Theoretical and Applied Genetics (2019) 132:3399–3411
https://doi.org/10.1007/s00122-019-03432-y

ORIGINAL ARTICLE

Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions

R. Rincent[1,2] · M. Malosetti[3] · B. Ababaei[4,5] · G. Touzy[1,2,6,7] · A. Mini[1] · M. Bogard[6] · P. Martre[4] · J. Le Gouis[1,2] · F. van Eeuwijk[3]

Received: 23 April 2019 / Accepted: 17 September 2019 / Published online: 27 September 2019

Key message **We propose new methods to predict genotype × environment interaction by selecting relevant environmental covariates and using an AMMI decomposition of the interaction.**

Abstract Farmers are asked to produce more efficiently and to reduce their inputs in the context of climate change. They have to face more and more limiting factors that can combine in numerous stress scenarios. One solution to this challenge is to develop varieties adapted to specific environmental stress scenarios. For this, plant breeders can use genomic predictions coupled with environmental characterization to identify promising combinations of genes in relation to stress covariates. One way to do it is to take into account the genetic similarity between varieties and the similarity between environments within a mixed model framework. Molecular markers and environmental covariates (EC) can be used to estimate relevant covariance matrices. In the present study, based on a multi-environment trial of 220 European elite winter bread wheat (*Triticum aestivum* L.) varieties phenotyped in 42 environments, we compared reference regression models potentially including ECs, and proposed alternative models to increase prediction accuracy. We showed that selecting a subset of ECs, and estimating covariance matrices using an AMMI decomposition to benefit from the information brought by the phenotypic records of the training set are promising approaches to better predict genotype-by-environment interactions ($G \times E$). We found that using a different kinship for the main genetic effect and the $G \times E$ effect increased prediction accuracy. Our study also demonstrates that integrative stress indexes simulated by crop growth models are more efficient to capture $G \times E$ than climatic covariates.

Introduction

Climate change is expected to increase the frequency and intensity of drought and heat stress and has been shown to already affect wheat yield in Europe (Brisson et al. 2010). The need to limit nitrogen input to reduce ground water pollution also results in more stressing conditions. Wheat being one of the three main staple crops, it is necessary to

Communicated by Mikko J. Sillanpaa.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03432-y) contains supplementary material, which is available to authorized users.

✉ R. Rincent
renaud.rincent@inra.fr

[1] INRA, UMR 1095 Génétique, Diversité et Ecophysiologie des Céréales, 5 Chemin de Beaulieu, 63100 Clermont-Ferrand, France

[2] Université Blaise Pascal, UMR 1095 Génétique, Diversité et Ecophysiologie des Céréales, 63178 Aubière Cedex, France

[3] Biometris, Wageningen University and Research Center, PO Box 100, 6700 AC Wageningen, The Netherlands

[4] LEPSE, INRA, Montpellier SupAgro, Université Montpellier, 34060 Montpellier, France

[5] Native Trait Research, Limagrain Europe, 63720 Chappes, France

[6] Arvalis Institut Du Végétal, 6 Chemin de la Côté Vieille, 31450 Baziège, France

[7] BIOGEMMA, Genetics and Genomics in Cereals, 63720 Chappes, France

increase its performance although the environment becomes less favourable. Selecting new varieties adapted to these new environmental conditions seems to be a viable alternative (IPCC 2013; Parent et al. 2018; Asseng et al. 2019). However, the multiplicity of environmental conditions prevents the evaluation of the selection candidates in all environments.

The use of molecular markers to predict the performance of potentially unphenotyped individuals thanks to genomic selection (GS, Meuwissen et al. 2001) models could help breeders to identify efficient varieties for a given environment. But in contrast to standard GS models, in which the genetic value is considered to be the same in any environment, we need to predict the genotype-by-environment interactions ($G\times E$). Recent theoretical developments based on the mixed model theory were proposed in that sense using training sets phenotyped in multi-environment trials (MET). It was first proposed to model $G\times E$ by attributing environment-specific marker effects (Schulz-Streeck et al. 2013; Lopez-Cruz et al. 2015; Crossa et al. 2016) or by modelling environmental covariances (Burgueño et al. 2012). In other studies, environmental covariates were introduced in the GS model (Heslot et al. 2014; Jarquín et al. 2014; Malosetti et al. 2016; Ly et al. 2017, 2018), which allows predictions in new environments. In these last studies, environments are characterized by environmental covariates (EC) which are directly introduced in the statistical model as in a factorial regression (Denis 1988) or are used to estimate covariances between environments. Environments with similar stress covariates are indeed expected to result in similar $G\times E$ pattern. Usually, a crop growth model (CGM) is used to simulate the phenological stages of the varieties, and the ECs are defined according to relevant stresses that are supposed to have an effect at critical growth stages (Brancourt-Hulmel 1999, 2000; Lecomte 2005). However, the available physical and climatic descriptors can be far from what the crop experienced in the field. Ly et al. (2017) proposed to use output of the CGM as EC and showed that it could capture more $G\times E$ variance than pedo-climatic EC.

We propose here to derive CGM outputs, specifically defined to reflect the stress experienced by the plants. It might be valuable to adapt the CGM with the objective of producing stress covariates as close as possible to what the crop experienced in the field. Another improvement to the existing models would be to define optimum subsets of ECs which best capture $G\times E$. In the existing literature, tens of climatic ECs are indeed included in the statistical model (with the exception of Ly et al. 2017, 2018 who considered only one EC). This might be a problem, because depending on the genotypes and environments that are considered, the ECs responsible for $G\times E$ will not necessarily be the same. Therefore, it may be interesting to select a subset of ECs to better model the covariance between environments. Another important point is that in the models proposed so far the estimation of covariance matrices is solely based on molecular markers and ECs, and so do not benefit from the information present in the phenotypic data collected in the MET. The $G\times E$ observed in the dataset, and for instance modelled with an additive main effect and multiplicative interaction (AMMI) decomposition, could be used to better estimate covariances. Finally, we supposed that the QTLs explaining the genetic main effect can be different from the QTLs explaining $G\times E$ and so we proposed to use a different kinship matrix for the two effects.

The objectives of the present study were to evaluate and improve prediction accuracy of $G\times E$ in a MET of elite winter wheat composed of 42 environments and 220 varieties. For this, we tested four possible improvements of the standard $G\times E$ models. First, we used the SiriusQuality CGM (Martre et al. 2006; Martre and Dambreville 2018) to derive integrative ECs reflecting the stress experienced by crops. Second, we selected subsets of ECs capturing most $G\times E$ variance. Third, we used an AMMI decomposition to improve the estimate of covariances between varieties and between environments using the phenotypic records available in the training set, in addition to markers and ECs. Fourth, we used a different kinship matrix for the main genetic effect and the $G\times E$ effect.

Materials and methods

Genetic material, genotyping and relatedness

We used a wheat panel composed of 220 European elite varieties of winter wheat (Ly et al. 2018; Rincent et al. 2018; Touzy et al. 2019). It was genotyped with the TaBW280K high-throughput genotyping array described in Rimbert et al. (2018). This array was designed to cover both genic and intergenic regions of the three subgenomes. Markers with minor allele frequency (MAF) below 5% or with heterozygosity or missing data rate above 5% were removed. Markers in strong linkage disequilibrium (LD) were filtered out using the pruning function of plink (Purcell et al. 2007) with a window of size 100 SNPs, a step of 5 SNPs and a LD threshold of 0.8. Eventually, we obtained 34,464 polymorphic high-resolution SNPs, with an average missing data rate of 1.0%. Missing values were imputed as the marker observed frequency.

Genotype of individual i at marker l ($M_{i,l}$) was coded as 1, 0.5, or 0 for homozygote for an arbitrarily chosen allele, heterozygote, and the other homozygote, respectively. Genomic relatedness (kinship) between individuals was estimated following VanRaden (2008):

$$K_{i,j} = \sum_{l=1}^{L} \frac{\left(M_{i,l} - p_l\right)\left(M_{j,l} - p_l\right)}{\mathrm{b}},$$

with $b = \sum_{l=1}^{L} p_l \times (1 - p_l)$, p_l being the allelic frequency of the reference allele in the corresponding diversity panel, L the number of markers.

Phenotypic data

Experimental designs and growth conditions

The panel was phenotyped for yield and heading time in a multi-environment trial composed of 42 environments located in France between 2012 and 2016 (Supplementary Table S1). These 42 environments correspond to 26 combinations of years and locations, with two treatments for 16 of them. Among these 16 combinations of years and locations, three had an irrigated (WW) and a rainfed (WD) treatment, one had a well-watered (WW) and a rainout shelter (RO) treatments, and 12 had a high-nitrogen (HN) and a low-nitrogen (LN) fertilizer treatments. For all environments except Cle16RO and Cle16WW, the replicates were composed of six blocks to which the varieties were attributed according to their earliness (split plot design with maturity as whole block treatment factor and genotype as sub plot treatment). Four check varieties were present in each block. One of the drought experiments was carried out under rainout shelters with an irrigated treatment next to it (Pheno3C phenotyping platform). For these two environments (Cle16RO and Cle16WW), the genetic material was slightly different: 228 individuals were phenotyped, with an overlap of 166 varieties with the panel used in the other environments. These genotypes were divided into eight groups of earliness, resulting in 8 blocks in the field, and a p-rep design was used (64 varieties were replicated twice in addition to the 4 checks present in each of the 8 blocks.)

Crops were sown at the recommended date and density at each site (Supplementary Table S1). Heading time was determined when 50% of ears were visible. At ripeness maturity, grains were collected using a combine harvester and grain yield was corrected to 0% moisture content.

Computation of adjusted means and estimation of heritabilities and variance components

The statistical models used to compute adjusted means and estimate generalized heritabilities (Cullis et al 2006) were defined to reflect the different experimental designs. In each environment with two replicates, 2-dimensional P-spline mixed models as implemented in the R package SpATS (Xose Rodriguez-Alvarez et al. 2018) were used, including the block and possibly the repetition as fixed effects. This spatial model adjusts for both global and local trends simultaneously. In the other environments, spatial effects were only corrected for replicate and block effects using the checks and if present the partially replicated varieties. The different statistical models used are presented in Supplementary Table S2.

Estimation of the variance components in a two-stage analysis

The adjusted means obtained above were used in a second stage together with their variance matrix to compute the environmental, genetic and interaction variances, as proposed in Smith et al. (2001): $\hat{Y}_{ij} = \mu + E_j + G_i + GE_{ij} + \varepsilon_{ij}$, where $\hat{Y}_{ij}$ is the adjusted mean of variety i in environment j estimated as above, and $\boldsymbol{E_j}$, $\boldsymbol{G_i}$ and $\boldsymbol{GE_{ij}}$ are random environmental, genotype and interaction effects, respectively. The residuals had the following distribution $\varepsilon_{ij} \sim N(0, E\sigma_e^2)$, with $E^{-1} = \mathrm{diag}(V^{-1})$, V being the covariance matrix computed at the first stage (Smith et al. 2001). ASREML-R (Gilmour et al. 2009) was used to estimate the variance components.

Environment data

For each environment, daily statistics on temperatures, radiation, precipitation and ETPs (Penman potential evapotranspiration from a crop canopy (mm/day) computed with Penman equation) were collected at close meteorological stations. These meteorological stations were next to the experimental site for most environments, and few kilometres away for the others. Altitude, longitude and latitude were also available. Management practices including sowing date, fertilization, irrigation and harvest date were available (Supplementary table S1). Soils were characterized for most environments by an estimate of soil depth, a texture analysis of the first layers (0–30 cm and 30–60 cm), and an analysis of the content of organic matter and nitrogen during winter. When soil information was missing, it was taken from the ARVALIS soil database.

AMMI analysis and extraction of a subset of environments and varieties

AMMI analysis of the full dataset

The AMMI decomposition is based on the following statistical model:

$$Y_{ij} = \mu_j + G_i + \sum_{k=1}^{K} \lambda_k a_{ik} b_{jk} + \varepsilon_{ij},$$

with Y_{ij} the adjusted mean of variety i in environment j, μ_j the mean of environment j, G_i the effect of variety i, a_{ik} is the genetic score of variety i on axis k, b_{jk} is the environmental score of environment j on axis k, λ_k is the singular value of axis k, and ε_{ij} are the residuals assumed to be independent and normally distributed. A weighted AMMI

could have been applied to this dataset (Rodrigues et al. 2014), but taking the weights into account in the mixed model analysis resulted in similar results, so we considered that an unweighted AMMI was sufficient to model $G\times E$ efficiently. The AMMI decomposition was based on three axes, capturing the most consistent interactions. These three axes captured 62% of the residual variability from the main effects model, which we believed to represent real $G\times E$ only and not error.

The interaction terms of the AMMI decomposition can be written as follows: $\Phi_{\mathrm{AMMI}} = A\Lambda B^t$, where A is the matrix of size ($N_{\mathrm{G}}\times 3$) of genetic scores, B is the matrix of size ($N_{\mathrm{E}}\times 3$) of environmental scores, and Λ is the diagonal matrix of size (3×3) with the singular values of the three axes on its diagonal.

Distances between the genotype-by-environment interactivity for varieties (D_G) and between the same interactivity for environments (D_E) were computed as the Euclidean distances between rows and columns of Φ_{AMMI}, respectively.

Sampling a subset of varieties and environments

Because of the important redundancy in the dataset and to limit computational time, it was reduced to a combination of 156 varieties and 20 environments. This sampling was based on the Ward (1963) hierarchical clustering of D_G and D_E, as implemented in the R package hclust. 156 groups of varieties were defined, and one variety was randomly sampled in each group, leading to a subset of 156 varieties. The same procedure was applied to the largest of the two clusters of environments, which was sub-clustered in 15 groups. This resulted in a subset of 20 environments (5 in group 1, and 15 in group 2), composed of 15 combinations of year and location, five of them having two treatments and ten others having only one treatment. All the analyses described below were applied to this subset with $N_{\mathrm{G}} = 156$ and $N_{\mathrm{E}} = 20$.

Similarly to the full dataset, an AMMI analysis was done, distance matrices D_G and D_E were computed and a hierarchical cluster analysis was applied to this subset.

Covariance between the interactivity of varieties (K_{AMMI}) and between interactivity of environments (W_{AMMI}) could then be computed as: $K_{\mathrm{AMMI}} = 1_G - \frac{D_G}{\max(D_G)}$ and $W_{\mathrm{AMMI}} = 1_E - \frac{D_E}{\max(D_E)}$, where 1_G and 1_E are matrices of 1 of size ($N_{\mathrm{G}} \times N_{\mathrm{G}}$) and ($N_{\mathrm{E}} \times N_{\mathrm{E}}$), respectively.

Environmental characterization using a crop growth model

Each environment was characterized by environmental covariates (EC, hereafter 'covariates') calculated from daily weather data and by variables (as in Heslot et al. 2014; Jarquin et al. 2014) and stress indices simulated by SiriusQuality. Because environmental conditions, and traits associated with environmental resilience, may have positive or negative effects depending on the developmental phase at which they occur (Tardieu 2012; Brancourt-Hulmel et al. 1999, 2000; Lecomte 2005), the environmental covariates (ECs) were calculated for each relevant developmental phase simulated in each environment by SiriusQuality.

Predicting development stage with the wheat crop growth model SiriusQuality

We used the wheat crop growth process-based model SiriusQuality (Martre et al. 2006; He et al. 2012; Martre and Dambreville 2018; https://www1.clermont.inra.fr/siriusquality/) that uses a modified version of the phenology model introduced in Jamieson et al. (1998). It was used to predict developmental stages of a virtual variety representative of the average earliness of the panel and dry mass stress indices. Soils rootable depth, silt, clay and sand percentages, and organic matter content were available for most environments.

The genotypic parameters of the phenology parameters of SiriusQuality (P, phyllochron; SLDL, day length response of leaf production; VAI, response of vernalization rate to temperature) of a virtual genotype representative of the average earliness of the panel were estimated based on the average heading date observed in the multi-environment trial using the Bayesian approach proposed in Rincent et al. (2017). It resulted in a root mean square error of 2.6 days. The model predicted for each environment the following stages based on the Zadoks scale (Zadoks et al. 1974): GS30 (pseudo stem erect), GS39 (flag leaf ligule just visible, male meiosis), GS55 (heading), GS65 (flowering), GS71 (grain water ripe, end of endosperm cell division) and GS91 (physiological maturity).

Estimation of climatic stress covariates

Two kinds of covariates were used to characterize the environmental conditions applied to the panel. The first kind of covariates are statistics on meteorological data by developmental stage (Supplementary Tables S3). These covariates are environmental characteristics, which do not take into account the physiology of the plants (except the average development stage of the panel). Seventy four of these covariates were estimated for each environment. Two of these covariates (ndt0f and st0f) had no variability and were removed.

Simulation of water, nitrogen and temperature stress with the crop growth model SiriusQuality

For the second kind of covariates, SiriusQuality was used to integrate the effect of environmental conditions on daily biomass production and to calculate ECs that

integrate the response of the crop to the severity of the stresses and their interactions (Supplementary Table S4). A daily dry matter stress index (DMSI) that directly relates the impacts of temperature, drought and N deficit, alone or in combination, to daily biomass loss was calculated as the ratio of actual to unlimited total aboveground biomass accumulation on that day:

$$\mathrm{DMSI}(d) = \frac{B_{\mathrm{act}}(d) - B_{\mathrm{act}}(d-1)}{B_{\mathrm{unl}}(d) - B_{\mathrm{act}}(d-1)}$$

where B_{act} is the biomass simulated under actual conditions and B_{unl} is the biomass simulated under unlimited temperature, water or N conditions. DMSI may range from 0 for very severe stress conditions with no biomass production to 1 for unstressed conditions with potential biomass production.

In order to calculate the DMSI, the source code of SiriusQuality was modified to enable it to execute each day simulations with the crop status under actual conditions from the previous day and with actual and unlimited temperature, water, and N, alone or in combination. Therefore, each day the model was executed five times. For unlimited water conditions, water limitation was removed by adding enough water to each soil layer to raise its moisture content back to field capacity each day. For unlimited N conditions, N was added to soil profile to provide the crop with sufficient N resources to fulfil its demand. For unlimited high-temperature conditions, simulated daily maximum canopy temperatures were capped to a temperature of 25 °C and simulated daily minimum canopy temperatures were reduced by the same amount to keep the daily canopy temperature range unchanged. Canopy temperature was modified instead of air temperature as SiriusQuality uses the former to simulate crop development and growth. For each developmental phase, the sum, maximum and average values of each DMSI were calculated. We could then integrate or take the maximum of these stress indexes for each developmental stage or for the whole crop cycle. SiriusQuality was also used to estimate the quantity of available nitrogen in the soil for each developmental phase, the nitrogen nutrition index (NNI, Justes et al. 1994) at flowering, and the post-flowering uptake of nitrogen. SiriusQuality produced 71 ECs in each environment (Supplementary Table S4). Four of these covariates (SQ_SF_Nfh_max, SQ_SF_Nhm_max, SQ_SF_Nfh_mean and SQ_SF_Nhm_mean) had no variability and were removed.

In total, 139 covariates were estimated in each environment, they were centred and scaled, and compiled in a matrix Ω of dimension ($N_{\mathrm{E}} \times 139$).

Estimation of an environmental covariance matrix using EC

We used the matrix of ECs (Ω) to estimate a covariance matrix between environments (W). In a first version of W, we applied the approach of Jarquin et al. (2014) in which all the ECs were used to estimate the covariance matrix (W_{all}). Environments with similar stresses are assumed to have similar $G\times E$ patterns. To compute W_{all}, we proceeded in two steps: first we computed the Euclidean distance matrix between environments (D_{Ω}) with the matrix of environmental covariates (Ω), and then the covariance matrix was computed as: $W_{\mathrm{all}}=1_E - \frac{D_{\Omega}}{\max(D_{\Omega})}$.

Prediction objectives and the corresponding cross-validation schemes

Four prediction objectives were considered: the prediction of observed varieties in observed environments (oGoE) or in new environments (oGnE), and the prediction of new varieties in observed environments (nGoE) or in new environments (nGnE). oGoE consists of predicting missing values in a multi-environment trial. This typically corresponds to the situation faced by breeders when some observations are missing in their trial networks. oGnE and nGoE are more ambitious because predictions are made in an environment or for a variety without any phenotypic information on it. nGnE is the most ambitious scenario, because both the variety and the environment are unobserved (no phenotypic record).

To evaluate the performance of the prediction models in these four situations, four cross-validation schemes were defined: CVrandom, CVnewG, CVnewE and CVnewGE.

oGoE was addressed by the CVrandom scheme, that is a fivefold cross-validation, in which the folds were randomly sampled from the dataset. oGnE was addressed by the CVnewE scheme, in which a new combination of year and location is predicted. This was a leave-one-out scheme, with the constraint that the two treatments of a same combination of year and location were both in training or both in validation. In other words, there were 15-folds, each fold corresponding to a combination of year and location. That way we took into account the fact that the two treatments in a year/location combination were not independent. nGoE was addressed by the CVnewG scheme, with a division in fivefolds consisting in five randomly selected groups of varieties. nGnE was addressed with CVnewGE scheme: here, the varieties were divided in fivefolds and the environments were divided in 15-folds, corresponding to the 15 combinations of year and location. So, in total this scenario had $5\times15=75$ folds.

Prediction accuracy was computed for each fold and each environment as the correlation between predictions and adjusted means in the validation set. For each

cross-validation scheme, the total procedure was repeated 10 times to get robust estimates of prediction accuracy.

Selection of environmental covariate subsets and use of AMMI scores to improve genotype-by-environment interaction predictions

In addition to the use of DMSI as stress covariates to better capture $G \times E$, we present in this section three main ways to improve prediction accuracy in the context of $G \times E$. First, we propose to select a subset of ECs efficient to capture $G \times E$ variance, instead of using all ECs without any filtering. In the second proposition, AMMI scores for varieties and environments were used to get a W estimate closer to the observed $G \times E$. Finally, we propose to evaluate the interest of using a different kinship matrix for the main genetic effect and the interaction effect, to take into account the fact that different QTLs can be involved in the two effects.

Improving the G × E predictions by selecting ECs for better W estimate

The idea here is that depending on the dataset, this will not necessarily be the same ECs that will be responsible of $G \times E$. Therefore, like for molecular markers in classical genomic prediction models, we can expect some of the variables to have a very low or no effect at all. In that case, it seems reasonable to select subset of variables. This would be also a way to deal with the strong redundancy present in the Ω matrix (139 ECs for 20 environments). This subset of ECs was determined in order to get an environmental covariance matrix (W_{sel}, estimated as above but with a subset of ECs) the most correlated to W_{AMMI}. For this, a stepwise forward procedure was applied to the 139 ECs: at each iteration, the EC that generated the highest increase of correlation between W_{sub} and W_{AMMI} was added to the subset (W_{sub} being the matrix W obtained with the subset of ECs). We considered that W_{sel} was obtained when the increase of correlation was below 0.01.

To illustrate the interest of using crop models to derive integrated ECs capturing $G \times E$, we applied the procedure to the subset of 72 climatic ECs, to the 67 integrated ECs and to the full set of 139 ECs, and compared the correlations between W_{sel} and W_{AMMI} in the three cases.

Using AMMI scores for varieties and environments to better model G × E.

The AMMI decomposition presented above is an efficient way to determine the main $G \times E$ trends in the dataset. The advantage of the resulting covariance matrices (K_{AMMI} and W_{AMMI}) is that they reflect the observed $G \times E$ and not only the a priori knowledge we have on the varieties (markers) and environments (ECs). These two covariance matrices can then be used to produce a $G \times E$ covariance matrix using a simple Kronecker product: $K_{\text{AMMI}} \otimes W_{\text{AMMI}}$. Using K_{AMMI} instead of K in the interaction term is also a way to take into account the fact that the QTLs underlying the genetic main effect and the interaction effect can be different. Therefore, we suppose that using a different kinship matrix for the two effects can be valuable.

Of course, in practice, the AMMI decomposition could be only applied to the training set. As a result, the genetic and environmental scores of the unobserved varieties and environments would be unknown. So, these scores need to be predicted to be able to estimate K_{AMMI} and W_{AMMI} for the full dataset (training and validation sets). For this, we fitted a genomic prediction model on the genetic scores, and a regression model on the environmental scores. These models can then be used to predict the scores of the new varieties and new environments using their marker and EC profiles (see next section for more details).

Statistical models used to make predictions in the cross-validation scenarios

In the first kind of models, varieties were related using the kinship matrix but there was no sharing of information between environments:

$$\text{EG}: \quad Y_{ij} = \mu_j + \boldsymbol{G_i} + \varepsilon_{ij}, \text{with} \quad \boldsymbol{G_i} \sim N\left(0, K \cdot \sigma^2_{g1}\right) \tag{1}$$

$$\begin{aligned} &\text{EG\_G} \times E: \quad Y_{ij} = \mu_j + \boldsymbol{G_i} + \boldsymbol{GE_{ij}} + \varepsilon_{ij}, \\ &\text{with} \quad \boldsymbol{GE_{ij}} \sim N\left(0, K \otimes \mathrm{I}_{N_E} \cdot \sigma^2_{g2}\right) \end{aligned} \tag{2}$$

I_{N_E} is an identity matrix of size the number of environments, and ⊗ is the Kronecker product. In model EG, the prediction of a variety is the same for any environment. In model EG_$G \times E$, $G \times E$ is taken into account, but there is no sharing of information between environments. As a consequence, this second model is only applicable to scenarios CVrandom and CVnewG.

In the second kind of models, the clustering of the environments in two groups was taken into account:

$$\begin{aligned} &\text{EG\_}(G \times E)_{\text{grp}}: \quad Y_{ij} = \mu_j + \boldsymbol{G_i} + \boldsymbol{GE_{ij}} + \varepsilon_{ij}, \\ &\text{with} \quad \boldsymbol{GE_{ij}} \sim N\left(0, K \otimes \mathrm{Z}_{\text{group}} \cdot \sigma^2_{g3}\right) \end{aligned} \tag{3}$$

where $\mathrm{Z}_{\text{group}}$ is a matrix of size ($N_{\text{E}} \times N_{\text{E}}$), with 1 s on the diagonal, and with off-diagonal elements equal to 1 or 0 depending on whether the two concerned environments were in the same group or not. Here, the prediction of a variety will be the same for environments in the same

group. We consider that there is no $G\times E$ within group of environments.

In the third kind of models, information between environments is shared thanks to an environmental covariance matrix estimated with the environmental covariates as proposed by Jarquin et al. 2014:

$$\begin{aligned}&\text{EG\_}G\times W_{\text{all}}: \quad Y_{ij}=\mu_j+\boldsymbol{G_i}+\boldsymbol{GE_{ij}}+\varepsilon_{ij},\\ &\text{with} \quad \boldsymbol{GE_{ij}}\sim N\left(0,K\otimes \mathrm{W}_{\text{all}}\cdot\sigma^2_{g4}\right)\end{aligned} \tag{4}$$

W_{all} was estimated as described above.

A second version of this model was defined

$$\text{EG\_}G\times W_{set}, \tag{5}$$

with W_{sel} instead of W_{all}. This was to estimate the effect of selecting an optimal subset of ECs on prediction accuracy. This model is not suitable for predictions to new environments, because the full dataset (including the validation set) was used to determine the subset of optimal ECs. So, we defined another model

$$\text{EG\_}G\times W_{selpred} \tag{6}$$

in which the procedure of EC selection was applied to the training set only.

In the last kind of models, the covariance matrices between environments and between varieties were obtained from the AMMI decomposition as defined above:

$$\begin{aligned}&\text{EG\_}(G\times W)_{\text{AMMI}}: \quad Y_{ij}=\mu_j+\boldsymbol{G_i}+\boldsymbol{GE_{ij}}+\varepsilon_{ij},\\ &\text{with} \quad \boldsymbol{GE_{ij}}\sim N\left(0,K_{\text{AMMI}}\otimes \mathrm{W}_{\text{AMMI}}\cdot\sigma^2_{g7}\right)\end{aligned} \tag{7}$$

This model was proposed to estimate the prediction accuracies that could be achieved if the covariance matrices were estimated with the $G\times E$ observed in the dataset. It is not suitable for prediction as the phenotypes of the validation set were used to estimate the covariances, but it could be used to place an upper bound to our prediction models.

To illustrate the interest of using a different kinship matrix for the main genetic effect and for the interaction term, we also tested the following model:

$$\begin{aligned}&\text{EG\_}G_{\text{AMMI}}\times W_{\text{all}}: \quad Y_{ij}=\mu_j+\boldsymbol{G_i}+\boldsymbol{GE_{ij}}+\varepsilon_{ij},\\ &\text{with} \quad \boldsymbol{GE_{ij}}\sim N\left(0,K_{\text{AMMI}}\otimes \mathrm{W}_{\text{all}}\cdot\sigma^2_{g8}\right)\end{aligned} \tag{8}$$

These last two models are again not suitable for predictions of new genotypes and environments because the AMMI decomposition is applied to the full dataset including varieties and environments in the validation set. In practice, the varieties and environments in the validation set would be unknown. However, these new varieties and new environments would be characterized by genotypic and environmental descriptors, that is genotypes and environmental covariates. So, we proposed a second version of model (7), in which the AMMI decomposition is made on the training set only. As above, the interaction of this AMMI can be written: $\Phi_{\text{AMMI}}=A\Lambda B^t$, but this time A is the matrix of size ($\mathrm{N_{G\_CS}}\times 3$) of genetic scores and B is the matrix of size ($\mathrm{N_{E\_CS}}\times 3$) of environmental scores, with $\mathrm{N_{G\_CS}}$ and $\mathrm{N_{E\_CS}}$ the number of varieties and of environments in the training set, respectively. As the AMMI was applied to the training set only, it was necessary to predict the genetic and environmental scores for the new varieties and environments. For this, in a second step, for each AMMI axis, the genetic scores of the new varieties (for CVnewG and CVnewGE) were predicted with a G-BLUP model calibrated on the genetic scores of the training varieties, and the environmental scores of the new environments (for CVnewE and CVnewGE) were predicted using a partial least square regression on the environmental covariates calibrated on the environmental scores of the training environments. For CVrandom, missing values were first imputed with model EG_$G\times W$, and then the AMMI was done on the full imputed dataset. We could then use these observed or predicted genetic and environmental scores to compute K_{AMMIpred} and $\mathrm{W_{AMMIpred}}$, as before. We obtained the following model:

$$\begin{aligned}&\text{EG\_}(G\times W)_{\text{AMMIpred}}: \quad Y_{ij}=\mu_j+\boldsymbol{G_i}+\boldsymbol{GE_{ij}}+\varepsilon_{ij},\\ &\text{with} \quad \boldsymbol{GE_{ij}}\sim N\left(0,K_{\text{AMMIpred}}\otimes \mathrm{W_{AMMIpred}}\cdot\sigma^2_{g9}\right)\end{aligned} \tag{9}$$

All the models are summarized in Table 1.

Results

Description of the MET and its subset

The experimental designs used in this MET allowed reaching high generalized heritability for GY (Supplementary Table S1), with an average of 0.84 and a minimum of 0.44 (Gre14WD) for the 42 environments. The average heritability was equal to 0.82 in the subset of 20 environments.

The correlation matrix (Fig. 1) and the hierarchical clustering (Fig. 2) on the full dataset for GY revealed strong similarities between environments: eight environments had a correlation above 0.85 with at least one other environment, and 29 had a correlation above 0.7 with at least one other environment. There were two major groups of environments: the first one comprised the five environments located in Gréoux, and the second one all other environments. It is interesting to note that in most cases the two treatments of a given combination of year and location clustered together, suggesting that these treatments had a low contribution to $G\times E$.

The analysis of the variance components revealed that the environmental variance was much higher than the genotype and the $G\times E$ variances in both the full dataset and

Table 1 Description of the eight models used to make predictions

Model abbreviation	Factors included		Structure of the cov. matrix for the $G\times E$ term	Suitable for prediction[b]
	Main effect (cov. matrix)			
	E[a]	G	$G\times E$	
EG (1)	x	K		Yes
EG_$G\times E$ (2)	x	K	$K\otimes \mathrm{I}_{N_E}$	Yes
EG_$(G\times E)_{grp}$ (3)	x	K	$K\otimes Z_{group}$	No
EG_$G\times W_{all}$ (4)	x	K	$K\otimes W_{all}$	Yes
EG_$G\times W_{sel}$ (5)	x	K	$K\otimes W_{sel}$	No
EG_$G\times W_{selpred}$ (6)	x	K	$K\otimes W_{selpred}$	Yes
EG_$(G\times W)_{AMMI}$ (7)	x	K	$K_{AMMI}\otimes W_{AMMI}$	No
EG_$G_{AMMI}\times W_{all}$ (8)	x	K	$K_{AMMI}\otimes W_{all}$	No
EG_$(G\times W)_{AMMIpred}$ (9)	x	K	$K_{AMMIpred}\otimes W_{AMMIpred}$	Yes

[a]The main environmental effect was fixed for all models

[b]This column specifies if the models are applicable in real-life experiment in which the phenotypes of the calibration set are unknown: four models (EG_$(G\times E)_{grp}$, EG_$(G\times W)_{AMMI}$, EG_$G_{AMMI}\times W_{all}$ and EG_$G\times W_{sel}$) are not applicable because they use information contained in the validation set, but they were tested to estimate the potential of new approaches. K is the kinship matrix estimated with the molecular markers. K_{AMMI} and W_{AMMI} are the genotype and environment covariance matrices obtained from the AMMI decomposition on the full dataset. $K_{AMMIpred}$ and $W_{AMMIpred}$ are the genotype and environment covariance matrices predicted from the AMMI decomposition on the calibration set. W_{all} and W_{sel} are the environment covariance matrices estimated with all ECs or with a subset, respectively. $W_{selpred}$ is the environment covariance matrix with a subset of ECs defined using the calibration set only

the subset (Table 2, Supplementary Fig. S1). These results show that the ratio of the $G\times E$ on the genotypic (G) variance was comparable between subset and full dataset, which consolidates the choice of the environments and varieties composing the subset.

Determination of an optimal subset of environmental covariates to capture genotype-by-environment interactions

The correlations between matrices W_{AMMI} and W_{sub} obtained with different subsets of ECs were plotted in Fig. 3. These correlations were always intermediate to high but varied markedly depending on the selected ECs. This figure clearly shows that working on an optimal subset of ECs could considerably increase the proportion of $G\times E$ that can be captured by W. The correlation between W_{sel} and W_{AMMI} was equal to 0.84, whereas the correlation between W_{all} and W_{AMMI} was equal to 0.56. Therefore, we expect the matrix W_{sel} to be more efficient than W_{all} to predict $G\times E$. Figure 3 also shows that the best integrative ECs are more informative than the best climatic ECs. The use of integrated ECs derived from SiriusQuality in addition to the climatic ECs permitted to increase the correlation between W_{sel} and W_{AMMI}. This correlation was equal to 0.84 when all ECs were used, 0.76 when only integrated ECs are used and 0.71 when only climatic ECs were used.

Only four climatic covariates (rdtmmf2, Nferti, rdtmw2, nd25ef) and three covariates simulated with *SiriusQuality* (SQ_SF_Tw_max, SQ_SF_T, SQ_SF_Nem) composed Wsel. Using these seven ECs instead of the full set of 139 ECs resulted in an increase in correlation with W_{AMMI} from 0.56 to 0.84. So, more $G\times E$ was explained with these seven ECs than with all ECs together. It is interesting to note that two of the three most important ECs were obtained by the crop growth model (SQ_SF_Tw_max and SQ_SF_T). The three first ECs are related to photothermal quotient between meiosis and flowering (rdtmmf2) and to the maximum value of DMSI for maximum daily temperature > 20 °C during the winter period (SQ_SF_Tw_max) and the cumulative DMSI for maximum daily temperature > 20 °C during the entire growth period (SQ_SF_T).

Evaluation of prediction accuracy of the proposed models

Prediction accuracies were intermediate to high for all scenarios, but it was higher for CVrandom and CVnewE than for CVnewG and CVnewGE (Table 3). There were large differences between the models, and their ranking was different from one scenario to another, but all combinations of models and cross-validation scenarios led to accuracies above 0.39. Concerning the first three models that do not include any EC, it was valuable to add an interaction term, particularly if the clustering in two groups of environments was taken into account (EG_$(G\times E)_{grp}$). The reference model EG_$G\times W_{all}$, similar to the proposition of Jarquin et al. (2014), performed better than EG_$(G\times E)_{grp}$ in CVrandom and CVnewG, but

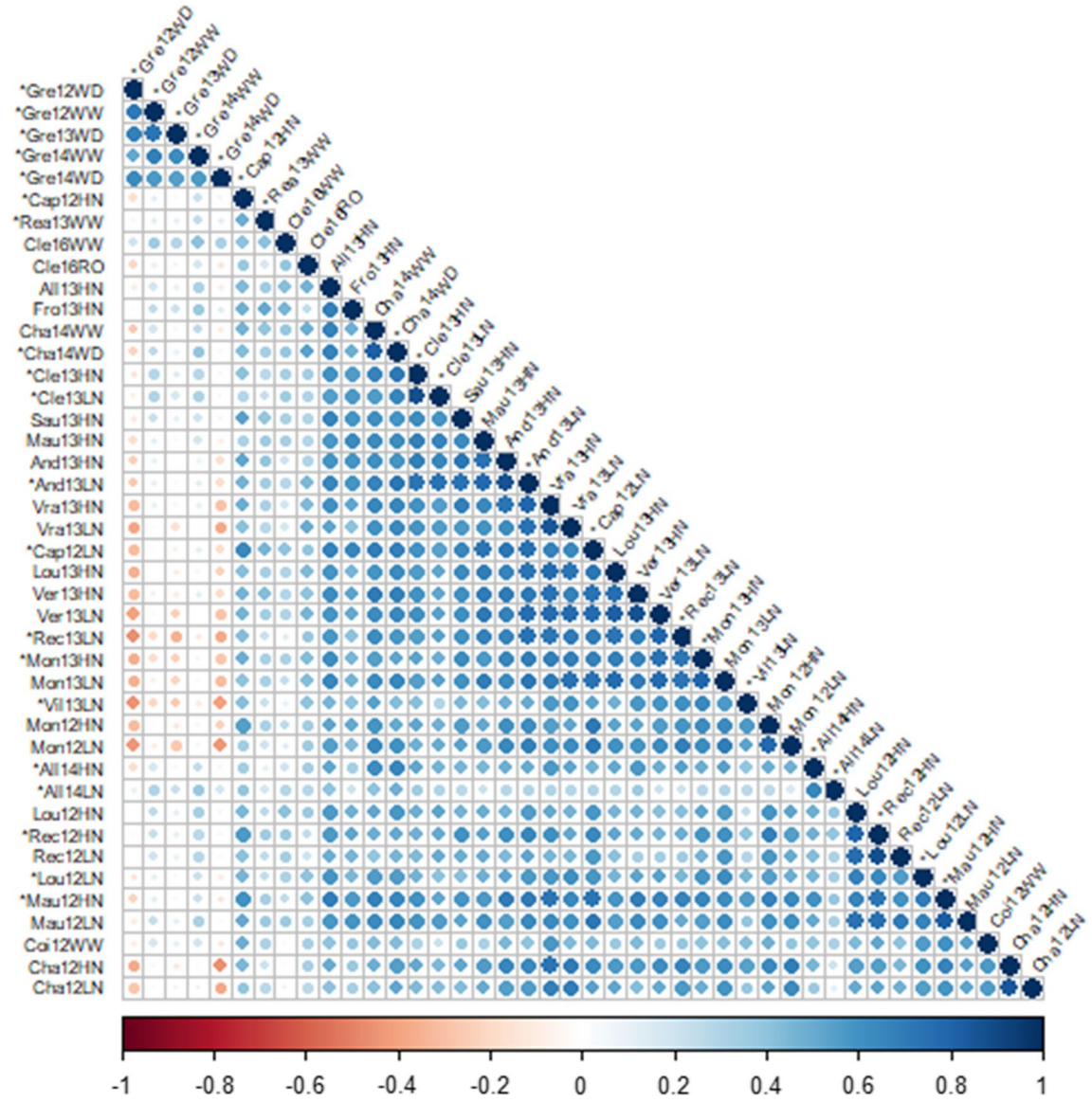


Fig. 1 Correlation matrix for the adjusted means of GY in the MET. Size and colours of the points are function of the pairwise correlations. The environments selected in the subset are indicated with a star in front of their name. Environments are sorted according to their level of correlation

worse in CVnewE and CVnewGE. The selection of subsets of EC (EG_$G \times W_{sel}$, EG_$G \times W_{selpred}$) did not result in any significant increase in prediction accuracy. The model including the AMMI decomposition (EG_$(G \times W)_{AMMI}$) on the dataset (including the validation set) was much better than all other models for each scenario. Most of this gain was lost when the AMMI decomposition was based on the training set only (EG_$(G \times W)_{AMMIpred}$), but it remained as good or better than the other models. Model EG_$G_{AMMI} \times W_{all}$ performed better than EG_$G \times W_{all}$, which confirms the interest of using a different kinship for the main effect and the interaction effect.

Discussion

Interactions between varieties and environments are common in crop MET (Chenu 2015). This $G\times E$ variation can be seen as an additional source of variability that could be used to increase the performance of varieties adapted to specific environmental scenarios (Tardieu et al. 2012). Taking this variability into account in breeding is difficult because there are infinite combinations of stress that can occur, even if some of them are more frequent than others. The important number of potential selection candidates also prevent evaluating all of them in MET. For these reasons, predicting $G\times E$ to identify optimal combinations of genes in specific

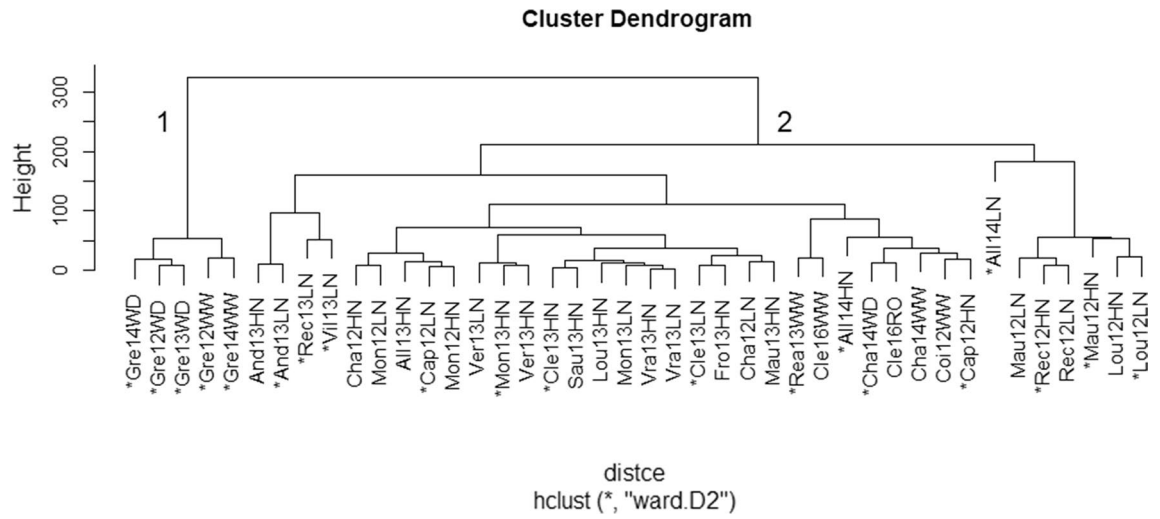


Fig. 2 Dendogram of the hierarchical cluster analysis (Ward 1963) on the interaction term of the AMMI decomposition. The environments selected in the subset are indicated with a star in front of their name. The two main clusters are numbered 1 and 2

Table 2 Variance components for GY in the full dataset and the subset composed of 20 environments and 156 varieties

Components	Full dataset	Subset
E	180.9 (40.0)	161.6 (52.5)
G	29.9 (2.9)	28.6 (3.5)
G× *E*	23.1 (0.5)	35.5 (1.2)

The estimations of the environmental (*E*), genetic (*G*) and genotype × environment interaction (*G*× *E)* variances are indicated, together with the standard errors in bracket

environments seems relevant and promising (Jarquin et al. 2014; Heslot et al. 2014; Millet et al. 2016). We compared the prediction accuracies of several reference models using ECs and molecular markers to estimate covariance structures for interaction terms. In addition, we proposed to improve these approaches by using more adapted ECs thanks to crop growth models, by selecting a subset of relevant ECs and by modelling *G*× *E* with AMMI decomposition.

A multi-environment trial with an intermediate interaction level

We used a large dataset composed of 42 environments with contrasting management and for a diversity panel of 220 European elite varieties of winter bread wheat (Ly et al. 2017). However, despite these environmental and genetic diversities, the amount of *G*× *E* was not as high as expected (Table 2, Figs. 1, 2), and of the same magnitude as the *G* variance. One explanation is that most environments were located in the Northern part of France in years 2012 and 2013. The *G*× *E* is clearly structured by the different environmental conditions that occurred in the South of France

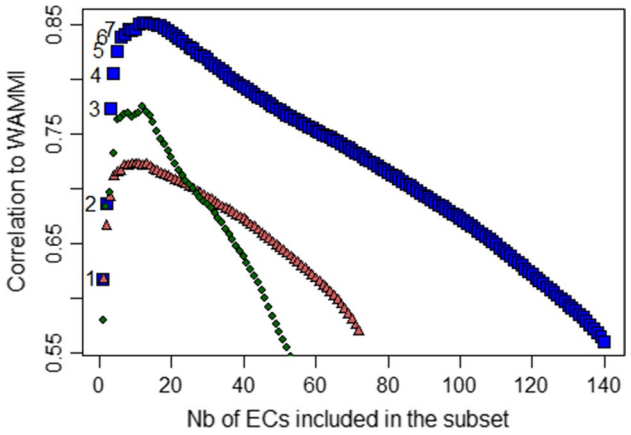


Fig. 3 Correlation between the environment covariance matrix obtained from the AMMI decomposition on the full dataset (W_{AMMI}) and covariance matrices obtained with different subsets of EC (W_{sub}) against the number of environmental covariates (ECs) included in the subset. ECs were added sequentially in order to get the maximal correlation with W_{AMMI}. The selection was based only on climatic ECs (red triangles), or only on integrated ECs derived from crop growth model SiriusQuality (green dots), or on both climatic and integrated ECs (blue squares). The seven first selected ECs were involved in the computation of W_{sel}: 1, rdtmmf2; 2, SQ_SF_Tw_max; 3, SQ_SF_T; 4, Nferti; 5, rdtmw2; 6, nd25ef; 7, SQ_SF_Nem (see Supplementary Table S3 for description of the covariates)

(Gréoux) against all the other environments. The two treatments of a same combination of year and location were similar in terms of interaction for most environments (Fig. 2). One explanation could be that the difference between the two treatments (level of water or nitrogen supply) were not sufficient in some cases, particularly for the N treatments, as revealed by the small differentials for GY between the two N treatments in most experiments. On the opposite, the two

Table 3 Prediction accuracy of the different statistical models in the four cross-validation scenarios for G

Models	Cross-validation scenarios				Average
	CVrandom	CVnewG	CVnewE	CVnewGE	
EG (1)	0.55	0.40	0.56	0.39	0.48
EG_$G \times E$ (2)	0.63	0.52	–	–	–
EG_$(G \times E)_{grp}$ (3)	0.71	0.50	0.72	0.50	0.61
EG_$G \times W_{all}$ (4)	0.74	0.52	0.69	0.45	0.60
EG_$G \times W_{sel}$ (5)	0.73	0.52	0.71	0.48	0.61
EG_$G \times W_{selpred}$ (6)	0.75	0.52	0.61	0.42	0.58
EG_$(G \times W)_{AMMI}$ (7)	0.81	0.72	0.81	0.71	0.76
EG_$G_{AMMI} \times W_{all}$ (8)	0.81	0.71	0.71	0.59	0.71
EG_$(G \times W)_{AMMIpred}$ (9)	0.76	0.55	0.69	0.49	0.62
Average	0.72	0.55	0.69	0.50	0.62

See Table 1 for the description of the models. CVrandom, CVnewG, CVnewE and CVnewGE are cross-validation schemes determined to address the prediction objectives oGoE, nGoE, oGnE and nGnE

water treatments of a same environment led to important interactions, probably because the treatments were more contrasting. This is particularly the case for the rainout shelter experiment (Cle16RO and Cle16WW).

Selecting relevant environmental covariates

Using all the predetermined ECs as proposed by Jarquin et al. 2014 allowed capturing an important part of the $G \times E$. The resulting environmental covariance (W_{all}) matrix had a correlation of 0.56 with the covariance matrix obtained from the AMMI decomposition (W_{AMMI}) corresponding to the observed $G \times E$. We showed that selecting a subset of relevant covariates could increase this correlation to 0.84, and this high correlation was obtained with only seven ECs. Therefore, it is clear that, like for molecular markers (Usai et al. 2009), a selection of relevant variables can increase the proportion of $G \times E$ captured by the model. The selection (or regularization) of the EC is probably more important than for the molecular markers, because of the important redundancy in the EC matrix. This is because the number of environments is limited and some of the ECs bring similar information.

Another important point is that the use of SiriusQuality to derive integrative stress indices allowed reaching much higher correlation than with climatic ECs only (Fig. 3). These stress indices have the property to take into account the development of the plant, whereas the climatic EC only reflect physical measurements (rain, temperature, radiation, soil texture). This confirms the study of Ly et al. (2017) who showed that crop models were efficient to estimate the stress that the plants experienced. The main advantage of the new stress indices used in this study (DMSI) compared with climate covariates and other indices simulated by crop growth models like the water supply/demand ratio (Chenu et al. 2011) or the N nutrition index (Ly et al. 2017) is that different types of stresses, alone or in combinations, can be quantified and their effects on crop growth can be compared directly. The use of CGM requires precise knowledge on the soil characteristics, crop management and climatic factors, and so an efficient use of CGM relies on systematic and accurate characterization of the environments.

The three more informative ECs are related to the photothermal quotient between meiosis and flowering (rdtmmf2), which reflects the balance between the incident irradiance available for growth and the potential growth of sinks driven by temperature, and temperatures above the optimal (SQ_SF_Tw_max, SQ_SF_T). This result was expected as the strong structure of the environments in two groups is related to a temperature gradient (Mediterranean versus oceanic climate). The four other ECs are related to N deficit (Nferti and SQ_SF_Nem), photothermal quotient during the winter period (rdtmw2) and the number of days with maximum daily temperature > 25 °C between heading and flowering (nd25ef). This could also be expected as there were numerous environments with limited N fertilization. It is more surprising that no water stress EC was selected, but this may be because temperature and water stress often occur simultaneously. In any case, the interpretation of the ECs selected by the algorithm should be considered with caution and validated with other datasets, as there are strong correlations between ECs.

An AMMI decomposition of G × E observed in the training set to estimate covariance matrices can increase prediction accuracy

As expected and found in other studies Jarquin et al. (2014), Ly et al. (2017, 2018), the scenario CVrandom resulted in the highest prediction accuracies. It appeared

more difficult to predict new varieties than new environments (Table 3).

Relatively high prediction accuracies were obtained with the main effect model (EG) confirming that the $G\times E$ variance was limited in this dataset. However, adding an interaction term (EG_G×E) potentially increased prediction accuracy, particularly when the grouping of the environments in two clusters (EG_$(G\times E)_{grp}$) was taken into account. It was difficult to get higher prediction accuracies than this reference model, because $G\times E$ variance is mostly due to the structure in two groups of environments. Using ECs to estimate covariance between environments resulted in slightly higher accuracies in scenario CVrandom and CVnewG, but lower in scenario CVnewE and CVnewGE (EG_$G\times W_{all}$, derived from Jarquin et al. 2014). Apparently using all the ECs was a poor strategy to estimate covariances. But selecting a subset of covariates (EG_$G\times W_{sel}$ and EG_$G\times W_{selpred}$) did not improve prediction accuracies, despite the fact that W_{sel} better explained $G\times E$ than W_{all} (Fig. 3).

Using the AMMI decomposition of the phenotypic data to estimate observed covariances between varieties and between environments seems to be very efficient. The covariance matrices obtained with an AMMI with only three axes (EG_$(G\times W)_{AMMI}$) resulted in much higher prediction accuracies than all the other models and for all scenarios. The accuracies were lower when the AMMI was only based on the training set and the genetic and environmental scores of the missing varieties and environments were predicted with the markers and the EC (EG_$G\times W_{AMMIpred}$). But they remained as good or better than all the other models. Considering these results and the fact that the AMMI was done with only three axes, and so is expected to capture only consistent $G\times E$ and no residual variance, we think that this strategy is very promising. It illustrates that it should be possible to reach much higher prediction accuracies if covariance matrices were better estimated than in reference models in which markers and environmental covariates are used without considering the observed $G\times E$. The lower accuracy obtained with EG_$G\times W_{AMMIpred}$ was probably due to the reduced size and interactivity found in this dataset. It would be interesting to test this model and compare it to the other models using datasets with more contrasting environments. We first plan to explore the potential of these models using crop growth model simulations with virtual varieties, to generate datasets with known properties (level of $G\times E$) and sufficient size.

It is also interesting to note that EG_$G_{AMMI}\times W_{all}$ performed better than EG_$G\times W_{all}$, which can be due to the fact that the full dataset was used to estimate the kinship matrix for the interaction term of the first model, but we can also suppose that using a different kinship matrix for the main genetic effect and the interaction effect could be helpful. We can indeed expect that different QTLs explain the two types of effects.

One important limit to our work is that in all models, we considered that all varieties are submitted to the same environmental conditions in a given environment. This is not realistic, in particular when there is a variability in earliness which allows some of the varieties to escape stress. One way to deal with this would be to compute EC specifically for each variety or for groups of similar varieties, as in Jarquin et al. (2014) and Heslot et al. (2014). We tested this strategy on our datasets, but it resulted in lower prediction accuracies (data not shown), which may be because the prediction of the developmental stages was not sufficiently accurate. We believe that a precise characterization of the environments composing the MET is a key element to get accurate $G\times E$ predictions. For instance, a precise characterization of the soil structure and depth is essential to estimate relevant ECs. This information is also essential to get reliable outputs from the crop models. In future works, a special focus should be given to collect precise and uniform descriptions of the trials. This would be essential to run efficient prediction and association mapping studies taking $G\times E$ into account.

Acknowledgements Management of the wheat multi-environment trials was financially supported by the French National Research National Agency under Investment for the Future (BreedWheat Project ANR-10-BTBR-03) and by FranceAgriMer. The Phéno3C platform was financially funded by the French National Research National Agency under the Investment for the Future Phenome project (ANR-11-INBS-12) and by the European Regional Development Fund (AV0011535). This publication has been written with the support of the AgreenSkills + fellowship programme which has received funding from the EU's Seventh Framework Programme under Grant Agreement No. FP7- 609398 (AgreenSkills + contract). PM was also supported by the EU project H2020 SolACE (Grant Agreement No. 727247).

Compliance with ethical standards

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

Ethical standards The authors declare that the experiments comply with the current laws of the countries in which the experiments were performed.

References

Asseng S, Martre P, Maiorano A et al (2019) Climate change impact and adaptation for wheat protein. Global Change Biol 25:155–173

Brancourt-Hulmel M (1999) Crop diagnosis and probe genotypes for interpreting genotype environment interaction in winter wheat trials. Theor Appl Genet 99:1018–1030. https://doi.org/10.1007/s001220051410

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

Brisson N, Gate P, Gouache D et al (2010) Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Res 119:201–212. https://doi.org/10.1016/j.fcr.2010.07.012

Burgueño J, de los Campos G, Weigel K, Crossa J, (2012) Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci 52:707. https://doi.org/10.2135/cropsci2011.06.0299

Chenu K (2015) Characterizing the crop environment: nature, significance and applications. Academic Press Ltd., London

Chenu K, Cooper M, Hammer GL, Mathews KL, Dreccer MF, Chapman SC (2011) Environment characterization as an aid to wheat improvement: interpreting genotype–environment interactions by modelling water-deficit patterns in North-Eastern Australia. J Exp Bot 62:1743–1755

Crossa J, de los Campos G, Maccaferri M, et al (2016) Extending the marker × environment interaction model for genomic-enabled prediction and genome-wide association analysis in durum wheat. Crop Sci 56:2193. https://doi.org/10.2135/cropsci2015.04.0260

Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. J Agric Biol Environ Stat 11:381–393. https://doi.org/10.1198/108571106X154443

Denis JB (1980) Analyse de la régression factorielle. Biom Praxim 20:1–34

IPCC Fifth Assessment Report. Climate change 2013: The physical science basis.

Gilmour AR, Gogel B, Cullis BR, Thompson R (2009) ASREML user guide release 3.0. VSN International, Hemel Hempstead

He J, Le Gouis J, Stratonovitch P et al (2012) Simulation of environmental and genotypic variations of final leaf number and anthesis date for wheat. Eur J Agron 42:22–33

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. Theor Appl Genet 127:463–480. https://doi.org/10.1007/s00122-013-2231-5

Jamieson PD, Semenov MA, Brooking IR, Francis GS (1998) Sirius: a mechanistic model of wheat response to environmental variation. Eur J Agron 8:161–179. https://doi.org/10.1016/S1161-0301(98)00020-3

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

Justes E, Mary B, Meynard J et al (1994) Determination of a critical nitrogen dilution curve for winter-wheat crops. Ann Bot 74:397–407. https://doi.org/10.1006/anbo.1994.1133

Lecomte C (2005) Experimental evaluation of varietal innovations. Proposition of genotype—environment analysis tools adapted to the diversity of needs and constraints of the professionals of the seeds industry. Diss AgroParisTech, p 262

Lopez-Cruz M, Crossa J, Bonnett D, et al (2015) Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. G3 Genes Genomes Genet 5:569–582.https://doi.org/10.1534/g3.114.016097

Ly D, Chenu K, Gauffreteau A et al (2017) Nitrogen nutrition index predicted by a crop model improves the genomic prediction of grain number for a bread wheat core collection. Field Crops Res 214:331–340. https://doi.org/10.1016/j.fcr.2017.09.024

Ly D, Huet S, Gauffreteau A et al (2018) Whole-genome prediction of reaction norms to environmental stress in bread wheat (*Triticum aestivum* L.) by genomic random regression. Field Crops Res 216:32–41. https://doi.org/10.1016/j.fcr.2017.08.020

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

Martre P, Dambreville A (2018) A model of leaf coordination to scale-up leaf expansion from the organ to the canopy. Plant Physiol 176:704–716

Martre P, Jamieson PD, Semenov MA et al (2006) Modelling protein content and composition in relation to crop nitrogen dynamics for wheat. Eur J Agron 25:138–154. https://doi.org/10.1016/j.eja.2006.04.007

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Millet EJ, Welcker C, Kruijer W et al (2016) Genome-wide analysis of yield in Europe: allelic effects vary with drought and heat scenarios. Plant Physiol 172(2):749–764

Parent B, Leclere M, Lacube S, Semenov MA, Welcker C, Martre P, Tardieu F (2018) Maize yields over Europe may increase in spite of climate change, with an appropriate use of the genetic variability of flowering time. Proc Nat Acad Sci 115:10642–10647

Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. https://doi.org/10.1086/519795

Rimbert H, Darrier B, Navarro J et al (2018) High throughput SNP discovery and genotyping in hexaploid wheat. PLoS ONE 13:e0186329. https://doi.org/10.1371/journal.pone.0186329

Rincent R, Kuhn E, Monod H et al (2017) Optimization of multi-environment trials for genomic selection based on crop models. Theor Appl Genet 130:1735–1752. https://doi.org/10.1007/s00122-017-2922-4

Rincent R, Charpentier JP, Faivre-Rampant P, Paux E, Le Gouis J, Bastien C, Segura V (2018) Phenomic selection is a low-cost and high-throughput method based on indirect predictions: proof of concept on wheat and poplar. G3 Genes Genomes Genet 8:3961–3972. https://doi.org/10.1534/g3.118.200760

Rodrigues PC, Malosetti M, Gauch HG, van Eeuwijk FA (2014) A weighted AMMI algorithm to study genotype-by-environment interaction and QTL-by-environment interaction. Crop Sci 54:1555–1570

Schulz-Streeck T, Ogutu JO, Gordillo A et al (2013) Genomic selection allowing for marker-by-environment interaction. Plant Breed 132:532–538. https://doi.org/10.1111/pbr.12105

Smith AB, Cullis BR, Gilmour AR (2001) The analysis of crop variety evaluation data in Australia. Aust N Z J Stat 43:129–1450

Tardieu F (2012) Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. J Exp Bot 63:25–31

Touzy G, Rincent R, Bogard M et al. (2019) Using environmental clustering to identify specific drought tolerance QTLs in bread wheat (*T. aestivum* L). Theor Appl Genet. https://doi.org/10.1007/s00122-019-03393-2

Usai MG, Goddard ME, Hayes BJ (2009) LASSO with cross-validation for genomic selection. Genet Res 91:427. https://doi.org/10.1017/S0016672309990334

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423. https://doi.org/10.3168/jds.2007-0980

Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244

Xose Rodriguez-Alvarez M, Boer MP, van Eeuwijk FA, Eilers PHC (2018) Correcting for spatial heterogeneity in plant breeding experiments with P-splines. Spat Stat 23:52–71. https://doi.org/10.1016/j.spasta.2017.10.003

Zadoks JC, Chang TT (1974) Konzak CF (1974) A decimal code for the growth stages of cereals. Weed Res 14:415–421

Supplementary material

Supplementary Table S1: Description of the 42 environments composing the multi-environment trial.

Name[a]	Subset[b]	Location	Altitude (m)	Longitude (°)	Latitude (°)	Treatment	Sowing date	N fertilizer (kgN ha^{-1})	Cumulative precipitation + irrigation (mm)		Average temperature (°C)		Maximum temperature (°C)		Generalized heritability	
									from sowing to flowering[c]	from flowering to maturity[c]	from sowing to flowering[c]	from flowering to maturity[c]	from sowing to flowering[c]	from flowering to maturity[c]	DOE[d]	GY
Cap12HN	no	Cappelle	150	3.1	50.3	HN	10/28/2011	228	406	155	7.6	15.9	28.1	30.2	0.99	0.91
Cap12LN	no	Cappelle	150	3.1	50.3	LN	10/28/2011	130	406	155	7.6	15.9	28.1	30.2	0.99	0.92
Cha12HN	no	Chalons	102	4.4	48.95	HN	10/22/2011	200	336.3	122	7.8	16.9	28.8	30.4	0.93	0.89
Cha12LN	no	Chalons	102	4.4	48.95	LN	10/22/2011	100	336.3	122	7.7	16.9	28.8	30.4	0.91	0.9
Coi12WW	no	Coings	166	1.75	46.86	HN	10/17/2011	190	467	88	8	17.4	28.4	32.0	0.99	0.92
Gre12WW	yes	Gréoux	300	5.86	43.75	Irrigated	11/17/2011	210	282	104	8.5	20.0	32.3	34.6	0.98	0.89
Gre12WD	yes	Gréoux	300	5.86	43.75	Dry	11/17/2011	210	247	37	8.4	20.0	32.3	34.6	0.96	0.65
Lou12HN	yes	Louville	144	1.82	48.19	HN	10/17/2011	150	364	123	7.9	16.4	27.6	31.6	0.97	0.93
Lou12LN	yes	Louville	144	1.82	48.19	LN	10/17/2011	50	337.6	97	7.9	16.4	27.6	31.6	0.98	0.92
Mau12HN	yes	Maule	54	1.85	48.91	HN	10/14/2011	180	403	115	8.5	17.4	28.5	32.5	0.98	0.95
Mau12LN	no	Maule	54	1.85	48.91	LN	10/14/2011	80	403	115	8.5	17.4	28.5	32.5	0.98	0.93
Mon12HN	no	Mons	85	3.01	49.88	HN	10/20/2011	200	359	217	7.5	15.4	27.2	28.7	0.97	0.81
Mon12LN	no	Mons	85	3.01	49.88	LN	10/20/2011	100	374	224	7.5	15.4	27.2	28.7	0.96	0.83
Rec12HN	yes	Réclainville	150	1.75	48.34	HN	10/28/2011	150	351	127	7.9	16.3	27.6	31.6	NA	0.91
Rec12LN	yes	Réclainville	150	1.75	48.34	LN	10/28/2011	50	351	127	7.9	16.3	27.6	31.6	NA	0.9
All13HN	no	Allonnes	150	1.65	48.34	HN	11/22/2012	201	372	29	7.2	19.0	27.9	31.3	0.98	0.94
And13HN	yes	Andelu	97	1.69	48.86	HN	10/27/2012	240	388	45	7	18.4	27	30.3	0.99	0.89
And13LN	yes	Andelu	97	1.69	48.86	LN	10/27/2012	160	388	45	6.9	18.1	27	30.3	0.99	0.9
Cle13HN	no	Clermont-Ferrand	331	3.15	45.46	HN	######	250	264	31	7	18.3	28.5	31.7	0.98	0.86
Cle13LN	no	Clermont-Ferrand	331	3.15	45.46	LN	######	175	264	31	7	18.3	28.5	31.7	0.88	0.9
Fro13HN	no	Froissy	160	2.22	49.57	HN	10/22/2012	180	372	55	6.8	18.2	28.3	32.3	0.99	0.69
Gre13WD	yes	Gréoux	300	5.86	43.75	Dry	######	225	522	2	7.9	18.8	26.4	33.5	0.96	0.6
Lou13HN	no	Louville	144	1.82	48.19	HN	######	200	391	35	7.1	18.1	27.9	30.5	0.99	0.87
Mau13HN	no	Maule	54	1.85	48.91	HN	10/25/2012	170	397	42	7	18.0	26.4	30.1	0.98	0.96
Mon13HN	no	Mons	85	3.01	49.88	HN	10/24/2012	200	348	71	6.2	17.7	26.2	31.7	0.98	0.85
Mon13LN	no	Mons	85	3.01	49.88	LN	10/24/2012	100	352	71	6.4	18.1	28.9	32.7	0.99	0.86
Rea13WW	yes	Réalville	100	1.5	44.14	Irrigated	11/14/2012	184	554	137	8.5	17.7	28.5	31.7	0.99	0.93
Rec13LN	yes	Réclainville	150	1.75	48.34	LN	10/28/2012	100	396	35	7.1	18.1	27.9	30.5	0.99	0.88
Sau13HN	no	Saultain	70	3.58	50.34	HN	10/27/2012	185	447	47	6.6	18.2	26.7	31.5	0.86	0.5
Ver13HN	no	Verneuil	91	2.5	48.38	HN	10/29/2012	200	495	74.4	6.7	18.1	28.9	30.1	0.98	0.92
Ver13LN	no	Verneuil	91	2.5	48.38	LN	10/29/2012	100	495	74.4	6.7	18.1	28.9	30.1	0.96	0.89
Vil13LN	yes	Villier	150	2.13	48.73	LN	10/23/2012	190	456	62	6.9	17.9	26.7	29.7	0.95	0.45
Vra13HN	no	Vraux	85	4.24	49.03	HN	10/23/2012	210	375	62	6.4	17.9	27.2	30.1	0.87	0.93
Vra13LN	no	Vraux	85	4.24	49.03	LN	10/23/2012	110	375	62	6.4	17.8	27.2	30.1	0.95	0.84
All14HN	yes	Allonnes	150	1.65	48.34	HN	10/16/2013	157	437	88	8.4	17.0	25.1	28.5	0.97	0.89
All14LN	yes	Allonnes	150	1.65	48.34	LN	10/16/2013	51	437	88	8.4	17.0	25.1	28.5	0.7	0.85
Cha14WW	no	Champigny	135	0.16	46.71	Irrigated	10/22/2013	180	601	84	8.6	16.9	27.7	30.7	0.99	0.90
Cha14WD	no	Champigny	135	0.16	46.71	Dry	10/22/2013	180	566	84	8.6	16.9	27.7	30.7	0.99	0.90
Gre14WW	yes	Gréoux	300	5.86	43.75	Irrigated	######	260	604	180	8.6	18.8	28	35.7	0.99	0.75
Gre14WD	yes	Gréoux	300	5.86	43.75	Dry	######	260	468	82	8.6	18.6	28	35.7	0.98	0.44
Cle16WW	yes	Clermont-Ferrand	331	3.15	45.46	Irrigated	######	70	306	100	8.5	17.7	29.7	33.4	0.94	0.78
Cle16RO	no	Clermont-Ferrand	331	3.15	45.46	Dry (shelters)	######	140	111	67	8.3	17.3	29.7	33.4	0.94	0.75

[a] HN, high N ; LN, low N ; WW, well-watered ; WD, water deficit ; RO, rainout shelter.

[b] Subset indicates if the environment was selected in the subset of 20.

[c] Flowering day predicted as Heading + 115 degree days and maturity predicted by the crop model

[d] NA, not available.

Supplementary Table S2: summary of the statistical models used to compute adjusted means and estimate heritabilities.

Environment	*Touzy et al. 2019* [a]	*R-package* [b]	*Statistical model (heritabilities)*[d]	*Statistical model (adj. means)*	*Nb full Rep.* [c]
All14HN, All14LN	*All14RF*	SpATS	$Y_{ijk} = \mu + \boldsymbol{S}_k + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ijk} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	2
And13HN, And13LN	*And13RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Cap12HN, Cap12LN	*Cap12RF*	SpATS	$Y_{ijk} = \mu + \boldsymbol{S}_k + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ijk} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	2
Cha12HN, Cha12LN	*Cha12RF*	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	1
Cle13HN, Cle13LN	*Cle13RF*	SpATS	$Y_{ik} = \mu + \boldsymbol{S}_k + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	2
Cle16WW, Cle16RO	*Cle16RF*	ASREML-R	$Y_{ibk} = \mu + B_b + \boldsymbol{S}_{bk(check)} + \mathbf{G}_i + \epsilon_{ibk}$	$Y_{ibk} = \mu + B_b + \boldsymbol{S}_{bk} + G_i + \epsilon_{ibk}$	1
Gre12WW, Gre12WD	*Gre12IR*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Gre13WD	*Gre13RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Gre14WW, Gre14WD	*Gre14IR*	SpATS	$Y_{ijk} = \mu + R_j + \alpha.RU_{ijk} + \boldsymbol{S}_{jk} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \alpha.RU_{ijk} + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Lou12HN, Lou12LN	*Lou12RF*	SpATS	$Y_{ik} = \mu + \boldsymbol{S}_k + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	2
Mau12HN, Mau12LN	*Mau12RF*	SpATS	$Y_{ik} = \mu + \boldsymbol{S}_k + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	2
Mon12HN, Mon12LN	*Mon12RF*	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_k + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_k + G_i + \epsilon_{ik}$	1,5
Mon13HN, Mon13LN	*Mon13RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Rea13WW	*Rea13RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	4
Rec12HN, Rec12LN	*Rec12RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Ver13HN, Ver13LN	*Ver13RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Vra13HN, Vra13LN	*Vra13RF*	ASREML-R	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Cha14WW, Cha14WD	*Cha14RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	2
Coi12WW	*Coi12RF*	SpATS	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk(check)} + \mathbf{G}_i + \epsilon_{ijk}$	$Y_{ijk} = \mu + R_j + \boldsymbol{S}_{jk} + G_i + \epsilon_{ijk}$	3
All13HN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1
Fro13HN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1
Lou13HN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1
Mau13HN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1
Rec13LN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1
Sau13HN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1
Vil13LN	-	ASREML-R	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + \mathbf{G}_i + \epsilon_{ik}$	$Y_{ik} = \mu + \boldsymbol{S}_{k(check)} + G_i + \epsilon_{ik}$	1

[a] Equivalence with the trial names used in Touzy et al. (2019),

[b] R package used to fit the models

[c] Number of full replicates of all varieties (note that check varieties were always replicated in each block.

[d] G_i is the varietal effect, R_j is replicate effect, S_{jk} and $S_{jk(checks)}$ are the block effects and the block effects estimated only with the checks, respectively. B_b is the rain-out shleter effect. α is the effect of the plot water capacity RU_{ijk}. ϵ_{ijk} are the residuals assumed to be independent and normally distributed. Random effects are indicated with bold letters.

For Gre14WD and Gre14WW a quantitative covariate representing the available water capacity for each plot (RU) was introduced in the model, and permitted to significantly increase heritabilities. This covariate, strongly related to local soil depth, was only available for these two environments.

Supplementary Table S3. Description of the 74 climatic covariates.

Type / name	Description[a]	Min[b]	Max[b]
Temperature			
stmpw	Sum of T_{mean} from sowing to beginning of stem elongation	773	990
stmpem	Sum of T_{mean} from beginning of stem elongation to meiosis	360	744
stmpmf	Sum of T_{mean} from meiosis to flowering	262	345
stmpef	Sum of T_{mean} from beginning of stem elongation to flowering	617	1042
stmpfh	Sum of T_{mean} from flowering to end of endosperm cell division	72	212
stmphm	Sum of T_{mean} from end of endosperm cell division to maturity	449	601
stmpfm	Sum of T_{mean} from flowering to maturity	634	662
mtmpw	Average T_{mean} from sowing to beginning of stem elongation	5	7
mtmpem	Average T_{mean} from beginning of stem elongation to meiosis	9	12
mtmpmf	Average T_{mean} from meiosis to flowering	12	16
mtmpef	Average T_{mean} from beginning of stem elongation to flowering	10	13
mtmpfh	Average T_{mean} from flowering to end of endosperm cell division	13	19
mtmphm	Average T_{mean} from end of endosperm cell division to maturity	16	20
mtmpfm	Average T_{mean} from flowering to maturity	16	20
Solar radiation			
sradw	Sum of R_g from sowing to beginning of stem elongation	49483	114830
sradem	Sum of R_g from beginning of stem elongation to meiosis	49150	102456
sradmf	Sum of R_g from meiosis to flowering	30080	53454
sradef	Sum of R_g from beginning of stem elongation to flowering	84474	146670
sradfh	Sum of R_g from flowering to end of endosperm cell division	7710	23307
sradhm	Sum of R_g from end of endosperm cell division to maturity	43435	77342
sradfm	Sum of R_g from flowering to maturity	63145	92406
mradw	Average R_g from sowing to beginning of stem elongation	384	957
mradem	Average R_g from beginning of stem elongation to meiosis	1221	1853
mradmf	Average R_g from meiosis to flowering	1583	2362
mradef	Average R_g from beginning of stem elongation to flowering	1336	1962
mradfh	Average R_g from flowering to end of endosperm cell division	1454	2619
mradhm	Average R_g from end of endosperm cell division to maturity	1568	1687
mradfm	Average R_g from flowering to maturity	1540	2669
sradmg	Sum of R_g from 100°Cd before meiosis to flowering	46264	64379
sradmm	Sum of R_g from 5 d before to 5 d after meiosis	12116	27391
ndi10m	Number of days when R_g < 10.45 MJ/m2 from meiosis minus 5 d to meiosis plus 5 d	0	5
sri10m	Sum of R_g < 10.45 MJ/m2 from meiosis minus 5 d to meiosis plus 5 d	0	3660
ndi10m2	Number of days when R_g < 10.45 MJ/m2 from flowering to maturity	0	7
sri10m2	Sum of Rg < 10.45 MJ/m2 from flowering to maturity	0	5770
Photothermal quotient			
rdtmw	PTQ from sowing to beginning of stem elongation (sradw / stmpw)	56	148
rdtmem	PTQ from beginning of stem elongation to meiosis (sradem / stmpem)	125	164
rdtmmf	PTQ from meiosis to flowering (sradmf / stmpmf)	103	185
rdtmef	PTQ from beginning of stem elongation to flowering (sradef / stmpef)	119	158
rdtmfh	PTQ from flowering to end of endosperm cell division (sradfh / stmpfh)	87	158
rdtmhm	PTQ from end of endosperm cell division to maturity (sradhm / stmphm)	95	135
rdtmfm	PTQ from flowering to maturity (sradfm / stmpfm)	97	141

rdtmw2	PTQ_{mean} from sowing to beginning of stem elongation (mradw / stmpw)	0.41	1.23
rdtmem2	PTQ_{mean} from beginning of stem elongation to meiosis (mradem / stmpem)	1.7	4
rdtmmf2	PTQ_{mean} from meiosis to flowering (mradmf / stmpmf)	4.9	8.3
rdtmef2	PTQ_{mean} from beginning of stem elongation to flowering (mradef / stmpef)	1.3	2.6
rdtmfh2	PTQ_{mean} from flowering to end of endosperm cell division (mradfh / stmpfh)	7.6	30.9
rdtmhm2	PTQ_{mean} from end of endosperm cell division to maturity (mradhm / stmphm)	3	4.83
rdtmfm2	PTQ_{mean} from flowering to maturity (mradfm / stmpfm)	2.4	4.1
Winter water excess			
watxw	Sum of $\Delta W > 0$ from sowing to beginning of stem elongation	36	399
Water deficit			
spietpw	Cumulative ΔW from sowing to beginning of stem elongation	-108	-17
spietpem	Cumulative ΔW from beginning of stem elongation to meiosis	-134	-64
spietpmf	Cumulative ΔW from meiosis to flowering	-76	-37
spietpef	Cumulative ΔW from beginning of stem elongation to flowering	-196	-117
spietpfh	Cumulative ΔW from flowering to end of endosperm cell division	-41	-8
spietphm	Cumulative ΔW from end of endosperm cell division to maturity	-138	-48
spietpfm	Cumulative ΔW from flowering to maturity	-170	-66
spietp1	Cumulative ΔW from 150°Cd before to 350°Cd after beginning of stem elongation	-69	101
nsddr	Maximum number of successive days with $\Delta W < 0$ between 150°Cd before and 350°Cd after beginning of stem elongation	6	21
ntddr	Number of days with $\Delta W < 0$ between 150°Cd before and 350°Cd after beginning of stem elongation	29	67
Winter frost			
ndefr	Number of days with a $T_{min} < -4$°C between 150°Cd before and 350°Cd after beginning of stem elongation	0	2
sti4	Sum of $T_{min} < -4$°C between 150°Cd before and 350°cd after beginning of stem elongation	-9.4	0
Spring frost			
ndefrs	Number of days with $T_{min} < 0$°C between beginning of stem elongation and flowering	0	8
sti4s	Sum of $T_{mean} < 0$°C between beginning of stem elongation and flowering	-15.3	0
ndt0f	Number of days with $T_{min} < 0$°C from heading to 300°Cd after heading	0	0
st0f	Sum of $T_{min} < 0$°C from heading to 300°Cd after heading	0	0
Heat stress			
nd25m	Number of days with $T_{max} > 25$°C from 5 d before to 5 d after meiosis	0	4
st25m	Sum of $T_{max} > 25$°C from 5 d before to 5 d after meiosis	0	118
nd25ef	Number of days with $T_{max} > 25$°C from heading to flowering	0	6
st25ef	Sum of $T_{max} > 25$°C from heading to flowering	0	164
nd25fh	Number of days with $T_{max} > 25$°C from flowering to end of endosperm cell division	0	6
st25fh	Sum of the $T_{max} > 25$°C from flowering to end of endosperm cell division	0	165
nd25hm	Number of days with $T_{max} > 25$°C from end of endosperm cell division to maturity	4	25
st25hm	Sum of $T_{max} > 25$°C from end of endosperm cell division to maturity	110	727

Nitrogen			
Nferti	Amount of N fertilization (kgN/ha)	50	260

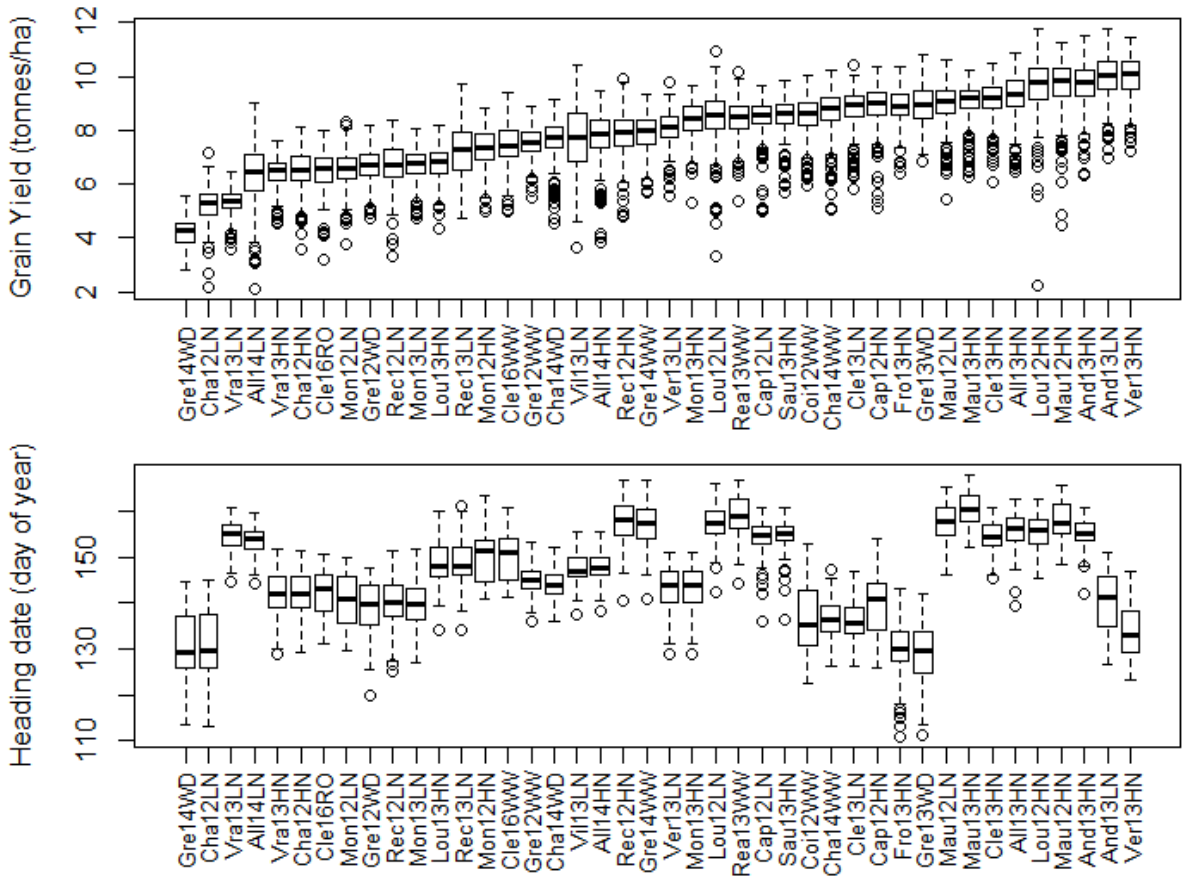
[a] R_g, daily solar radiation; T_{mean}, daily average temperature; T_{max}, daily maximum temperature; T_{min}, daily minimum temperature; **PTQ**, photothermal quotient; PTQ_{mean}, daily average photothermal quotient; ΔW, daily water balance (precipitation + irrigation – Penman evapotranspiration). [b] Minimum and maximum value in the 42 environments.

Supplementary Table S4. Description of the 71 integrated environmental covariates simulated using the wheat crop model *SiriusQuality*.

Type / name	Description [a]	Min[b]	Max[b]
Nitrogen			
SQ_NNIflo	N nutrition index at flowering	0.5	1.1
SQ_PostFloUptake	Post flowering uptake of N	2	55
SQ_AvailableN	Total available N	182	373
SQ_SF_N	Sum of $DMSI_N$ from sowing to maturity	210	271
SQ_SF_Nw	Sum of $DMSI_N$ from sowing to beginning of stem elongation	112	183
SQ_SF_Nem	Sum of $DMSI_N$ from beginning of stem elongation to meiosis	33	80
SQ_SF_Nmf	Sum of $DMSI_N$ from meiosis to flowering	16	21
SQ_SF_Nfh	Sum of $DMSI_N$ from flowering to end of endosperm cell division	6	11
SQ_SF_Nhm	Sum of $DMSI_N$ from end of endosperm cell division to maturity	14	29
SQ_SF_N_max	Maximum of $DMSI_N$ from sowing to maturity	0.14	1
SQ_SF_Nw_max	Maximum of $DMSI_N$ from sowing to beginning of stem elongation	0.98	1
SQ_SF_Nem_max	Maximum of $DMSI_N$ from beginning of stem elongation to meiosis	0.14	1
SQ_SF_Nmf_max	Maximum of $DMSI_N$ from meiosis to flowering	0.86	1
SQ_SF_Nfh_max	Maximum of $DMSI_N$ from flowering to end of endosperm cell division	1	1
SQ_SF_Nhm_max	Maximum of $DMSI_N$ from end of endosperm cell division to maturity	1	1
SQ_SF_Nw_mean	Average of $DMSI_N$ from sowing to beginning of stem elongation	0.99	1
SQ_SF_Nem_mean	Average of $DMSI_N$ from beginning of stem elongation to meiosis	0.92	1
SQ_SF_Nmf_mean	Average of $DMSI_N$ from meiosis to flowering	0.98	1
SQ_SF_Nfh_mean	Average of $DMSI_N$ from flowering to end of endosperm cell division	1	1
SQ_SF_Nhm_mean	Average of $DMSI_N$ from end of endosperm cell division to maturity	1	1
Water deficit			
SQ_SF_W	Sum of $DMSI_W$ from sowing to maturity	182	269
SQ_SF_Ww	Sum of $DMSI_W$ from sowing to beginning of stem elongation	112	183
SQ_SF_Wem	Sum of $DMSI_W$ from beginning of stem elongation to meiosis	31	76
SQ_SF_Wmf	Sum of $DMSI_W$ for water from meiosis to flowering	13	21
SQ_SF_Wfh	Sum of $DMSI_W$ from flowering to end of endosperm cell division	2.9	11
SQ_SF_Whm	Sum of $DMSI_W$ from end of endosperm cell division to maturity	4.7	29
SQ_SF_W_max	Maximum $DMSI_W$ from sowing to maturity	0.1	1
SQ_SF_Ww_max	Maximum $DMSI_W$ from sowing to beginning of stem elongation	0.8	1
SQ_SF_Wem_max	Maximum $DMSI_W$ from beginning of stem elongation to meiosis	0.16	1
SQ_SF_Wmf_max	Maximum $DMSI_W$ from meiosis to flowering	0.97	1
SQ_SF_Wfh_max	Maximum $DMSI_W$ from flowering to end of endosperm cell division	0.28	1
SQ_SF_Whm_max	Maximum $DMSI_W$ from end of endosperm cell division to maturity	0.09	1
SQ_SF_Ww_mean	Average $DMSI_W$ from sowing to beginning of stem elongation	0.99	1
SQ_SF_Wem_mean	Average $DMSI_W$ from beginning of stem elongation to meiosis	0.82	1
SQ_SF_Wmf_mean	Average $DMSI_W$ from meiosis to flowering	0.7	1
SQ_SF_Wfh_mean	Average $DMSI_W$ from flowering to end of endosperm cell division	0.37	1
SQ_SF_Whm_mean	Average $DMSI_W$ from end of endosperm cell division to maturity	0.33	1
Temperature			
SQ_SF_T	Sum of $DMSI_T$ from sowing to maturity	210	270
SQ_SF_Tw	Sum of $DMSI_T$ from sowing to beginning of stem elongation	112	183
SQ_SF_Tem	Sum of $DMSI_T$ from beginning of stem elongation to meiosis	33	80
SQ_SF_Tmf	Sum of $DMSI_T$ from meiosis to flowering	16	21
SQ_SF_Tfh	Sum of $DMSI_T$ from flowering to end of endosperm cell division	6	11
SQ_SF_Thm	Sum of $DMSI_T$ from end of endosperm cell division to maturity	14	29
SQ_SF_T_max	Maximum $DMSI_T$ from sowing to maturity	0.35	0.97

SQ_SF_Tw_max	Maximum $DMSI_T$ from sowing to beginning of stem elongation	0.9	1
SQ_SF_Tem_max	Maximum $DMSI_T$ from beginning of stem elongation to meiosis	0.81	1
SQ_SF_Tmf_max	Maximum $DMSI_T$ from meiosis to flowering	0.97	1
SQ_SF_Tfh_max	Maximum $DMSI_T$ from flowering to end of endosperm cell division	0.86	1
SQ_SF_Thm_max	Maximum $DMSI_T$ from end of endosperm cell division to maturity	0.35	0.97
SQ_SF_Tw_mean	Average $DMSI_T$ from sowing to beginning of stem elongation	1	1
SQ_SF_Tem_mean	Average $DMSI_T$ from beginning of stem elongation to meiosis	1	1
SQ_SF_Tmf_mean	Average $DMSI_T$ from meiosis to flowering	1	1
SQ_SF_Tfh_mean	Average $DMSI_T$ from flowering to end of endosperm cell division	0.98	1
SQ_SF_Thm_mean	Average $DMSI_T$ from end of endosperm cell division to maturity	0.96	1
Water deficit x nitrogen x temperature			
SQ_SF_WNT	Sum of $DMSI_{WNT}$ from sowing to maturity	182	269
SQ_SF_WNTw	Sum of $DMSI_{WNT}$ from sowing to beginning of stem elongation	112	183
SQ_SF_WNTem	Sum of $DMSI_{WNT}$ from beginning of stem elongation to meiosis	31	76
SQ_SF_WNTmf	Sum of $DMSI_{WNT}$ from meiosis to flowering	13	21
SQ_SF_WNTfh	Sum of $DMSI_{WNT}$ from flowering to end of endosperm cell division	3	11
SQ_SF_WNThm	Sum of $DMSI_{WNT}$ from end of endosperm cell division to maturity	4.7	29
SQ_SF_WNT_max	Maximum $DMSI_{WNT}$ from sowing to maturity	0.09	0.92
SQ_SF_WNTw_max	Maximum $DMSI_{WNT}$ from sowing to beginning of stem elongation	0.8	1
SQ_SF_WNTem_max	Maximum $DMSI_{WNT}$ from beginning of stem elongation to meiosis	0.14	1
SQ_SF_WNTmf_max	Maximum $DMSI_{WNT}$ from meiosis to flowering	0.4	1
SQ_SF_WNTfh_max	Maximum of $DMSI_{WNT}$ from flowering to end of endosperm cell division	0.28	1
SQ_SF_WNThm_max	Maximum of $DMSI_{WNT}$ from end of endosperm cell division to maturity	0.09	1
SQ_SF_WNTw_mean	Average of $DMSI_{WNT}$ from sowing to beginning of stem elongation	0.99	1
SQ_SF_WNTem_mean	Average of $DMSI_{WNT}$ from beginning of stem elongation to meiosis	0.82	1
SQ_SF_WNTmf_mean	Average of $DMSI_{WNT}$ from meiosis to flowering	0.7	1
SQ_SF_WNTfh_mean	Average of $DMSI_{WNT}$ from flowering to end of endosperm cell division	0.37	1
SQ_SF_WNThm_mean	Average of $DMSI_{WNT}$ from end of endosperm cell division to maturity	0.34	0.99

[a] **$DMSI_N$**, dry mass stress index for nitrogen deficit; **$DMSI_W$**, dry mass stress index for water deficit; **$DMSI_T$**, dry mass stress index for high temperature; **$DMSI_{WNT}$**, dry mass stress index for water and nitrogen deficit and high temperature.
[b] Minimum and maximum value in the 42 environments.



Supplementary Figure S1: Boxplots of grain yield (tonnes/ha at 0% humidity) and heading date (day of year) in the multi-environment trial composed of 42 environments.