López OA, Montesinos-López A, Crossa J, de los Campos G, Alvarado G, Mondal S, Rutkoski J, González-Pérez L (2017c) Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods* 13:1–23
- Mrode RA (2014) Linear models for the prediction of animal breeding values. CABI, Boston, MA
- Oliveira R, Teixeira-Pinto A (2015) Analyzing multiple outcomes: is it really worth the use of multivariate linear regression? *J Biom Biostat* 6:256. <https://doi.org/10.4172/2155-6180.1000256>

- Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach PLoS ONE 7: e32253. <https://doi.org/10.1371/journal.pone.0032253>
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, ISBN 3-900051-07-0. <http://www.R-project.org/>.
- Rutkoski JE, Heffner EL, Sorrells ME (2011) Genomic selection for durable stem rust resistance in wheat. *Euphytica* 179:161–173. <https://doi.org/10.1007/s10681-010-0301-1>
- Rutkoski J, Poland J, Mondal S, Autrique E, González-Pérez L, Crossa J, Reynolds M, Singh R (2016) Predictor traits from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3* 6:2799–2808
- Schaeffer LR (2006) Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet* 123:218–223
- Schulthess AW, Wang Y, Miedaner T, Wilde P, Reif JC, Zhao Y (2016) Multiple-trait and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theor Appl Genet* 129:273–287
- Schulthess AW, Zhao Y, Longin CFH, Reif JC (2017). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (*Triticum aestivum* L.). *Theor Appl Genet*. <https://doi.org/10.1007/s00122-017-3029-7>.
- Srivastava MS, Khatri CG (1979) An introduction to multivariate statistics. Elsevier North Holland, New York, NY, USA
- Stone JV (2004) Independent component analysis: a tutorial introduction. MIT Press, Cambridge, Massachusetts, ISBN 0-262-69315-1
- Teixeira FRF, Nascimento M, Nascimento ACC, Silva FF, Cruz CD, Azevedo CF, Paixão DM, Barroso LMA, Verardo LL, Resende MDV, Guimarães SEF, Lopes PS (2016) Factor analysis applied to genome prediction for high-dimensional phenotypes in pigs. *Genet Mol Res* 15:gm.15028231
- Thompson R (1977) Estimation of quantitative genetic parameters. In: Pollak E, Kempthorne O, Bailey TB (eds) Proceedings of the International Conference on Quantitative Genetics. Iowa State University Press, Ames, IO, pp 639–657
- Van Raden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* 91:4414–4423