ORIGINAL ARTICLE

Projecting results of zoned multi-environment trials to new locations using environmental covariates with random coefficient models: accuracy and precision

Harimurti Buntaran[1] · Johannes Forkman[2] · Hans-Peter Piepho[1]

Abstract

Key message **We propose the utilisation of environmental covariates in random coefficient models to predict the genotype performances in new locations.**

Abstract Multi-environment trials (MET) are conducted to assess the performance of a set of genotypes in a target population of environments. From a grower's perspective, MET results must provide high accuracy and precision for predictions of genotype performance in new locations, i.e. the grower's locations, which hardly ever coincide with the locations at which the trials were conducted. Linear mixed modelling can provide predictions for new locations. Moreover, the precision of the predictions is of primary concern and should be assessed. Besides, the precision can be improved when auxiliary information is available to characterize the targeted locations. Thus, in this study, we demonstrate the benefit of using environmental information (covariates) for predicting genotype performance in some new locations for Swedish winter wheat official trials. Swedish MET locations can be stratified into zones, allowing borrowing information between zones when best linear unbiased prediction (BLUP) is used. To account for correlations between zones, as well as for intercepts and slopes for the regression on covariates, we fitted random coefficient (RC) models. The results showed that the RC model with appropriate covariate scaling and model for covariate terms improved the precision of predictions of genotypic performance for new locations. The prediction accuracy of the RC model was competitive compared to the model without covariates. The RC model reduced the standard errors of predictions for individual genotypes and standard errors of predictions of genotype differences in new locations by 30–38% and 12–40%, respectively.

Introduction

The main goal of a plant breeding programme is to develop well-adapted genotypes in a target population of environments (TPE). Multi-environment trials (MET) are conducted to evaluate candidate genotypes in the TPE, and to understand and exploit the pattern of genotype × environment interactions (GEI) in the TPE. GEI is the differential response of genotypes across different environments (Kang and Gorman 1989). GEI in a TPE can be exploited to make more targeted predictions and recommendations on cultivars. This is of particular interest when there is crossover interaction, which poses a challenge when selecting genotypes for broad adaptation.

Identification of environmental covariates that are responsible for GEI is useful to enhance the predictive capability of MET analyses (Heslot et al. 2014) and evaluate the adaptability of the genotypes to the new target environment. The most commonly used types of environmental covariates are soil and meteorological covariates (van Eeuwijk et al. 2016). Incorporating environmental covariates in the GEI analysis has been done by factorial regression (Denis 1988; Piepho et al. 1998; van Eeuwijk and Elgersma 1993). Furthermore, environmental covariates have been used in a linear mixed model framework, such as in quantitative trait loci (QTL)

✉ Hans-Peter Piepho
  hans-peter.piepho@uni-hohenheim.de

  Harimurti Buntaran
  harimurti.buntaran@uni-hohenheim.de

  Johannes Forkman
  johannes.forkman@slu.se

[1] Biostatistics Unit, Institute of Crop Science, University of Hohenheim, Fruwirthstraße 23, 70599 Stuttgart, Germany

[2] Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Box 7043, 750 07 Uppsala, Sweden

biparental mapping to dissect the response of marker effects to the environmental covariates known as QTL × environment interactions (Boer et al. 2007; Crossa et al. 1999; Malosetti et al. 2004) or as eco-physiological QTL (van Eeuwijk et al. 2005). Environmental covariates have also been introduced in genomic selection (Gillberg et al. 2019; Heslot et al. 2014; van Eeuwijk et al. 2019).

The regression on environmental covariates is usually modelled by fixed effects. This type of modelling is appropriate when only studying the pattern of GEI at the tested locations. Such models are also appropriate for making predictions in an unstructured TPE. However, when the TPE is sub-divided into zones, it is necessary to model genotypic effects as random in order to borrow strength between zones (Buntaran et al. 2019, 2020). If such modelling is coupled with factorial regression approaches, genotype-specific regression coefficients must be modelled as random effects as well, giving rise to what are known as random coefficient (RC) models (Longford 1993; Milliken and Johnson 2002).

Although MET are usually designed to cover the whole TPE, none of the trials in an MET coincides exactly with a grower's field or location. Thus, grower's fields, the real target of breeding, must be seen as new locations in the TPE. In the same vein, it may be said that the MET analysis is usually used to produce predictions of tested genotypes for a new location, making use of the information from tested locations. Predicting genotype performance in a new location is akin to predicting values that have no records at all. Henderson (1977) showed that best linear unbiased prediction (BLUP) can be used to predict breeding values for the animals that had no records. BLUP, therefore, can also be used to obtain genotype predictions in a new location that has no records.

Since growers' fields hardly ever coincide with the trials' location and, in practice, cultivar yield will never reach the exact same value as the predicted mean values from the MET, reporting the precision or precision measures as quantified by standard errors and prediction intervals is highly desirable. Without this, growers are left with having no information regarding the precision in the predictions that are reported. The key challenge is that the standard errors of predictions of variety means obtained from the routine analysis of MET are only valid for the locations where the trials were carried out, but not for the untested locations or growers' field. However, the precision of the predictions for the untested locations is crucial in order to assist growers in selecting a cultivar for their farm or field.

Accuracy refers to how close the value of the statistic is to a supposed ''true value''. In other words, accuracy measures how close an estimate $\widehat{\theta}$ of a parameter θ is to the ''true value'' of θ (Kotz et al. 2006). Accuracy can be evaluated via a cross-validation (CV) study by estimating prediction error (Hastie et al. 2009). From a CV study, accuracy can be measured in terms of mean squared error (MSE), which consists of variance and squared bias of $\widehat{\theta}$. Conversely, the precision of an estimator $\widehat{\theta}$, measures how narrow the distribution of $\widehat{\theta}$ clusters about its expected value (Kotz et al. 2006). The precision of $\widehat{\theta}$ is the reciprocal of the variance of $\widehat{\theta}$.

In this study, we explore the use of the RC models for improving the precision of yield predictions in some new locations, which represent growers' fields, and evaluate the prediction accuracy. The precision of predictions is assessed with standard errors of predictions of genotypic values (SEPV) and standard errors of the predictions of pairwise differences of genotypic values (SEPD). The prediction accuracy was evaluated via a CV study.

Materials and methods

Dataset

In this study, a dataset from Swedish official cultivar testing in 2016 was used. The dataset comprises 25 genotypes of winter wheat tested in 18 locations. In addition, the dataset includes information about four new locations. The 22 locations are stratified into three zones: South, Middle, and North (Buntaran et al. 2019). Two of the four new locations are located in the North zone, and two in the South zone. The new locations have no observations of yield for any of the genotypes but do have data on environmental covariates. The layout of the trials at each location was an α-design with two replicates. The available covariates were soil properties, i.e. pH, clay content, and humus content. The covariates are location-specific.

Statistical models

A two-stage fully-efficient stage-wise approach was used, which forwards the full variance–covariance matrix of adjusted means from stage I (Damesa et al. 2017; Piepho et al. 2012) to the final analysis in stage II of a stage-wise analysis. In stage I, each location was analysed individually using a linear mixed model (LMM). The general equation of an LMM is written as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{1}$$

where $\mathbf{Y}$ is the vector of yield is the design matrix for fixed-effects of vector $\boldsymbol{\beta}$, $\mathbf{Z}$ is the incidence matrix for random-effects of vector $\mathbf{u}$, and $\mathbf{e}$ is the vector residuals. The distributions of the random effects and the residuals are $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$ and $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$. Then, the distribution of the response is $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V})$, where $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$.

In our case, in stage I, the fixed effects $\mathbf{X}\boldsymbol{\beta}$ for the jm-th trial are

$$\mu_{ijm} = \phi_{jm} + \tau_{ijm}$$

where the ϕ_{jm} is the intercept of the j-th location nested in the m-th zone, and τ_{ijm} is the effects of i-th genotype in the j-th location nested in the m-th zone, for $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J_m$, and $m = 1, 2, \dots, M$. The letter I denotes the number of genotypes, M is the number of zones, J_m is the number of locations within the m-th zone, and $J = \sum_{m=1}^{M} J_m$ is total number of locations. The random effects $\mathbf{Zu}$ and $\mathbf{e}$ are

$$r_{jkm} + b_{jklm} + e_{ijklm}$$

where $r_{jkm} \sim N(0, \sigma_r^2)$ is the random effect of the k-th replication in the j-th location nested in the m-th zone, $b_{jklm} \sim N(0, \sigma_b^2)$ is the random effect of the s-th incomplete block of the k-th replication in the j-th location nested in the m-th zone, and e_{ijklm} is the residual plot error associated with the observation y_{ijklm} in vector $\mathbf{Y}$. In a single-stage analysis, the replication effect is random with necessity because it is nested within location, which is a random factor. Thus, to mimic this approach in the fully efficient two-stage analysis, the replication effect is assigned to be random.

From stage I analysis, we obtained the adjusted means at the j-th location in the m-th zone, $\hat{\mu}_{ijm}$, which are estimated by best linear unbiased estimation (BLUE), and e_{ijm}, the error associated with the adjusted means. The adjusted means and the associated error were used for stage II analysis.

In stage II, the analysis was conducted using a model that comprises zone effects since the locations are stratified into three zones. For the stage II model, the vector $\mathbf{Y}$ consists of the means $\hat{\mu}_{ijm}$ from stage I and the error vector $\mathbf{e}$ has sub-vectors $\mathbf{e}_j$ with elements e_{ijm}, with $\mathrm{var}(\mathbf{e}_j) = \mathbf{R}_j$. The variance–covariance matrix $\mathbf{R}_j$ of adjusted means in the j-th location is obtained in stage I by using residual maximum likelihood (REML). The overall variance–covariance matrix of $\mathbf{e}$ is block diagonal, i.e.

$$\mathrm{var}(\mathbf{e}) = \mathbf{R} = \oplus_{j=1}^{J} \mathbf{R}_j = \begin{pmatrix} \mathbf{R}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{R}_2 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{R}_J \end{pmatrix}$$

In stage II, seven fixed-genotype-effect (FG) and seven random-genotype-effect (RG) statistical models were fitted. The entries in vectors $\mathbf{X\beta}$ and $\mathbf{Zu}$ of the FG and RG models are presented in Tables 1, 2, respectively. In stage II, for all 14 models, the location effect nested in the m-th zone is random, whereas the effect of the zone, ζ_m, is fixed. In the RG models, the genotype × zone interaction effect, $(g\zeta)_{im}$, is random because genotype is a random factor. The random genotype main effect and genotype × zone interaction effect allow exploiting information across zones when computing

Table 1 The seven fixed-genotype-effect (FG) models

Model	$\mathbf{X\beta}$	$\mathbf{Zu}$*	Covariance structure
FG1	$\alpha + \zeta_m + \tau_i + (\tau\zeta)_{im}$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^{M} \mathbf{G}_{l_m}, \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$
			$\mathbf{G}_{\tau l} = \oplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}, \mathbf{G}_{(\tau l)_m} = \sigma_{(\tau l)_m}^2 \mathbf{I}$
FG2	$\alpha + \zeta_m + \tau_i + (\tau\zeta)_{im}$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \sigma_l^2 \mathbf{I}$
			$\mathbf{G}_{\tau l} = \sigma_{\tau l}^2 \mathbf{I}$
FGC	$\alpha + \beta x_{jm} + \zeta_m + \tau_i + (\tau\zeta)_{im} + \lambda_m x_{jm}$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^{M} \mathbf{G}_{l_m}, \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$
			$\mathbf{G}_{\tau l} = \oplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}, \mathbf{G}_{(\tau l)_m} = \sigma_{(\tau l)_m}^2 \mathbf{I}$
FGCQ	$\alpha + \beta x_{jm} + \gamma x_{jm}^2 + \zeta_m + \tau_i + (\tau\zeta)_{im} + \lambda_m x_{jm}$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^{M} \mathbf{G}_{l_m}, \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$
			$\mathbf{G}_{\tau l} = \oplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}, \mathbf{G}_{(\tau l)_m} = \sigma_{(\tau l)_m}^2 \mathbf{I}$
FGI1	$\alpha + \beta_1 x_{jm} + \gamma x_{jm}^2 +$ $\zeta_m + \lambda_m x_{jm} +$ $\left(\tau_i + \beta_{2i} x_{jm} + \beta_{3i} x_{jm}^2\right) +$ $\left((\tau\zeta)_{im} + \beta_{4_{im}} x_{jm} + \beta_{5_{im}} x_{jm}^2\right)$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^{M} \mathbf{G}_{l_m}, \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$ $\mathbf{G}_{\tau l} = \oplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}, \mathbf{G}_{(\tau l)_m} = \sigma_{(\tau l)_m}^2 \mathbf{I}$
FGI2	$\alpha + \beta x_{jm} + \gamma x_{jm}^2 + \zeta_m + \tau_i + \lambda_m x_{jm}$ $+ \left((\tau\zeta)_{im} + \beta_{4_{im}} x_{jm} + \beta_{5_{im}} x_{jm}^2\right)$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^{M} \mathbf{G}_{l_m}, \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$ $\mathbf{G}_{\tau l} = \oplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}, \mathbf{G}_{(\tau l)_m} = \sigma_{(\tau l)_m}^2 \mathbf{I}$
FGI3	$\alpha + \beta x_{jm} +$ $\gamma x_{jm}^2 + \zeta_m + (\tau\zeta)_{im} + \lambda_m x_{jm}$ $+ \left(\tau_i + \beta_{2i} x_{jm} + \beta_{3i} x_{jm}^2\right)$	$l_{jm} + (\tau l)_{ijm}$	$\mathbf{G}_l = \oplus_{(m=1)}^{M} \mathbf{G}_{(l_m)}, \mathbf{G}_{(l_m)} = \sigma_{(l_m)}^2 \mathbf{I}$ $\mathbf{G}_{\tau l} = \oplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}, \mathbf{G}_{(\tau l)_m} = \sigma_{(\tau l)_m}^2 \mathbf{I}$

*The index i refers to the i-th genotype, j refers to the j-th location, and m refers to the m-th zone

Table 2 The seven random-genotype effect (RG) models

Model	$\mathbf{X\beta}$	$\mathbf{Zu}$*	Covariance structure
RG1	$\alpha + \zeta_m$	$l_{jm} + g_i + (g\zeta)_{im} + (gl)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^M \mathbf{G}_{l_m},\ \mathbf{G}_{l_m} = \sigma_{l_m}^2\, \mathbf{I}$ $\mathbf{G}_g = \sigma_g^2 \mathbf{I}$ $\mathbf{G}_{g\zeta} = \sigma_{g\zeta}^2 \mathbf{I}$ $\mathbf{G}_{gl} = \oplus_{m=1}^M \mathbf{G}_{(gl)_m},\ \mathbf{G}_{(gl)_m} = \sigma_{(gl)_m}^2 \mathbf{I}$
RG2	$\alpha + \zeta_m$	$l_{jm} + g_i + (g\zeta)_{im} + (gl)_{ijm}$	$\mathbf{G}_l = \sigma_l^2 \mathbf{I}$ $\mathbf{G}_g = \sigma_g^2 \mathbf{I}$ $\mathbf{G}_{g\zeta} = \sigma_{g\zeta}^2 \mathbf{I}$ $\mathbf{G}_{gl} = \sigma_{gl}^2 \mathbf{I}$
RGC	$\alpha + \beta x_{jm} + \zeta_m + \lambda_m x_{jm}$	$l_{jm} + g_i + (g\zeta)_{im} + (gl)_{ijm}$	$\mathbf{G}_{l_{jm}} = \oplus_{m=1}^M \mathbf{G}_{l_m},\ \mathbf{G}_{l_m} = \sigma_{l_m}^2\, \mathbf{I}$ $\mathbf{G}_{g_i} = \sigma_g^2 \mathbf{I}$ $\mathbf{G}_{g\zeta} = \sigma_{g\zeta}^2 \mathbf{I}$ $\mathbf{G}_{gl} = \oplus_{m=1}^M \mathbf{G}_{(gl)_m},\ \mathbf{G}_{(gl)_m} = \sigma_{(gl)_m}^2 \mathbf{I}$
RGCQ	$\alpha + \beta x_{jm} + \gamma x_{jm}^2 + \zeta_m + \lambda_m x_{jm}$	$l_{jm} + g_i + (g\zeta)_{im} + (gl)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^M \mathbf{G}_{l_m},\ \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$ $\mathbf{G}_g = \sigma_g^2 \mathbf{I}$ $\mathbf{G}_{g\zeta} = \sigma_{g\zeta}^2 \mathbf{I}$ $\mathbf{G}_{gl} = \oplus_{m=1}^M \mathbf{G}_{(gl)_m},\ \mathbf{G}_{(gl)_m} = \sigma_{(gl)_m}^2 \mathbf{I}$
RC1	$\alpha + \beta x_{jm} + \gamma x_{jm}^2 + \zeta_m + \lambda_m x_{jm}$	$l_{jm} + \left(a_i + b_i x_{jm} + c_i x_{jm}^2\right)$ $+ \left(d_{im} + h_{im} x_{jm} + p_{im} x_{jm}^2\right)$ $+ (gl)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^M \mathbf{G}_{l_m},\ \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$ $\begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} \sim iid\, N\left(\mathbf{0}, \mathbf{G}_{g_i}\right)$ $\mathbf{G}_g = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma_c^2 \end{bmatrix}$ $\begin{bmatrix} d_{im} \\ h_{im} \\ p_{im} \end{bmatrix} \sim iid\, N\left(\mathbf{0}, \mathbf{G}_{(g\zeta)_{im}}\right)$ $\mathbf{G}_{g\zeta} = \begin{bmatrix} \sigma_d^2 & \sigma_{dh} & \sigma_{dp} \\ \sigma_{dh} & \sigma_h^2 & \sigma_{hp} \\ \sigma_{dp} & \sigma_{hp} & \sigma_p^2 \end{bmatrix}$ $\mathbf{G}_{gl} = \oplus_{m=1}^M \mathbf{G}_{(gl)_m},\ \mathbf{G}_{(gl)_m} = \sigma_{(gl)_m}^2 \mathbf{I}$
RC2	$\alpha + \beta x_{jm} + \gamma x_{jm}^2 + \zeta_m + \lambda_m x_{jm}$	$l_{jm} + g_i$ $+ \left(d_{im} + h_{im} x_{jm} + p_{im} x_{jm}^2\right)$ $+ (gl)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^M \mathbf{G}_{l_m},\ \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$ $\mathbf{G}_g = \sigma_g^2 \mathbf{I}$ $\begin{bmatrix} d_{im} \\ h_{im} \\ p_{im} \end{bmatrix} \sim iid\, N\left(\mathbf{0}, \mathbf{G}_{(g\zeta)_{im}}\right)$ $\mathbf{G}_{(g\zeta)_{im}} = \begin{bmatrix} \sigma_d^2 & \sigma_{dh} & \sigma_{dp} \\ \sigma_{dh} & \sigma_h^2 & \sigma_{hp} \\ \sigma_{dp} & \sigma_{hp} & \sigma_p^2 \end{bmatrix}$ $\mathbf{G}_{gl} = \oplus_{m=1}^M \mathbf{G}_{(gl)_m},\ \mathbf{G}_{(gl)_m} = \sigma_{(gl)_m}^2 \mathbf{I}$
RC3	$\alpha + \beta x_{jm} + \gamma x_{jm}^2 + \zeta_m + \lambda_m x_{jm}$	$l_{jm} + \left(a_i + b_i x_{jm} + c_i x_{jm}^2\right)$ $+ (g\zeta)_{im} + (gl)_{ijm}$	$\mathbf{G}_l = \oplus_{m=1}^M \mathbf{G}_{l_m},\ \mathbf{G}_{l_m} = \sigma_{l_m}^2 \mathbf{I}$ $\mathbf{G}_{g\zeta} = \sigma_{g\zeta}^2 \mathbf{I}$ $\begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} \sim iid\, N\left(\mathbf{0}, \mathbf{G}_{g_i}\right)$ $\mathbf{G}_g = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma_c^2 \end{bmatrix}$

*The index i refers to the i-th genotype, j refers to the j-th location, and m refers to the m-th zone

zone-based prediction of genotypes (Buntaran et al. 2019; Kleinknecht et al. 2013; Piepho and Möhring 2005; Piepho et al. 2016).

In the seven FG models, two models are LMM without covariates but with different variance–covariance structure for location and genotype×location effects, two models are LMM with covariate modelled by fixed effects, and three models are LMM with covariate interactions terms. Below, we describe each model in Table 1:

1. FG1 is the baseline LMM without covariate. The random-effect terms in this model are the effect of location nested in the m-th zone, $l_{jm} \sim N(0, \sigma^2_{l_m})$, and the interaction effect of genotype and location nested in the m-th zone $(\tau l)_{ijm} \sim N(0, \sigma^2_{(\tau l)_m})$. The variance–covariance structure for the location effect, l_{jm}, is a heterogeneous zone-specific covariance structure, $\mathbf{G}_l = \bigoplus_{m=1}^{M} \mathbf{G}_{l_m}$, where $\mathbf{G}_{l_m} = \sigma^2_{l_m} \mathbf{I}$ with $\mathbf{I}$ a J_m-dimensional identity matrix. The structure for the genotype×location effect, $(\tau l)_{ijm}$, is also a heterogeneous zone-specific covariance structure is $\mathbf{G}_{\tau l} = \bigoplus_{m=1}^{M} \mathbf{G}_{(\tau l)_m}$, where $\mathbf{G}_{(\tau l)_m}$ is the $J_m I$-dimensional diagonal matrix $\sigma^2_{(\tau l)_m} \mathbf{I}$, assuming that all genotypes were tested in all locations.

2. FG2 is the FG1 model but with homogeneous variance–covariance structure for the location and genotype×location effects. Thus, the location effect, l_{jm} has the J-dimensional structure $\mathbf{G}_l = \sigma^2_l \mathbf{I}$, and for the genotype×location effect $(\tau l)_{ijm}$, the JI-dimensional structure is $\sigma^2_{\tau l} \mathbf{I}$.

3. FGC is a model with a covariate. The covariate values, x_{jm}, are location-specific. In this model, the covariate is modelled by a linear trend. The notation for fixed regression terms involving the covariate is $\beta x_{jm} + \lambda_m x_{jm}$, where β is the fixed effect for the slope of the covariate and λ_m is the fixed effect for the zone-specific slope of the covariate for the m-th zone. The variance–covariance structures for genotype, location, genotype×zone, and genotype×location effects are the same as in the FG1 model.

4. FGCQ is the FGC model with an additional quadratic term for the covariate, x^2_{jm}. The slope for this quadratic term is denoted as γ. Note that in this model, the zone-specific interaction term of the quadratic covariate was not included because it was not significant. Thus, we decided to settle for the FGCQ model without the zone-specific quadratic term.

5. FGI1 is a model with covariate interactions terms for genotype, τ_i, and genotype×zone, $(\tau \zeta)_{im}$. Hence, this model has genotype and genotype×zone specific coefficients for the intercepts and slopes. FGI1 is the most complex model because it has linear and quadratic terms for the covariate interacting with genotype and genotype×zone. The model for covariate interaction in the genotype term is $\left(\tau_i + \beta_{2_i} x_{jm} + \beta_{3_i} x^2_{jm} \right)$ where τ_i is the genotype-specific intercept for the i-th genotype, β_{2_i} is the linear genotype-specific slope for the i-th genotype, and β_{3_i} is the quadratic genotype-specific slope for the i-th genotype. The model for covariate interaction in the genotype×zone term is $\left((\tau \zeta)_{im} + \beta_{4_{im}} x_{jm} + \beta_{5_{im}} x^2_{jm} \right)$, where $(\tau \zeta)_{im}$ is the genotype×zone-specific intercept for the i-th genotype in the m-th zone, $\beta_{4_{im}}$ is the linear genotype×zone-specific slope for the i-th genotype in the m-th zone, and $\beta_{5_{im}}$ is the quadratic genotype×zone-specific slope for the i-th genotype in the m-th zone.

6. FGI2 is a reduced version of the FGI1 model. The regression coefficients for the linear and quadratic term in the genotype main effect are removed.

7. FGI3 is a reduced version of the FGI1 model. The regression coefficients for the linear and quadratic term in the genotype×zone term are removed.

For the seven RG models, the differences compared to the seven FG models is that the genotype effect was random and there were three RC models due to the interactions between the covariate and genotype and genotype×zone effects. In the seven RG models, two models are LMM without covariates but with different variance–covariance structures for location and genotype×location effects, two models are LMM with covariate modelled by fixed effects, and three models are LMM with random coefficients for the regression on covariates. Below, we describe each model in Table 2:

1. RG1 is the baseline LMM without covariate and random coefficients. The random-effect terms in this model are the effect of genotype, $g_i \sim N(0, \sigma^2_g)$, the effect of location nested in the m-th zone, $l_{jm} \sim N(0, \sigma^2_{l_m})$, the interaction effect of genotype and zone, $(g\zeta)_{im} \sim N(0, \sigma^2_{g\zeta})$, and the interaction effect of genotype and location nested in the m-th zone, $(gl)_{ijm} \sim N(0, \sigma^2_{(gl)_m})$. The variance–covariance structure for the genotype effect, g_i, is $\mathbf{G}_g = \sigma^2_g \mathbf{I}$. The structure for the genotype×zone effect, $(g\zeta)_{im}$, is $\mathbf{G}_{g\zeta} = \sigma^2_{g\zeta} \mathbf{I}$. For the location effect, l_{jm}, the variance–covariance structure is a heterogeneous zone-specific covariance structure, heterogeneous zone-specific covariance structure, $\mathbf{G}_l = \bigoplus_{m=1}^{M} \mathbf{G}_{l_m}$, where $\mathbf{G}_{l_m} = \sigma^2_{l_m} \mathbf{I}$ with $\mathbf{I}$ a J_m-dimensional identity matrix. The structure for the genotype×location effect, $(gl)_{ijm}$, is also a heterogeneous zone-specific covariance structure is $\mathbf{G}_{gl} = \bigotimes_{m=1}^{M} \mathbf{G}_{(gl)_m}$, where $\mathbf{G}_{(gl)_m}$ is the $J_m I$-dimensional diagonal matrix $\sigma^2_{(gl)_m} \mathbf{I}$, assuming that all genotypes were tested in all locations.

2. RG2 is the RG1 model but with homogeneous variance–covariance structure for the location and genotype×location effects. Thus, the location effect, l_{jm} has

the J-dimensional structure $\mathbf{G}_l = \sigma_l^2 \mathbf{I}$, and for the genotype×location effect $(gl)_{ijm}$, the JI-dimensional structure is $\sigma_{gl}^2 \mathbf{I}$.

3. RGC is a model with a covariate. The covariate values, x_{jm}, are location-specific. In this model, the covariate is modelled by a linear trend. The notation for the fixed regression terms involving the covariate is $\beta x_{jm} + \lambda_m x_{jm}$, where β is the fixed effect for the slope of the covariate and λ_m is the fixed effect for the zone-specific slope of the covariate for the m-th zone. The variance–covariance structures for the genotype, location, genotype×zone, and genotype×location effects are the same as in the RG1 model. Thus, the random effects comprise no regression terms.

4. RGCQ is the RGC model with an additional quadratic term for the covariate, x_{jm}^2. The slope for this quadratic term is denoted as γ. As in the FGCQ model, the zone-specific interaction term of the quadratic covariate was not included because it was not significant. Thus, we decided to settle for the RGCQ model without the zone-specific quadratic term.

5. RC1 is a model with random coefficients. Since the covariate is location-specific, it was not possible to fit random coefficients for location. Thus, a RC model was fitted for the genotype term, g_i, and the genotype×zone term, $(g\zeta)_{im}$. Hence, this model has genotype and genotype×zone-specific coefficients for the intercepts and slopes. RC1 is the most complex model because it has linear and quadratic terms for the covariate, and allows a covariance between intercept and slope effects, meaning that $\mathbf{G}_g$ and $\mathbf{G}_{g\zeta}$ are unstructured variance–covariance matrices. It is essential to allow the covariance between intercept and slope to be a free parameter in order to ensure invariance with respect to translation and scale transformation of the covariates (Longford 1993; Piepho and Ogutu 2002). Note that when the covariance structure is diagonal, then the structure is only invariant to scale transformations but not to translations (Wolfinger 1996). The model for random coefficients in the genotype term was $g_i = a_i + b_i x_{jm} + c_i x_{jm}^2$, where a_i is the random genotype-specific intercept for the i-th genotype, b_i is the random linear genotype-specific slope for the i-th genotype, and c_i is the random quadratic genotype-specific slope for the i-th genotype. The model for random coefficients in the genotype×zone term was $(g\zeta)_{im} = d_{im} + h_{im} x_{jm} + p_{im} x_{jm}^2$, where d_{im} is the random genotype×zone specific intercept for the i-th genotype in the m-th zone, h_{im} is the random linear genotype×zone-specific slope for the i-th genotype in the m-th zone, and c_i is the random quadratic genotype×zone-specific slope for the i-th genotype in the m-th zone.

6. RC2 is a reduced version of model of RC1. The random regression coefficients for the linear and quadratic terms in the genotype main effect are dropped completely, so its covariance structure is $\mathbf{G}_g = \sigma_g^2 \mathbf{I}$.

7. In the RC3 model, only the random coefficients for the genotype×zone term is removed completely, so its covariance structure is $\mathbf{G}_{g\zeta} = \sigma_{g\zeta}^2 \mathbf{I}$.

Covariate selection and scaling

The covariate selection was done by extending the fixed-effects part of the FG1 model with the covariates as follows:

$$\alpha + \zeta_m + \tau_i + (\tau\zeta)_{im} + \beta_1 x_{1jm} + \beta_2 x_{2jm}$$
$$+ \beta_3 x_{3jm} + \beta_4 x_{4jm} + \beta_5 x_{5jm} + \beta_6 x_{6jm}$$

Parameters β_1 and β_2 are the coefficients of linear and quadratic terms for the scaled clay covariate, respectively, β_3 and β_4 are the coefficients of linear and quadratic terms for the scaled pH covariate, respectively, and β_5 and β_6 are the coefficients of linear and quadratic terms for the scaled humus covariate, respectively. The pH and humus covariates were standardised to mean 0 and standard deviation 1. The clay covariate was scaled as $(\text{clay} - 40)/10$ since this scaling resulted in non-negative variance estimates for the RC1 model. The fixed effect tests were adjusted with Kenward-Roger denominator degrees of freedom (Kenward and Roger 1997). If the F-test of a covariate effect was not significant at $\alpha = 5\%$, the covariate was dropped. The quadratic terms were tested first. The linear terms were tested linear only if the quadratic had to be dropped.

Predictions of genotypes in new locations

All models in Tables 1 and 2, which are special cases of Eq. 1, were fitted using the mixed model equations (MME) in Eq. 2 (Henderson 1950). The information used to estimate the $\mathbf{G}$ and $\mathbf{R}$ matrices in this MME are the records from the tested locations. Since the $\mathbf{G}$ and $\mathbf{R}$ matrices are unknown, they are estimated from the data via REML, producing estimates $\hat{\mathbf{G}}$ and $\hat{\mathbf{R}}$:

$$\begin{bmatrix} \mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{Z} + \hat{\mathbf{G}}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{y} \end{bmatrix} \quad (2)$$

The solution for $\hat{\boldsymbol{\beta}}$, which is the empirical BLUE (EBLUE), is obtained via generalised least-squares

$$\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X} \right)^{-1} \mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{y} \quad (3)$$

and $\widehat{\mathbf{u}}$, the empirical best linear unbiased predictor (EBLUP) of $\mathbf{u}$, is

$$\widehat{\mathbf{G}}\mathbf{Z}'\widehat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}) \tag{4}$$

McLean et al. (1991) discussed three types of inference space, i.e. the broad, narrow, and intermediate spaces. Estimates for these inference spaces are computed based on estimable functions $\mathbf{K}'\boldsymbol{\beta}$, which are linear combinations of fixed effects only, and predictable functions $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$, which consist of linear combinations of fixed effects and random effects $\mathbf{M}'\mathbf{u}$. The matrix $\mathbf{K}'$ consists of coefficients 1 and 0, where 1 indicates fixed effects needed in the estimable function, and covariate values for models that include covariates. The matrix $\mathbf{M}'$ contains the coefficients 1 and 0 according to the relevant random effects in the predictable function, and covariates for the RC models. The estimable function results in the BLUE, while the predictable function results in the BLUP. The differences between three types of inference space are explained in the Appendix.

In this study, the intermediate inference space (McLean et al 1991) is useful because it determines which genotype is the best in the particular environment. Hence, we focus on this inference space. The predictable functions for the EBLUPs in the new location can be expressed as

$$w = \mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u} + \mathbf{M}_0'\mathbf{u}_0 \tag{5}$$

where $\mathbf{K}'\boldsymbol{\beta}$ is the estimable function of fixed effects for zone. This term also includes fixed covariate terms for the models that use covariates. The matrix $\mathbf{K}'$ selects the fixed effects for the targeted zone where the new location is located, and the covariates for models with covariate. The term $\mathbf{M}'\mathbf{u}$ is the predictable function involving the genotype main effect and the genotype×zone interaction effects of the zone where the new location located, as well as any random covariate terms. The matrix $\mathbf{M}'$ selects the relevant random effects that can be estimated from the tested location; thus, it comprises the main effects of genotype and the genotype×zone interactions. For RC models, the matrix $\mathbf{M}'$ includes intercept and slope for the random-effects terms that have interactions with covariates. The last term, $\mathbf{M}_0'\mathbf{u}_0$, is the predictable function of the location main effect and genotype×location interaction for the new location, with $\text{var}(\mathbf{u}_0) = \mathbf{G}_0$. The matrix $\mathbf{M}'$ includes any covariate terms for the RC models. In the FG models, there is no $\mathbf{M}'\mathbf{u}$ term because the genotype and zone effects are fixed, and their interactions effects are also fixed.

The prediction of $\mathbf{M}_0'\mathbf{u}_0$ is always zero because there is no information on these effects for the new location. Thus, the random effect $\mathbf{u}_0$ is not estimated as such, but its distribution is needed in order to obtain the standard errors of predictions of genotypic values (SEPV) and standard errors of the predictions of pairwise differences of genotypic values (SEPD) for the genotypes in the new location. An important point to be observed here is that the random effects $\mathbf{u}_0$ for a new location are stochastically independent of the BLUP of $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$, which can be regarded as the conditional expectation of w, given fixed and random effects of the model for the observed data. Thus, the estimable effects can be expressed as a conditional mean as follows:

$$\eta = E(w|\boldsymbol{\beta}, \mathbf{u}) = \mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u} \tag{6}$$

Furthermore, on account of the independence assumption, $\text{var}(\mathbf{u}) = \mathbf{G}$, $\text{var}(\mathbf{u}_0) = \mathbf{G}_0$, and $\text{var}([\mathbf{u}, \mathbf{u}_0]) = \text{diag}(\mathbf{G}, \mathbf{G}_0)$, the conditional variance of w is:

$$\text{var}(w|\boldsymbol{\beta}, \mathbf{u}) = \text{var}(\mathbf{M}_0'\mathbf{u}_0) = \mathbf{M}_0'\mathbf{G}_0\mathbf{M}_0 \tag{7}$$

Precision measures: SEPV, SEPD, and prediction intervals

SEPV and SEPD in new locations

In the further derivation, it is key to observe that the estimate of η in (6) and the random effects u_0 for the new locations are stochastically independent. Thus, the total variance of the prediction of w in (5) is simply the sum of the variance of the estimate of η in (6) and the variance in (7). Hence, the SEPV in the new locations are computed as follows:

$$SEPV = \sqrt{\text{var}(\widehat{\eta}) + \text{var}(w|\boldsymbol{\beta}, \mathbf{u})} \tag{8}$$

where $\text{var}(\widehat{\eta})$ is the square of standard errors of EBLUPs for zone-based genotype average, and for our models $\text{var}(w|\boldsymbol{\beta}, \mathbf{u})$ is the sum of the variance components of effects for location and genotype×location. For the models with covariates, $\text{var}(w|\boldsymbol{\beta}, \mathbf{u})$ includes the variances and covariances for intercepts and slopes of random coefficients. In the FG models, $\text{var}(\widehat{\eta})$ is the square of the standard errors of the EBLUE of η for the zone-based genotype average.

For plant breeders and growers, the differences between genotypes are more informative than point estimates for individual genotypes. Thus, the pairwise prediction of differences and the SEPD needs to be computed. The predictions of pairwise differences can be computed using coefficients 1 and -1 in the $\mathbf{K}$ and $\mathbf{M}$ matrices according to the difference of interest genotypes.

The SEPD for the new locations are computed as:

$$SEPD = \sqrt{VDIFF(\widehat{\eta}) + 2 \times \sigma_{gl}^2} \tag{9}$$

where $VDIFF(\hat{\eta})$ is the square of the standard error of a difference of EBLUPs or EBLUEs for zone-based genotype averages, and σ_{gl}^2 is the variance component of genotype×location effects. In the FG models the σ_{gl}^2 refers to $\sigma_{\tau l}^2$. In (9), the latter is assumed homogeneous between zones, as in models FG2 and RG2. Alternatively, the variance may be assumed zone-specific, as in the other models listed in Tables 1 and 2.

Prediction intervals for the genotypes in the new locations

Meeker et al. (2017) define a prediction interval for a single future observation is an interval that will, with a specified degree of confidence, cover a future randomly selected observation from a distribution with pre-specified coverage probability $(1 - \alpha)$. Following this definition, the prediction intervals for genotype performances are regarded as a prediction interval to contain a single future observation. The important assumption of this interval is that the previously sampled locations and the future one can be regarded as random samples from the same distribution. The genotype×location effects for the new location can be regarded as a random sample from the same distribution as that of the genotype×location effects for locations in the dataset. The prediction interval for w is centred at η and the approximate $(1 - \alpha) \times 100\%$ prediction interval is given by:

$$\hat{\eta} \pm z_{1-\alpha/2}\text{SEPV} \tag{10}$$

where $z_{1-\alpha/2}$ is the $(1 - \alpha/2) \times 100\%$ quantile of the standard normal distribution. The prediction interval for pairwise differences between the i-th and i'-th genotype is given by:

$$DIFF(\hat{\eta}) \pm z_{1-\alpha/2}\text{SEPD} \tag{11}$$

where $\text{DIFF}(\hat{\eta})$ is the difference in predictions between two genotypes.

Accuracy measure

Cross-validation for model selection

A CV study can measure the prediction errors of the model using the mean squared error of prediction (MSEP) of difference. We conducted a leave-one-out CV for model comparison and selection. To mimic the prediction for new locations, we left one location out at a time and assigned its data as the validation set, and the data from the remaining locations as the training set. For the models with covariate, the covariate in the validation set was used for predictions. The MSEP was computed similarly to the MSEP proposed by Piepho (1998) for measuring the prediction accuracy of the models. The MSEP is a standard statistic for evaluating predictive accuracy. Let y

Table 3 Fixed effect tests for covariate selection by the extended FG1 model

Effect*	Numerator DF	Denominator DF	F Value	$Pr > F$
G	24	299	8.91	$< .0001$
Z	2	5.29	0.18	0.8394
G×Z	48	268	1.27	0.1212
†Clay	1	7.69	3.26	0.1100
†Clay squared	1	8.2	6.84	0.0303
‡pH	1	5.35	2.57	0.1656
‡pH squared	1	3.39	0.07	0.8094
‡Humus	1	6.87	0.01	0.9312
‡Humus squared	1	7.52	0.00	0.9993

*G, genotype; Z, zone; G×Z, genotype×zone

†Clay content was scaled to $(\text{clay} - 40)/10$

‡pH and humus were scaled to mean 0 and standard deviation 1

and z denote the observed and predicted values, respectively, and let I denote the total number of genotypes, and J the total number of locations. The assessment was measured based on the discrepancies between observed $(y_{ij} - y_{i'j})$ and predicted $(z_{ij} - z_{i'j})$ pairwise differences, since the main interest in cultivar trials is in prediction of differences between genotypes rather than performance of individual genotype (Piepho, 1998). The discrepancies between the observed and predicted pairwise differences from the 18 folds were accumulated and the MSEP was computed from this accumulation. The MSEP is computed as follows:

$$MSEP = \frac{\sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{i \neq i'}^{I} \left[y_{ij} - y_{i'j} - \left(z_{ij} - z_{i'j} \right) \right]^2}{JI(I-1)}.$$

The model producing the smallest MSEP is preferable because it predicts the yield differences in the validation set most accurately.

Implementation of the models

All models were implemented in SAS 9.4 (SAS Institute 2013) and ASReml-R 4.1.0.130 (Butler et al. 2017). We briefly describe the implementation strategy and visualisations for each package in the Supplementary Material.

Visualisation

In this section, some visualisation methods are proposed. In SAS, a prediction intervals plot to present predictions with prediction intervals can be generated using PROC SGPLOT. In this plot, the prediction intervals generated based on Eq. 10. A heatmap to present genotype pairwise differences can be generated using PROC TEMPLATE and

Table 4 Covariance parameter estimates of seven fixed-genotype-effect (FG) models

Covariance parameter	Subject[†]	Group	Estimate*						
			FG1	FG2	FGC	FGCQ	FCI1	FCI2	FCI3
Intercept	L	North	24,606.00		29,725.00	8596.23	11,052.00	11,052.00	8595.45
		Middle	82,207.00		33,466.00	30,888.00	38,357.00	38,357.00	30,893.00
		South	16,142.00		17,068.00	5625.77	6414.33	6414.33	5710.33
				45,118.00					
Intercept	G×L	North	598.57		598.60	598.59	206.18	206.18	534.64
		Middle	1182.36		1182.38	1182.38	1173.21	1173.21	1148.01
		South	1032.08		1032.28	1033.13	117.76	117.76	265.19
				952.72					

*Covariance parameters were estimated by REML

[†]L, location; G×L, genotype×location

PROC SGRENDER. The heatmap comprises the *p*-values of the genotype pairwise differences. These p-values are based on the *z*-test. The significance level for the pairwise differences was at $\alpha = 5\%$. For Bonferroni adjustment, the significance level α has to be divided by the total number of pairs or the p-values multiplied with the number of pairs (Hochberg and Tamhane 1987). In this case, the α was divided with 300 since there were 300 pairs. The *z* score was obtained by division of each predicted pairwise difference by its corresponding SEPD. Another way to present genotype pairwise differences is by generating a table comprising a letter representation of all-pairwise differences can be constructed with the insert-and-absorb algorithm (Piepho 2004). This algorithm is implemented in a SAS macro (Piepho 2012) and can be obtained from https://biostatistik.uni-hohenheim.de/fileadmin/einrichtungen/biostatistik/Tools_und_Macros/SAS-Macros/mult.sas. The basis of significance for this table is the same as the significance in the heatmap. In R, the prediction intervals plot and a heatmap can be generated using the *ggplot2* (Wickham 2016) and *corrplot* (Wei and Simko 2017) packages, respectively.

Results

The fixed-effects test results of the extended FG1 model for covariate selection are given in Table 3. Based on the F-test, the only significant covariate was the squared scaled clay content. Thus, in this study, we decided to use only the scaled linear and quadratic clay content as covariates for all 14 models.

In this section, we present the results of all models fitted by SAS. The results from ASReml-R are available as the Supplementary Materials. The fit statistics report including the deviance, information criterion, ΔDeviance and ΔAIC of the seven FG models and the seven RG models are given in Tables S8 and S9, respectively, in the Supplementary

Materials. The fit statistics from ASReml-R are given in Tables S3 and S4 for the seven FG models and seven RG models, respectively.

The covariance parameter estimates of the seven FG models are presented in Table 4. The corresponding ASReml-R outputs are presented in Table S1. The variance estimates for location and genotype×location effects in each zone were highly heterogeneous, as shown for the FG1 model. Compared to the FG2 model, the homogeneous variance structure for the location effect seems not appropriate because the homogeneous variance estimate for location was far larger than the heterogeneous variance estimate for location in the North and South zones, and far smaller for location in the Middle zone. For the genotype×location effect, the homogeneous variance estimate was larger than the heterogeneous genotype×location variance estimate for the North zone, but smaller than for the Middle and South zones. Thus, again, the FG2 model might be less appropriate than the FG1 model. The use of both a linear and a quadratic covariate term (the FGCQ model) decreased the estimate of the inter-location variance compared to the use of only a linear term (the FGC model). However, the estimate of the genotype×location variance did not change much compared to the estimate of the inter-location variance. When the covariates interacted with genotype and genotype×zone, as in the FGI1 model, the estimates of the inter-location variance were larger than the FGCQ model, but the estimate of the variance for the South zone in the genotype×location dropped considerably. There was no difference in any variance component estimates in the FGI1 and FGI2 models. Thus, unexpectedly, when the model comprised fixed effects of genotype×zone and covariates interactions, dropping the covariate interactions with the genotype main effect, variance component estimates did not change. In the FGI3 model, only the genotype main effect interacted with the covariates. Here, the heterogeneous variance component

Table 5 Covariance parameter estimates of seven random-genotype-effect (RG) models

Covariance parameter	Subject[†]	Group	Estimate*						
			RG1	RG2	RGC	RGCQ	RC1	RC2	RC3
Intercept	L	North	24,613.00		29,729.00	8598.12	8608.21	8601.49	8601.40
		Middle	82,195.00		33,452.00	30,874.00	30,887.00	30,872.00	30,817.00
		South	16,120.00		17,048.00	5618.09	5717.27	5715.54	5704.45
		–	–	45,148.00	–	–	–	–	–
Intercept	G		613.46	638.70	613.48	613.55	–	609.96	-
Intercept	G×Z		65.30	60.61	65.28	65.21	–	–	98.88
Intercept	G×L	North	584.45		584.48	584.48	261.99	257.87	551.31
		Middle	1157.63		1157.64	1157.61	1058.87	1060.31	1134.55
		South	1057.85		1058.05	1058.91	309.83	299.28	321.24
		–	–	941.55	–	–	–	–	–
Intercept (1)	G	–	–	–	–	–	460.48	–	527.32
Covariance (2,1)	G	–	–	–	–	–	−121.12	–	−46.90
Linear term slope (2)	G	–	–	–	–	–	61.25	–	415.82
Covariance (3,1)	G	–	–	–	–	–	−9.45	–	−0.43
Covariance (3,2)	G	–	–	–	–	–	30.38	–	207.89
Quadratic term slope (3)	G	–	–	–	–	–	2.53	–	111.97
Intercept (1)	G×Z	–	–	–	–	–	216.39	175.72	–
Covariance (2,1)	G×Z	–	–	–	–	–	113.50	102.30	–
Linear term slope (2)	G×Z	–	–	–	–	–	363.11	453.44	–
Covariance (3,1)	G×Z	–	–	–	–	–	−39.53	−19.84	–
Covariance (3,2)	G×Z	–	–	–	–	–	123.37	169.48	–
Quadratic term slope (3)	G×Z	–	–	–	–	–	98.35	107.75	–

*Covariance parameters were estimated by REML

[†]G, genotype; L, location; G×L, genotype×location; G×Z, genotype×zone

Table 6 Averages of SEPV over 25 genotypes of each new location of all 14 models

SEPV	Model	Location			
		N01	N02	S01	S02
		———— $g \times m^{-2}$ ————			
	RC2	113.13	121.42	88.10	89.38
	RC3	114.23	122.10	88.14	89.20
	RC1	113.18	121.46	88.16	89.45
	FGI3	114.98	122.89	88.19	89.20
	RGCQ	114.24	122.06	91.65	92.57
	FGCQ	114.73	122.52	92.51	93.42
	FGI1	127.76	142.03	92.88	97.02
	FGI2	127.76	142.03	92.88	97.02
	RG1	174.11	174.11	141.55	141.55
	FG1	174.42	174.42	142.17	142.17
	RGC	207.01	215.71	150.87	147.69
	FGC	207.27	215.96	151.45	148.28
	RG2	235.22	235.22	231.91	231.91
	FG2	235.50	235.50	232.21	232.21

Table 7 Averages of SEPD over 25 genotypes of each new location of all 14 models

SEPD	Model	Location			
		N01	N02	S01	S02
		———— $g \times m^{-2}$ ————			
	RC2	29.49	32.54	28.92	30.98
	RC1	29.52	32.50	29.22	31.26
	FGI1	31.36	37.77	24.74	33.16
	FGI2	31.36	37.77	24.74	33.16
	RC3	36.85	37.03	29.61	30.30
	RG1	36.97	36.97	48.39	48.39
	RGCQ	36.97	36.97	48.41	48.41
	RGC	36.97	36.97	48.39	48.39
	FG1	40.03	40.03	51.53	51.53
	FGCQ	40.03	40.03	51.55	51.55
	FGC	40.03	40.03	51.53	51.53
	FGI3	41.62	42.37	29.79	30.16
	RG2	45.84	45.84	45.85	45.85
	FG2	49.55	49.55	49.66	49.66

Table 8 Covariance parameter estimates of RC2 model with correlation between slope and intercept

Covariance parameter	Subject*	Group	Estimate
Intercept	L	North	8601.49
		Middle	30,882.00
		South	5715.53
Intercept	G		609.93
Intercept	G×L	North	257.83
		Middle	1060.30
		South	299.35
Intercept	G×Z		175.42
Linear term slope (2)	G×Z		453.35
Quadratic term slope (3)	G×Z		107.72
Corr(2,1)	G×Z		0.36
Corr(3,1)	G×Z		−0.14
Corr(3,2)	G×Z		0.77

*G, genotype; L, location; G×S, genotype×location; G×Z, genotype×zone

Table 9 The means squared error of prediction differences of 14 models

Ranking	Model	MSEP ($g^2 \times m^{-4}$)
1	RGCQ	3681
2	RG1	3681
3	RGC	3681
4	RG2	3688
5	RC2	3797
6	RC3	3839
7	RC1	3848
8	FG2	4000
9	FGCQ	4008
10	FG1	4008
11	FGC	4008
12	FGI3	4393
13	FGI1	7217
14	FGI2	7217

estimates for location and genotype×location effects were similar to the estimates in the FGCQ model.

Table 5 presents the covariance parameter estimates of the seven RG models. The ASReml-R results for these models are given in Table S2. Unlike the FG models, the RG models included genotype and genotype×zone variance components. The pattern of estimates for inter-location and genotype×location variances were similar to the FG1, FG2, FGC, and FGCQ models. Interestingly, for the RC1, RC2, and RC3 models, the estimate of the genotype×location variance for the South zone dropped substantially compared to the other RG models. However, the estimate of the inter-location variance for the North zone was higher for the RC3 model than for the other RC models.

The average SEPV and SEPD, over 25 genotypes by location and by all 14 models, are presented in Tables 6 and 7, respectively. The results from ASReml-R are given in Table S5 and S6, respectively. The highest SEPV were observed using the RG2 and FG2 models. These two models had no covariate and used a homogeneous variance structure for the location and genotype×location terms. The RC2 and RC3 models were the two models with the smallest SEPV. The RC2 model had the smallest SEPV in the locations N01, N02, and S01, while the RC3 model had the smallest SEPV only in the location S02. However, the difference between the RC2 and the RC3 models for location S02 was subtle. The FGI3 model, with covariate interaction only with the genotype main effect, also had a small SEPV compared with other FG models and even compared with the RG models without any covariate terms. It is also apparent that modelling a heterogeneous variance structure for location and genotype×location terms improved the SEPV compared to the homogeneous variance structure because the RG and FG models had smaller SEPV than the RG2 and FG2. Compared to the FG and RG model, the SEPV in the RC2 model was improved by 37–38% for the untested locations in the South zone, and by 30–35% for the untested locations in the North zone.

The SEPD values were smaller than the SEPV values because the computation in Eq. 9 was only based on the terms involving genotype effects. In general, the RC2 model had the smallest SEPD except in location N02. However, again, the difference in SEPD between models RC1 and RC2 was small at this location. The FGI1 and FGI2 model had the same and the smallest SEPD among the FG models. In the FGI1 model, both genotype and genotype×zone terms had interactions with the covariates, while in the FGI2 model, only the genotype×zone term interacted with the covariate.

The heterogeneous variance structure for location and genotype×location was beneficial since the SEPD in the FG1 and RG1 models were smaller than in the FG2 and RG2 models. The SEPD depended on the covariate as well. For the untested locations in the South zone, the SEPD were smaller in the model where the covariate interacted with genotype and genotype×zone terms. Concomitantly, when the covariate had no interactions with genotype and genotype×zone terms, the SEPD of the untested locations in the South zone were higher than in the North zone. The SEPD in the RC2 model decreased by 36–40% for the untested locations in the South zone and by 12–20% for the untested locations in the North zone, which implies that the genotype×location variance estimate in the North zone is much higher than in the South zone. This is also shown in Table 5.

Table 8 presents variance and correlations estimates of the RC2 model, since this was a promising model based on

Genotype	Lower 95% Prediction Limit	Prediction	Upper 95% Prediction Limit	
28949	853.93	1026.6	1199.3	
28128	824.09	996.76	1169.4	
27125	817.32	989.99	1162.7	
27599	811.05	983.73	1156.4	
24521	806.1	978.76	1151.4	
27130	803.04	975.71	1148.4	
27110	802.58	975.25	1147.9	
28209	802.54	975.22	1147.9	
27600	799.88	972.55	1145.2	
27593	799.48	972.15	1144.8	
24984	798.68	971.35	1144	
26742	795.75	968.44	1141.1	
27590	791.26	963.94	1136.6	
27543	790.61	963.28	1136	
28954	780.9	953.59	1126.3	
27592	780.39	953.07	1125.7	
25512	777.62	950.29	1123	
27669	772.78	945.45	1118.1	
25965	770.5	943.17	1115.9	
27546	769.29	941.97	1114.6	
27605	767.09	939.75	1112.4	
23286	766.85	939.53	1112.2	
27548	764.24	936.91	1109.6	
23524	746.99	919.65	1092.3	
22455	716.37	889.03	1061.7	

Yield (g/ha) ± prediction intervals

Fig. 1 Predictions and 95% prediction intervals for each genotype in location S01 by the RC2 model by PROC SGPLOT

the averages of SEPV and SEPD. The correlation between the linear slope term and the intercept was 0.36, which is low. However, the correlation between the quadratic term and the intercept was -0.14, which is very low as well. The correlation between quadratic slope and the linear slope was quite high, i.e. 0.76. Thus, there was no problem with collinearity between these two terms. The results of ASReml-R are given in Table S7.

The MSEP of the 14 models are listed in Table 9. The model with the smallest MSEP predicted yield differences in the validation set most accurately. The RG models outperformed the FG models because their MSEP were smaller than the FG models. Hence, although both the FGI1 and FGI2 models had the exact same SEPD, and their SEPD were fairly competitive to the SEPD in the RC models, these

two models were the least performant since their MSEP were the highest. The RG1, RGC, and RGCQ models had the smallest MSEP, and their values were the same when they were rounded. The RC model with the random coefficients for the genotype×zone term (RC2) ranked fifth but considering that the three smallest MSEP were equal, the MSEP of the RC2 model can be considered as the third smallest. Clearly, the pattern shows that the models with covariate interaction in the genotype or genotype×zone terms were less performant than the models without such interactions. Based on the MSEP, the best models were RGCQ, RGC, RG1, RG2, and RC2.

We demonstrate how the proposed visualisations of the results can be implemented in SAS and R. Predicted values of the genotypes and the 95% prediction intervals

Genotype	22455	23286	23524	24521	24984	25512	25965	26742	27110	27125	27130	27543	27546	27548	27590	27592	27593	27599	27600	27605	27669	28128	28209	28949
23286	.04035																							
23524	.14468	.24578																						
24521	.00094	.08758	.02036																					
24984	.00217	.13512	.03696	.39883																				
25512	.01692	.35500	.14485	.16175	.23341																			
25965	.03064	.44977	.20789	.10916	.16483	.40287																		
26742	.00300	.15881	.04547	.36069	.45987	.26554	.19160																	
27110	.00143	.10853	.02731	.45165	.44633	.19366	.13350	.40699																
27125	.00023	.04059	.00753	.34876	.25968	.08460	.05289	.22837	.30498															
27130	.00137	.10555	.02627	.45802	.44012	.18979	.13026	.40090	.49360	.31084														
27543	.00509	.20571	.06567	.29623	.38979	.32688	.24333	.42932	.33902	.17768	.33373													
27546	.03361	.46640	.22035	.10131	.15447	.38683	.48334	.18055	.12473	.04845	.12159	.23049												
27548	.04858	.46394	.27516	.07362	.11685	.32157	.41423	.13820	.09218	.03299	.08997	.18103	.43058											
27590	.00480	.19958	.06303	.30415	.39885	.31874	.23654	.43829	.34789	.18316	.34137	.49091	.22374	.17453										
27592	.01328	.32002	.12372	.18708	.26377	.46170	.36627	.29732	.22166	.10100	.21699	.36213	.35075	.28811	.35368									
27593	.00200	.12990	.03486	.40964	.48895	.22494	.15829	.44899	.45741	.26863	.45095	.37944	.14853	.11138	.38817	.25433								
27599	.00053	.06295	.01339	.43187	.33426	.12410	.08060	.29890	.38475	.41431	.39060	.23993	.07419	.05283	.24640	.14442	.34434							
27600	.00192	.12678	.03362	.41488	.48349	.22070	.15467	.44356	.46276	.27341	.45645	.37422	.14536	.10889	.38294	.25032	.49454	.34958						
27605	.03934	.49694	.24352	.08829	.13684	.35762	.45282	.16086	.10959	.04099	.10663	.20765	.46936	.46083	.20150	.32242	.13094	.06367	.12817					
27669	.02546	.41876	.18559	.12470	.18488	.43350	.46856	.21310	.15078	.06163	.14754	.26857	.45211	.38380	.26130	.39598	.17784	.09283	.17390	.42183				
28128	.00009	.02370	.00385	.26691	.18968	.05419	.03206	.16413	.22850	.40730	.23356	.12314	.02905	.01912	.12839	.06535	.19729	.32610	.20138	.02405	.03802			
28209	.00143	.10864	.02743	.45131	.44671	.19447	.13419	.40744	.49963	.30470	.49323	.33955	.12541	.09277	.34807	.22196	.45779	.38451	.46315	.11009	.15152	.22771		
28949	.00000	.00131	.00011	.04879	.02781	.00411	.00194	.02230	.03775	.10281	.03902	.01427	.00168	.00095	.01512	.00546	.02991	.06911	.03054	.00130	.00248	.15075	.03784	
28954	.01290	.31381	.12042	.19167	.26995	.45475	.35976	.30421	.22732	.10434	.22248	.36904	.34399	.28199	.36053	.49284	.26045	.14916	.25645	.31610	.38942	.06800	.22780	.00583

P-Value scale: 0.4996 (top) — .00017 (bottom)

Fig. 2 Heatmap of p-value of the genotype pairwise differences in location S01 by the RC2 model by PROC TEMPLATE and PROC SGRENDER. The significance level was adjusted using Bonferroni adjustment

for location S01 based on the RC2 model are depicted in Fig. 1. Figure from R is given in Figure S4. Figure 1 presents the upper and lower prediction limits, which were obtained from Eq. 10. For the prediction of pairwise differences of genotypes in location S01 by the RC2 model, a heatmap containing p-values is given in Fig. 2. The heatmap produced by R is presented in Figure S5. The heatmap can be useful for data with a large number of genotypes and can provide an easy way to detect, by colour coding, which pairwise comparisons of genotypes are significant. The red shading indicates significant differences of pairwise genotypes predictions at $\alpha = 5\%$ with Bonferroni adjustment. The SAS and R codes for this purpose are available in the electronic Supplementary Material.

The pairwise differences displayed by a letter-based representation using the insert-and-absorb algorithm (Piepho 2004) are given in Table 10. The drawback of presenting letter-based representation of multiple comparisons is the limitation of the alphabet numbers and the clutter when many letters are required. When the needed number of letters exceeds 26, then this method will not work with the Latin alphabet, which can occur for a very large number

of genotypes. The SAS code to generate Table 10 in the electronic Supplementary Materials.

Discussion

In this study, it is shown that using clay in the quadratic term as a covariate and employing RC models reduced the SEPV averages for all new locations by 30–38%, and the SEPD averages by 12–40%. This shows that the RC models can improve the precision of predictions of genotypes performance and the precision of genotypes comparisons. By using the covariates in the random coefficients term, the SEPV is evaluated at specific values of the covariates (Milliken and Johnson 2002), which can substantially decrease the SEPV of the RC models compared to the models without any random coefficients. Between the 14 models, RC2 was the model with the smallest average SEPV for three new locations, except in location S02, the SEPV of the RC2 model was 0.18 higher than the RC3 model. Thus, it was minor. For the average SEPD, the RC2 model had only a marginally larger average SEPD in location N02 than the RC1 model.

Table 10 The letter representation of all-pairwise differences by the insert-and-absorb algorithm (Piepho 2004)

Genotype	Yield (g/ha)*	Letters		
28,949	1026.60	a		
28,128	996.76	a	b	
27,125	989.99	a		c
27,599	983.73	a		c
24,521	978.76	a		c
27,130	975.71	a		c
27,110	975.25	a		c
28,209	975.22	a		c
27,600	972.55	a		c
27,593	972.15	a		c
24,984	971.35	a		c
26,742	968.44	a		c
27,590	963.94	a		c
27,543	963.28	a		c
28,954	953.59	a		c
27,592	953.07	a		c
25,512	950.29	a		c
27,669	945.45	a		c
25,965	943.17	a		c
27,546	941.97	a		c
27,605	939.75	a		c
23,286	939.53	a		c
27,548	936.91	a		c
23,524	919.65		b	c
22,455	889.03			c

*Means not sharing any letter are significantly different at the 5% level of significance with Bonferroni adjustment (Piepho, 2018)

Nevertheless, the difference in average SEPD between the RC1 and RC2 models was small. The RC2 model utilised the linear and quadratic term in the genotype×zone interaction effects, but not in the genotype main effects. Figure S1 presents the quadratic regression per genotype and shows that the variation between genotypes is not that large, as the lines of genotypes are close to each other.

For this reason, the inclusion of random coefficients in the genotype term may not be worthwhile. The random coefficients in the genotype×zone term were more beneficial than in the genotype main-effect term, since the SEPD was higher when using the RC3 model than when using the RC2 model. Also, Figure S2 shows that the variation between genotypes×zone effects is large, as the lines of genotypes×zone are more widely spread out compared to Figure S1.

The model selection can be made by jointly considering SEPV and SEPD with MSEP from the CV study. In terms of MSEP, the FG models are clearly not favourable since their MSEP were higher than those of the RG models. This poor performance is expected because in difference to the

RG models, the FG models cannot borrow strength across zones. Among the RG models, there was no clear winner when jointly considering the SEPV, SEPD, and MSEP. The precision was certainly improved via the RC2 model, but its MSEP was not the smallest. The models with the smallest MSEP were RGCQ, RGC, and RG. Nevertheless, the difference of the MSEP between the RC2 and the RG, RGC, RGCQ model was 116 $g^2 \times m^{-4}$. Thus, the RC2 model predictions were less accurate $\left(\sqrt{\text{MSEP}} = 10.73 \text{ g} \times m^{-2} \right)$ compared to those three models. With the RG1 model, the prediction accuracy was slightly better, but the intervals and the uncertainty were larger compared to the RC2 model. Thus, the RC2 model is preferred based on the SEPV, SEPD, and the MSEP. The RC2 model minimised the uncertainty as indicated by lower SEPV and SEPD, although its MSEP was not the smallest.

Our study is similar to Jarquin et al. (2014) in that both use RC models to improve prediction accuracy. The major difference is that Jarquin et al. (2014) used an extensive number of environmental covariates with marker data and utilised this information by computing an environmental kinship matrix, for which a single variance component was fitted. Thus, implicitly, the model used by Jarquin et al. (2014) assumes that the slopes for the different covariates have the same variance and that there is no correlation between them. By contrast, in our study, a vast number of covariates and marker data were not available, but we allowed for heterogeneity in variance between slopes, and for covariance between slopes and intercepts to maintain the invariance feature of RC models (Longford 1993; Piepho and Ogutu 2002; Wolfinger 1996). It is acknowledged that we can afford to do so because we do not have a vast number of covariates. With an increasing number of covariates, fitting such RC models becomes more challenging, and it is not obvious how this can best be done. One option to circumvent numerical problems is to fit a low-rank approximation to the unstructured variance–covariance matrix for intercepts and slopes, i.e. a factor-analytic model (Jennrich and Schuchter 1986). Fitting a factor-analytic model guarantees that the variance–covariance matrix is positive definite. If the order of the factor-analytic model equals the number of slope terms plus the intercept, the model is equivalent to the unstructured model, whereas lower-rank approximations are obtained by reducing the order. Moreover, with a large number of covariates, covariate selection will be beneficial. R-square (R^2) for mixed models (Piepho 2019) is an option for covariate selection, as demonstrated in Hadasch et al. (2020). The best approach to accommodating a larger number of covariates certainly deserves further research.

The BLUPs for the untested locations are the predictions obtained from the observed locations, and these values are reported to growers. However, these BLUPs do not equal

the yield that the growers will get in practice. Furthermore, the standard errors of the predictions based on the observed locations cannot be applied to the grower's fields since the trials hardly ever coincide with these. Hence, the prediction intervals for the untested locations need to be computed and presented so that the growers have a view of how precise each cultivar's prediction is for their fields. To compute valid standard errors for the untested locations, the location effect needs to be modelled as random to account for the uncertainty of effects for the untested locations. Note that the random location main effect, as well as the random cultivar-location interaction effects, are part of the deviations from the regression curves. Prediction intervals are passed on the random variation around the curves. Hence, only if these effects are indeed present and modelled as random, can we compute valid prediction intervals for new locations. If the location effect is fixed, which is known as factorial regression, then the untested locations' uncertainty cannot be obtained since there is no variance component estimate for a fixed effect and the effect itself is unobservable. Besides, when the location effect is fixed, the standard errors are only valid for the locations where the trials are carried out. The next step, after having developed the uncertainty measure of the untested locations, could be to apply the RC model for a whole country to present maps of yield predictions using environmental covariates based on the available GIS data in the Swedish official cultivar trials.

When the covariate was modelled using the quadratic term, the averages of SEPV dropped considerably. Furthermore, using RC models, the reduction of the averages SEPV and SEPD depends on the covariate value in the new locations; Table 8 shows that the reduction of SEPV and SEPD varied between the four new locations. Gozdowski et al. (2017) reported that the relationships between the fine soil fractions for the 0–60 cm layers and yield are curvilinear, which also was the case in this study. The clay content used in this study was measured for the 0–25 cm layer. The SEPV depends on the variance estimates of location and genotype×location. The SEPV will decrease when the variance of location or genotype×location decrease. For the SEPD, however, using a covariate did not decrease SEPD because the SEPD only depends on the terms associated with genotype. Moreover, all variance estimates of effects for genotype, genotype×zone, and genotype×location, are similar for the RG1, RG2, RGC, and RGCQ models, while for the RC models these three terms have different variance estimates compared to the RG1, RG2, RGC, and RGCQ models (Table 5). For the FG models, the same pattern occurred in that the genotype×location variance estimates of FG1, FG2, FGC, and FGCQ models were similar.

The covariate scale is crucial when implementing RC models. We found this covariate scale issue when we fitted the RC1 model. This scaling issue occurred when the

random coefficients were fitted to both genotype and genotype×zone effects:

1. When the value subtracted from the covariate was lower than the covariate's mean, and the result was divided by 10, the variance component of the linear term of the random coefficients of the genotype main effect was 0.
2. The same results were obtained when the covariate was standardised to mean 0 and standard deviation 1 and when the value subtracted from the covariate was the covariate's mean, and the result was divided by 10.

The genotype main effect had a lower variance than the genotype×zone interaction effect, as shown in Figures S1 and S2. Thus, there was likely competition between genotype main effects and genotype×zone interaction effects in absorbing the variance because in the RC2 model, in which the random coefficients were only fitted for the genotype×zone interaction effect, the scaling of (1) and (2) caused no issue in the variance component estimates and the convergence. Besides, using the scaling of (1) and (2) in the RC2 model did not change the SEPV and SEPD for all new locations due to the invariance property ensured by allowing a covariance among random coefficients (unstructured variance–covariance). Other scalings such as subtraction-of-the-minimum and covariate-centring were also attempted. When these scalings were used, the RC models did not converge. Thus, the covariate scaling is essential for model convergence and to obtain the appropriate variance parameter estimates. The clay covariate was scaled for all fitted models by $(clay - 40)/10$. We used this scaling because it yielded a positive definite variance–covariance matrix. The choice of covariates is important in the fixed-effects and random-coefficients parts since they contribute to the improvement or worsening of the prediction precision. Furthermore, the scaling and linear transformations of the covariate are also essential since it can affect convergence and the estimates of covariance parameters.

Zone stratification improved precision because information can be borrowed between zones. However, it should be emphasized that a clear agro-ecological division is needed in order to improve precision for each zone. In this study, the North zone had slightly higher uncertainty than the South zone, which shows that in the North zone, geographical and environmental conditions might be more varied and predictions more uncertain. The SEPV of a zone can be improved by conducting more trials within a zone.

The RC models are useful for MET analysis because the covariate information, which is changing from one to the other trials, can be exploited in the random effects. This feature is useful to measure the adaptability of the genotypes in new environments, given the covariates of the new environment are available. In fact, the benefit of using a covariate in

the RC models depends on choosing the appropriate covariate. Covariates can initially be selected based on the biological considerations. Still, it is necessary to check whether these covariate candidates improve the model fit.

The computational approach for predictions in the new locations is different from Henderson (1977), who considered predictions for animals that were not in the records. The main difference is that in our models the location-specific random effects are independent between the observed dataset and the new location, whereas with the Henderson (1977) approach the random effects for animals with records are correlated by pedigree with the animals having no records. Moreover, if there is no zone stratification, the predictions of genotypes in the new locations are merely the summation of genotype BLUPs and the grand mean.

In our study, the precision of these predictions was additionally improved by incorporating covariates using RC models in terms of decreasing standard errors of predictions and standard error of pairwise differences of genotype predictions. Using RC models requires that covariates are available at each location. When the dataset is augmented with data of the new location, the computing strategy needs to be adjusted due to fact that there are no observations on the response in the new locations. For SAS, the augmented dataset could be used to conduct the analysis. However, for ASReml-R, the augmented dataset could not be used due to a different functionality of the package. Thus, the computing strategy needed to be adjusted by excluding the new locations from the dataset to conduct the analysis. The SEPV and SEPD of the new locations had to be computed after the analysis was completed in ASReml-R (see Supplementary Materials).

We demonstrated that predicting genotype yield for some new locations can be enhanced by inclusion of environmental covariates by borrowing information from other zones. A recent paper by Neyhart et al. (2021) followed the same spirit of predicting genotype performance in unobserved environments. Neyhart et al. (2021) assessed the genome-wide predictions in the unobserved environments for both between and within breeding generations. Resende et al. (2020) recently proposed the geospatial (geographic information system) genotype–environment interaction (GIS–GEI) method within an enviromics framework. This framework involves the joint analysis of MET data accounting for phenotypic, genotypic and envirotypic sources of information. It is anchored into a geoprocessing environment that employs enviromic markers, e.g. time-trend climate data, landscape or and management treatment information, obtained by means of modern envirotyping techniques. Our proposed RC modelling approach is ideally suited for integration in an enviromics-driven GIS–GEI framework. Moreover, the RC modelling can be used in the Bayesian framework as proposed by Theobald et al. (2002), who used a Bayesian method for making predictions with incorporating environmental covariates.

Conclusion

This study showed that the RC models can be used to improve the precision of the yield predictions of winter wheat genotypes in some new locations. The RC model RC2 was competitive, with regards to MSEP, compared to the RG, RGC, and RGCQ models. The RC2 model, with random coefficients of linear and quadratic terms in the genotype×zone effect, can be recommended based on joint consideration of precision in predictions and accuracy. The RC models improved the precision of the predictions for a new location by utilising covariate information in the new location in the random effects part, and by borrowing information from other zones via genetic correlation between zones. The crucial keys for improving the precision of the predictions in the RC models are the selection of suitable covariates, suitably scaling the covariate, modelling the appropriate trend for the covariate in the fixed-effects part, and using an unstructured variance–covariance for the random coefficients. The scale of the covariate is essential to obtain reliable variance component estimates and avoid convergence issues. Failing to select suitable covariates and to model the trend for the covariate also in the fixed-effects part may lead to larger SEPV and SEPD for new locations. Breeders can use RC models to determine the adaptability of tested genotypes in new environments. Agronomists and growers can use RC models to identify the best locally adapted genotypes.

Appendix

The differences between three types of inference space in McLean et al. (1991) are explained as follows:

1. The broad inference space refers to inference based exclusively on fixed effects and estimable functions, $\mathbf{K}'\boldsymbol{\beta}$. For example, the estimable function for the adjusted mean of the Middle zone is obtained by choosing $\mathbf{K}'$ accordingly, i.e. $\mathbf{K}' = [1,1,0,\dots,0]$ gives $\mathbf{K}'\boldsymbol{\beta} = \mu + \zeta_1$, where the first coefficient of 1 refers to the grand mean, μ, and the second to the fixed effect of the Middle zone, ζ_1. This is so, because the levels of the zone effect are ordered alphabetically such that the effect of the Middle zone comes first.

2. The narrow inference space refers to predictable functions, but is not subject-specific. In the example for the broad inference space, the average of the Middle zone is determined by $\mu + \zeta_1$. There are no random effects involved. Alternatively, we may consider the average of the Middle zone conditionally on the random genotype

effects, which corresponds to the narrow inference space. The predictable function then is $\mu + \zeta_1 + \bar{g} + \left(\overline{g\zeta}\right)_1$, where $\bar{g}$ is the average of the main random genotype effect, $I^{-1} \sum g_i$, and $\left(\overline{g\zeta}\right)_1$ is the genotype average in the random genotype × zone interaction effect in the Middle zone, $I^{-1} \sum \sum (\zeta g)_{i1}$.

3. The intermediate inference space also involves random effects, but in difference to the narrow inference space, the intermediate inference space is subject-specific. For example, the performance of i-th genotype in the j-th location of the Middle zone is given by $\mu + \zeta_1 + s_{j1} + g_i + (g\zeta)_{i1} + (gs)_{ij1}$. The estimable function $\mathbf{K}'\boldsymbol{\beta}$ is $\mu + \zeta_1$ and the predictable function $\mathbf{M}'\mathbf{u}$ is $s_{j1} + g_i + (g\zeta)_{i1} + (gs)_{ij1}$.

Code availability The SAS and R code are available as electronic Supplementary Materials.

Data availability The data used in this study are available as electronic Supplementary Materials.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest.

Consent to participate (include appropriate statements) 'Not applicable'.

Consent for publication (include appropriate statements) 'Not applicable'.

Ethics approval (include appropriate approvals or waivers) 'Not applicable'.

References

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

Buntaran H, Piepho H-P, Hagman J, Forkman J (2019) A cross-validation of statistical models for zoned-based prediction in cultivar testing. Crop Sci 59:1544–1553. https://doi.org/10.2135/cropsci2018.10.0642

Buntaran H, Piepho H-P, Schmidt P, Rydén J, Halling M, Forkman J (2020) Cross-validation of stage-wise mixed-model analysis of Swedish variety trials with winter wheat and spring barley. Crop Sci 60:2221–2240. https://doi.org/10.1002/csc2.20177

Butler DG, Cullis B, Gilmour A, Gogel BJ, Thompson R (2017) ASReml-R reference manual, version 4. University of Wollongong, Wollongong

Crossa J, Vargas M, van Eeuwijk FA, Jiang C, Edmeades GO, Hoisington D (1999) Interpreting genotype × environment interaction in tropical maize using linked molecular markers and environmental covariables. Theoret Appl Genetics 99:611–625. https://doi.org/10.1007/s001220051276

Damesa TM, Möhring J, Worku M, Piepho H-P (2017) One step at a time: stage-wise analysis of a series of experiments. Agron J 109:845–857. https://doi.org/10.2134/agronj2016.07.0395

Denis JB (1988) Two-way analysis using covariates. Statistics 19:123–132. https://doi.org/10.1080/02331888808802080

Gillberg J, Marttinen P, Mamitsuka H, Kaski S (2019) Modelling G×E with historical weather information improves genomic prediction in new environments. Bioinformatics 35:4045–4052. https://doi.org/10.1093/bioinformatics/btz197

Gozdowski D, Leszczyńska E, Stępień M, Rozbicki J, Samborski S (2017) Within-field variability of winter wheat yield and grain quality versus soil properties. Commun Soil Sci Plant Anal 48:1029–1041. https://doi.org/10.1080/00103624.2017.1323091

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

Hastie T, Tibshirani R, Friedman J (2009) The Elements of statistical learning. Springer Series in Statistics. Springer Series in Statistics, Springer, New York

Henderson CR (1950) Estimation of genetic parameters. Ann Math Stat 21:309–310

Henderson CR (1977) Best linear unbiased prediction of breeding values not in the model for records. J Dairy Sci 60:783–787. https://doi.org/10.3168/jds.S0022-0302(77)83935-0

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

Hochberg Y, Tamhane AC (1987) Multiple comparison procedures. John Wiley & Sons Inc, New York

Jarquín D et al (2014) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theoret Appl Genet 127:595–607. https://doi.org/10.1007/s00122-013-2243-1

Jennrich RI, Schluchter MD (1986) Unbalanced repeated-measures models with structured covariance matrices. Biometrics 42:805–820. https://doi.org/10.2307/2530695

Kang MS, Gorman DP (1989) Genotype×environment interaction in maize. Agron J 81:662–664. https://doi.org/10.2134/agronj1989.00021962008100040020x

Kenward MG, Roger JH (1997) Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997

Kleinknecht K, Möhring J, Singh KP, Zaidi PH, Atlin GN, Piepho HP (2013) Comparison of the performance of best linear unbiased estimation and best linear unbiased prediction of genotype effects from zoned Indian maize data. Crop Sci 53:1384–1391. https://doi.org/10.2135/cropsci2013.02.0073

Kotz S, Read CB, Balakrishnan N, Vidakovic B, Johnson NL (2006) Accuracy and Precision. In: Encyclopedia of Statistical Sciences, vol 1. Wiley: New Jersey

Longford NT (1993) Random coefficient models. Oxford University Press, New York

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

McLean RA, Sanders WL, Stroup WW (1991) A unified approach to mixed linear models. Am Stat 45:54–64. https://doi.org/10.1080/00031305.1991.10475767

Meeker WQ, Hahn GJ, Escobar LA (2017) Statistical intervals. John Wiley & Sons Inc, Hoboken

Milliken GA, Johnson DE (2002) Analysis of messy data, volume III: analysis of covariance. CRC Press, Boca Raton

Neyhart JL, Gutiérrez L, Smith KP (2021) Using environmental similarities to design training sets for genomewide selection. Crop Sci 61:396–409. https://doi.org/10.1002/csc2.20303

Piepho H-P (1998) Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Theoret Appl Genet 97:195–201

Piepho H-P (2004) An algorithm for a letter-based representation of all-pairwise comparisons. J Comput Graph Stat 13:456–466. https://doi.org/10.1198/1061860043515

Piepho H-P (2012) A SAS macro for generating letter displays of pairwise mean comparisons. Commun Biometry Crop Sci 7:4–13

Piepho H-P (2018) Letters in mean comparisons: what they do and don't mean. Agron J 110:431–434. https://doi.org/10.2134/agronj2017.10.0580

Piepho H-P (2019) A coefficient of determination (R^2) for generalized linear mixed models. Biometrical J 61:860–872. https://doi.org/10.1002/bimj.201800270

Piepho H-P, Möhring J (2005) Best linear unbiased prediction of cultivar effects for subdivided target regions. Crop Sci 45:1151–1159. https://doi.org/10.2135/cropsci2004.0398

Piepho H-P, Ogutu JO (2002) A simple mixed model for trend analysis in wildlife populations. J Agric Biol Environ Stat 7:350. https://doi.org/10.1198/108571102366

Piepho H-P, Denis J-B, van Eeuwijk FA (1998) Predicting cultivar differences using covariates. J Agric Biol Environ Stat 3:151–162. https://doi.org/10.2307/1400648

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

Piepho H-P et al (2016) Stability analysis for a countrywide series of wheat trials in Pakistan. Crop Sci 56:2465–2475. https://doi.org/10.2135/cropsci2015.12.0743

Resende RT, Piepho H-P, Rosa GJM, Silva-Junior OB, e Silva FF, de Resende MDV, Grattapaglia D, (2020) Enviromics in breeding: applications and perspectives on envirotypic-assisted selection. Theoret Appl Genet. https://doi.org/10.1007/s00122-020-03684-z

SAS Institute (2013) SAS for Windows 9.4. SAS Inst., Cary, NC

Theobald CM, Talbot M, Nabugoomu F (2002) A Bayesian approach to regional and local-area prediction from crop variety trials. J Agric Biol Environ Stat 7:403–419. https://doi.org/10.1198/108571102230

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

van Eeuwijk FA, Malosetti M, Yin X, Struik PC, Stam P (2005) Statistical models for genotype-by-environment data: from conventional ANOVA models to eco-physiological QTL models. Aust J Agr Res 56:883–894. https://doi.org/10.1071/AR05153

van Eeuwijk FA, Bustos-Korts DV, Malosetti M (2016) What should students in plant breeding know about the statistical aspects of genotype × environment interactions? Crop Sci 56:2119–2140. https://doi.org/10.2135/cropsci2015.06.0375

van Eeuwijk FA et al (2019) Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. Plant Sci 282:23–39. https://doi.org/10.1016/j.plantsci.2018.06.018

Wei T, Simko V (2017) R package "corrplot": visualization of a correlation matrix version 0.84. Available from https://github.com/taiyun/corrplot

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York

Wolfinger RD (1996) Heterogeneous variance: covariance structures for repeated measures. J Agric Biol Environ Stat 1:205–230. https://doi.org/10.2307/1400366