









## ORIGINAL ARTICLE

Special Section: Multi-Omics Prediction in Plant Breeding

# Integration of physiological and remote sensing traits for improved genomic prediction of wheat yield

Guillermo García-Barríos<sup>1,2</sup>  | Carlos A. Robles-Zazueta<sup>2,3</sup>  |  
 Abelardo Montesinos-López<sup>4</sup>  | Osva A. Montesinos-López<sup>5</sup>  | Matthew P. Reynolds<sup>2</sup>  |  
 Susanne Dreisigacker<sup>2</sup>  | José A. Carrillo-Salazar<sup>1</sup> | Liana G. Acevedo-Siaca<sup>2,6</sup>  |  
 Margarita Guerra-Lugo<sup>2</sup> | Gilberto Thompson<sup>2</sup> | José A. Pecina-Martínez<sup>1</sup> |  
 José Crossa<sup>2,7</sup> 

<sup>1</sup>Graduate Program in Genetic Resources and Productivity, Colegio de Postgraduados, Texcoco, Estado de Mexico, Mexico<sup>2</sup>International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Estado de Mexico, Mexico<sup>3</sup>Department of Plant Breeding, Hochschule Geisenheim University, Geisenheim, Germany<sup>4</sup>University Center of Exact Sciences and Engineering, Universidad de Guadalajara, Guadalajara, Jalisco, Mexico<sup>5</sup>Faculty of Telematics, Universidad de Colima, Colima, Colima, Mexico<sup>6</sup>Horticulture and Product Physiology, Wageningen University, Wageningen, The Netherlands<sup>7</sup>Graduate Program in Socioeconomics, Statistics, and Informatics, Colegio de Postgraduados, Texcoco, Estado de Mexico, Mexico**Correspondence**

Carlos A. Robles-Zazueta, Hochschule Geisenheim University, Von-Lade-Straße 1, 65366 Geisenheim, Germany and International Maize and Wheat Improvement Center (CIMMYT), carretera Mexico-Veracruz km 45, Texcoco, Estado de Mexico 56237, Mexico.  
 Email: [carlos.robleszazueta@hs-gm.de](mailto:carlos.robleszazueta@hs-gm.de)

José Crossa, International Maize and Wheat Improvement Center (CIMMYT), carretera Mexico-Veracruz km 45, Texcoco, Estado de Mexico 56237, Mexico.  
 Email: [j.crossa@cgiar.org](mailto:j.crossa@cgiar.org)

**Abstract**

Genomic selection is an extension of marker-assisted selection by leveraging thousands of molecular markers distributed across the genome to capture the maximum possible proportion of the genetic variance underlying complex traits. In this study, genomic prediction models were developed by integrating phenological, physiological, and high-throughput phenotyping traits to predict grain yield in bread wheat (*Triticum aestivum* L.) under three environmental conditions: irrigation, drought stress, and terminal heat stress. Model performance was evaluated using both five-fold cross-validation and leave-one-environment-out (LOEO) schemes. Under five-fold cross-validation, the model incorporating vegetation indices derived from spectral datasets from the grain-filling phase achieved the highest accuracy. In LOEO validation, the model that included days to heading performed best under irrigation, whereas under drought stress, the model utilizing vegetation indices from the

**Abbreviations:** ALA, average leaf angle; C<sub>ab</sub>, chlorophyll a + b content; CI<sub>green</sub>, green chlorophyll index; CI<sub>red-edge</sub>, red-edge chlorophyll index; DTH, days to heading; F<sub>cover</sub>, fraction of vegetation cover; FIPAR, fraction of intercepted photosynthetically active radiation; GE, genotype by environment interaction; GEBV, genomic estimated breeding value; HTP, high-throughput phenotyping; LAI, leaf area index; LAIC<sub>ab</sub>, leaf area index based on chlorophyll content; LOEO, leave-one-environment-out; MCARI, modified chlorophyll absorption ratio index; NDVI, normalized difference vegetation index; NIR, near-infrared; NRMSE, normalized root mean square error; SNP, single nucleotide polymorphism; SpRef, spectral reflectance; SR, simple ratio.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

**Funding information**

International Wheat Yield Partnership; Foundation for Food and Agriculture Research, Grant/Award Number: Heat and Drought Wheat Improvement Consortium (HeDWIC) project; Bill and Melinda Gates Foundation, Grant/Award Numbers: Accelerating Genetic Gains in Maize, Wheat (AGG) project

vegetative stage showed the highest accuracy. Under terminal heat stress, three models performed best: one incorporating genotype by environment interaction, one using vegetation indices during the vegetative stage, and one integrating spectral reflectance data from both the vegetative and grain-filling phases. Although incorporating multiple covariates can improve prediction accuracy or reduce the normalized root mean square error, using an extended model with all available covariates is not recommended due to the marginal predictive accuracy gains, increases in phenotyping, costs and complexity of data collection analysis. Overall, our findings show the importance of tailored phenomic inputs to specific environmental contexts to optimize genomic prediction of wheat yield.

**Plain Language Summary**

Genomic selection is an approach for improving grain yield in plant breeding by applying statistical models that incorporate genome-wide marker data. Genomic selection enables the prediction of genetic merit before phenotypic evaluation, thereby reducing the need for extensive field trials. In this study, we examined whether integrating genomic data with phenological, physiological, and remote sensing traits could enhance the accuracy of wheat grain yield predictions. Genomic prediction models were evaluated under three environmental conditions: irrigation, drought, and terminal heat stress. Vegetation indices and information derived from drone data collected after heading substantially improved our model's prediction accuracies. These findings provide practical insights for breeders seeking to optimize trait selection for identifying high-yielding wheat lines across diverse environments.

**1 | INTRODUCTION**

One of the hardest challenges for our society this century will be to meet the food demands of a growing population. Meeting this demand sustainably is increasingly complicated by harsher climatic conditions, diminishing water resources for irrigation, shrinking arable land, and increased soil degradation (Senapati et al., 2022).

Genetic improvement offers a sustainable solution to this challenge, as its outcomes are permanent and inheritable. Unlike nutritional and phytosanitary interventions, which require continuous inputs, genetic improvements are passed on from one generation to the next, making them a powerful tool to enhance agricultural sustainability (Acquaah, 2007; Mueller & Van Eenennaam, 2022).

Wheat (*Triticum aestivum* L.) breeders develop new cultivars that meet the expectations of farmers, the industry, and consumers in terms of grain yield, yield stability, agronomic traits (e.g., short stature and lodging resistance), tolerance to biotic stresses (e.g., rusts, smuts, powdery mildew, and fusarium head blight), abiotic stresses (e.g., drought resistance and winter hardiness), as well as nutritional aspects and end-use quality. For example, semihard to hard wheat is used

for bread, soft wheat for confectionery products and cookies, while durum wheat is used for semolina products (e.g., couscous) and pasta (Acquaah, 2007; Fradgley et al., 2024; Reynolds & Braun, 2022).

In wheat breeding programs, methods such as pedigree, bulk, single seed descent, and combinations of these methods have been predominantly used to advance generations and develop new cultivars. There is also a growing interest in leveraging biotechnology and bioinformatics to enhance selection gains (Alahmad et al., 2022).

One adopted approach in wheat breeding is genomic selection, which utilizes multiple single nucleotide polymorphisms (SNPs) to predict individuals' genetic merit (Meuwissen et al., 2001). In this method, a subset of individuals that have been both genotyped and phenotyped (the training population) is used to train a model that estimates SNP effects. The model then predicts genomic estimated breeding values (GEBVs) for a testing population that has been only genotyped (Crossa et al., 2017; Mueller & Van Eenennaam, 2022; H. Wang, 2023). Genomic selection can increase the rate of gain in a breeding program by increasing selection intensity, selection accuracy, or reducing selection cycle length (Da Silva et al., 2025; Escamilla et al., 2025; F. Wang et al., 2025).

The accuracy of GEBVs is determined by factors such as marker density, training population size, relatedness between training and testing populations, trait heritability, linkage disequilibrium, statistical methods employed, genotype by environment (GE) interactions, nonadditive effects, and the quality of phenotypic data (Bančić et al., 2023; Esposito et al., 2023; Jung et al., 2024; Montesinos-López et al., 2020; Sandro et al., 2024; Schrauf et al., 2020; Sehgal et al., 2020; Xu et al., 2021).

In this context, high-throughput phenotyping (HTP) methods have gained attention for their ability to provide detailed and noninvasive data on primary and secondary traits throughout the crop cycle (Alemu et al., 2024; Da Silva et al., 2025; Schrauf et al., 2020). HTP platforms often equipped with RGB, multispectral, hyperspectral, or light detection and ranging (LiDAR) sensors; enable the measurement of canopy architecture, photosynthetic efficiency, and biomass accumulation at varying spatial and temporal scales (Araus & Cairns, 2014). By delivering precise phenotypic information on physiological traits such as radiation use efficiency, biomass accumulation, photosynthesis, and chlorophyll fluorescence estimated at leaf and canopy scales (Robles-Zazueta et al., 2021), HTP can be used to refine genomic prediction models for complex traits (Crain et al., 2018; van Eeuwijk et al., 2019; Walter et al., 2019).

HTP has been integrated into genomic prediction models through various approaches. For instance, previous studies have employed relationship matrices derived from hyperspectral data in genomic best linear unbiased prediction (GBLUP) to predict grain yield in wheat. This research approach demonstrated that relationship matrices derived from hyperspectral reflectance collected in canopies in the field, and leaf and grains in the lab could effectively predict yield with accuracies comparable to those obtained using marker or pedigree data (Krause et al., 2019; Rincenc et al., 2018). Another application of spectral reflectance (SpRef) index data was described by Herr et al. (2024) who used these indices in univariate models as covariates and in multivariate models as secondary response variables to predict grain yield in wheat. Similarly, physiological traits such as canopy temperature, chlorophyll content (SPAD), membrane thermostability, senescence rates (stay-green), and normalized difference vegetation index (NDVI) have been used in combination with marker data to implement genomic prediction of yield in wheat. The multi-kernel models integrating physiological and genomic data achieved prediction accuracies 35%–169% higher than models that only used genomic information (Guo et al., 2020).

To complement these statistical approaches, radiative transfer models (RTMs) present a new alternative for breeders to add biological meaning into genomic prediction frameworks as RTMs are mechanistic models that link plant structural and biochemical traits with remote sensing infor-

## Core Ideas

- Incorporating high-throughput phenotyping data as covariates improves the accuracy of genomic prediction models.
- Vegetation indices and PROSAIL-derived traits measured at grain filling substantially increased accuracy of genomic prediction models.
- Genomic prediction under drought stress was less accurate, highlighting challenges in predicting adverse environments.

mation retrieved from satellites, unmanned aerial vehicles, or point-based spectroradiometers (Jacquemoud et al., 2009). In particular, the two most widely used RTMs, SAIL (scattering by arbitrary inclined leaves) and PROSPECT (leaf optics model), have been combined to develop the PROSAIL (PROSPECT + SAIL) model. The inversion of the PROSAIL model allows the user to retrieve physiological traits such as leaf chlorophyll, carotenoid and water content, leaf area index (LAI), average leaf angle (ALA) distribution, the fraction of vegetation cover ( $F_{\text{cover}}$ ), or the fraction of intercepted photosynthetically active radiation (FIPAR) (Duan et al., 2014; Féret et al., 2017). Phenotyping approaches such as vegetation indices, hyperspectral reflectance relationship matrices, the use of thermal imagery, or the use of RTMs such as PROSAIL in synergy with genomics can help to optimize wheat yield predictions. For these reasons, the objectives of this study were as follows: (i) to evaluate the responses of bread wheat to contrasting environmental conditions (irrigated, drought, and heat stress) using phenological, physiological, and HTP traits and (ii) to compare the predictive performance of genomic prediction models that integrate these traits collected during the vegetative and grain-filling stages.

## 2 | MATERIALS AND METHODS

### 2.1 | Study population and field trials

The study population is known as “Molecular Panel” (MOL-PAN) which is comprised of 240 bread wheat genotypes selected from a large and diverse genetic pool of approximately 1500 genotypes (pre-MOLPAN). The selection of MOLPAN lines was done based on its contrasting yield, visual biomass, phenology, and plant height under irrigation and drought stress conditions. Genetic fingerprint information was also taken into consideration to ensure that diversity (e.g., landrace-derived parents and commercial cultivars) was well-represented within the panel. The trials were conducted at the Campo Experimental Norman E. Borlaug (CENEB) at

International Maize and Wheat Improvement Center (CIMMYT), near Ciudad Obregón, Sonora, Mexico (27°20' N, 109°54' W). The panel was grown in three environments: (1) irrigation, (2) drought with only a single irrigation during the crop cycle at the time of sowing, and (3) terminal heat stress achieved through late sowing to expose the plants to high temperatures during the anthesis and grain-filling stages.

Wheat plants were grown under an  $\alpha$ -lattice experimental design, with two replicates per genotype in each environment (480 plots in total). In the irrigation and terminal heat stress environments, each plot consisted of six rows spaced 20 cm apart, with a plot length of 1.5 m, and drip irrigation. In the drought environment, each genotype was established on two raised beds of 2 m per plot. Sowing took place on November 16 and 30, 2022, for the irrigation and drought environments, respectively, and on March 1, 2023, for the terminal heat environment. The harvest dates were May 12, 2023 (irrigation), April 28, 2023 (drought stress), and June 9, 2023 (terminal heat stress). Grain yield was measured following the methodology described by Pietragalla and Pask (2013) and adjusted to grams per square meter.

## 2.2 | Assessment of heading date

Days to heading (DTH) was recorded as the number of days until approximately 50% of the plants or culms in the experimental plot exhibited heading. In this study, heading was defined either as the point when spikes halfway emerged (GS55) or as reaching Zadoks Stage 59, following the criteria established by Zadoks et al. (1974).

## 2.3 | Stomatal conductance and chlorophyll fluorescence measurements

Stomatal conductance ( $g_s$ ) and chlorophyll fluorescence traits were measured using a handheld combined leaf porometer-fluorometer (LI-600PF, Li-COR Biosciences). Measurements were made on the flag leaf during grain filling in the three environments and were conducted from 10:00 a.m. to 1:00 p.m. to minimize the effects from plant circadian rhythm and from measuring during the hottest hours of the day, which can stimulate stomatal closure and affect the measurement of  $g_s$ . Two plants were measured per plot with each measurement lasting ~30 s per leaf. The subsamples were averaged per plot for subsequent data analyses.

From these measurements the quantum yield of photosystem II ( $\Phi_{PSII}$ ) was calculated, which estimates the efficiency at which light absorbed by PSII is used to reduce electron acceptor  $Q_A$  and drive photochemistry. The following equation was used:

$$\Phi_{PSII} = \frac{F_{m'} - F_s}{F_{m'}} \quad (1)$$

where  $F_{m'}$  is the maximum fluorescence yield in a leaf exposed to actinic light and  $F_s$  is the steady-state fluorescence yield of a leaf exposed to actinic light.

For transpiration measurements, the following equation was used with the manufacturer software:

$$\text{Transpiration} = \frac{\mu (W_{\text{sam}} - W_{\text{ref}})}{s} \quad (2)$$

where  $\mu$  is the air flow rate in  $\mu\text{mol s}^{-1}$ ,  $W_{\text{sam}}$  is the water vapor content in an air stream interacting with the measured leaf,  $W_{\text{ref}}$  is the water vapor content in an air stream before interacting with the leaf both measured as  $\text{mmol H}_2\text{O mol}^{-1}$  air, and  $s$  is the leaf area in  $\text{cm}^2$ .

## 2.4 | Multispectral imaging

High-resolution aerial images were collected using a RedEdge-3 multispectral camera (Micasense) mounted on a small airplane flying at an altitude of ~100 m. The ground sampling distance was ~6 cm, and the images were georeferenced using ground control points distributed across the experimental station. Image calibration was performed using the white calibration target provided by the manufacturer (Micasense). Images were collected at least twice during the vegetative (from canopy closure to booting) and grain-filling (from anthesis to milk grain) growth periods in each environment. Subsequently, for each growth period, an average of the spectral bands was calculated.

Image analysis was conducted using photogrammetry software (Pix4D) to stitch images and reconstruct the plots. Reflectance values in the red, green, blue, red-edge, and near-infrared (NIR) regions were extracted using ArcGIS (ESRI). These values were used to calculate vegetation indices (Table 1), including NDVI, modified chlorophyll absorption ratio index (MCARI), red-edge chlorophyll index ( $CI_{\text{red-edge}}$ ), green chlorophyll index ( $CI_{\text{green}}$ ), and simple ratio (SR). Furthermore, the inversion of the PROSAIL model was implemented using the *prospect* R package (Féret & de Boissieu, 2024) to simulate traits such as LAI,  $F_{\text{cover}}$ , ALA, FIPAR, chlorophyll a + b content ( $C_{\text{ab}}$ ), and LAI based on chlorophyll content ( $LAIC_{\text{ab}}$ ) (Jacquemoud et al., 2009).

## 2.5 | Genotypic analysis

Genotyping by sequencing was performed using the Trait-Genetics 25K SNP array, specifically developed for wheat. This array incorporates markers from various sources,

**TABLE 1** Vegetation indices and PROSAIL-derived physiological traits used as covariables to complement genomic prediction models. The formulas for PROSAIL-derived traits are “NA” as their calculations are based on whole spectral ranges as opposed to specific wavelengths.

Abbreviation	Proxy trait	Equation	Reference
<b>Vegetation indices</b>			
CI <sub>green</sub>	Chlorophyll content and LAI	$CI_{green} = \frac{\rho_{NIR}}{\rho_{green}} - 1$	Hunt et al. (2011)
NDVI	Biomass, chlorophyll and water content, LAI, response to drought and heat stress, and yield	$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}$	Jackson & Huete (1991)
MCARI	Chlorophyll content and LAI	$MCARI = [(\rho_{700} - \rho_{670}) - 0.2(\rho_{700} - \rho_{550})] \left( \frac{\rho_{700}}{\rho_{670}} \right)$	Haboudane (2004)
CI <sub>red-edge</sub>	Chlorophyll content, LAI, and biomass	$CI_{red-edge} = \frac{\rho_{NIR}}{\rho_{red-edge}} - 1$	Hunt et al. (2011)
SR	Biomass	$SR = \frac{\rho_{NIR}}{\rho_{red}}$	Xue & Su (2017)
<b>PROSAIL</b>			
ALA	Average leaf angles in the canopy	NA	Féret & de Boissieu (2024)
C <sub>ab</sub>	Chlorophyll a + b content	NA	Féret & de Boissieu (2024)
FIPAR	Fraction of intercepted photosynthetically active radiation	NA	Féret & de Boissieu (2024)
F <sub>cover</sub>	Fraction of vegetation cover	NA	Féret & de Boissieu (2024)
LAI	Leaf area index	NA	Féret & de Boissieu (2024)
LAI <sub>C<sub>ab</sub></sub>	Leaf area index based on chlorophyll content	NA	Féret & de Boissieu (2024)

Abbreviations:  $\rho$ , reflectance; CI, chlorophyll index; LAI, leaf area index; MCARI, modified chlorophyll absorbance reflectance index; NA, not applicable; NDVI, normalized difference vegetation index; NIR, near-infrared; SR, simple ratio.

including the Affymetrix Axiom platform, the Illumina Infinium Wheat 90 and 660K SNP arrays, and other publicly available wheat SNP resources. The dataset underwent multiple filtering steps, excluding SNPs with more than 50% missing data and a minor allele frequency below 0.05. Missing genotypic data were subsequently imputed using the mean genotype value per marker and after filtering 17,263 SNPs remained.

## 2.6 | Statistical analysis

For the traits measured in this study the best linear unbiased estimator (BLUE) was calculated using the R package lme4 (Bates et al., 2015) and implemented through the graphical user interface META-R version 6.04 (Alvarado et al., 2020) as follows:

$$Y_{ijrb} = \mu + E_i + \text{Line}_j + \text{Rep}_r + \text{IB(rep)}_{\text{IB(Rep)}} + \epsilon_{ijrb} \quad (3)$$

where  $Y_{ijrb}$  is the value of the response variable of the fixed effects of line  $j$ th ( $i = 1, \dots, I$ ) in the  $i$ th environment ( $j = 1, \dots, J$ ) measured in the fixed effects of  $r$ th replicate ( $r = 1, 2$ ) and the random effect of the IB(Rep)th incomplete block within the  $r$ th replicate ( $b = 1, 2$ ), assumed to have a mean 0 and variance  $\sigma_{\text{IB(Rep)}}^2$ ,  $\mu$  is the general mean, and  $\epsilon_{ij}$  are independent random errors assumed to be a normal variable with mean 0 and variance  $\sigma_{ijrb}^2$ .

In this study, GBLUP models were used to estimate the breeding values of bread wheat for grain yield using two validation approaches to assess the model accuracy, a five-fold cross-validation (CV) to predict unseen genotypes within the same environment and a leave-one-environment-out (LOEO) approach to evaluate performance in unknown environments to simulate breeding scenarios in novel environmental conditions. In the following subsections, we describe the 12 main models used in our study:

### 2.6.1 | Models E + G + GE and E + G

One of the statistical GBLUP models evaluated in this study, E + G + GE, assumed that the response variable is modeled as follows:

$$Y_{ij} = \mu + E_i + g_j + Eg_{ij} + \epsilon_{ij} \quad (4)$$

where  $Y_{ij}$  is the grain yield adjusted value of line  $j$ th measured in environment  $i$ th,  $\mu$  is the general mean,  $E_i$  is the fixed effect of the environment  $i$ th ( $i = 1, \dots, I$ ),  $g_j$  ( $j = 1, \dots, J$ ) is the random effects of lines,  $Eg_{ij}$  is the random interaction GE effect, and  $\epsilon_{ij}$  are independent random errors assumed to be a normal variable with mean 0 and variance  $\sigma^2$ .

Furthermore, it is assumed that the vector of random effect of lines  $g = (g_1, \dots, g_J)^T$  is distributed according to a multivariate normal distribution

$N_J(0_J, \sigma_g^2 G)$ , while the interaction GE random effects  $Eg = (Eg_{11}, \dots, Eg_{1J}, \dots, Eg_{21}, \dots, Eg_{IJ})^T$  follow a  $N_{IJ}(0_{IJ}, \sigma_{Eg}^2 (I_I \otimes G))$ , with  $0_J$  (mean vector of  $g$ ),  $0_{IJ}$  (mean vector of  $Eg$ ),  $I_I$  and  $\otimes$ , the null vector of size  $J$ , the null vector of size  $IJ$ , the identity matrix of dimensions  $I \times I$  and the Kronecker product, respectively, and  $G$  is the genomic relationship matrix (VanRaden, 2008).

The model (E + G) is the same as the one specified in (4) but without the random interaction effects  $Eg_{ij}$ :

$$Y_{ij} = \mu + E_i + g_j + \epsilon_{ij} \quad (5)$$

### 2.6.2 | Models E + G + GE + DTH and E + G + DTH

Another evaluated statistical model, E + G + GE + DTH, was obtained by adding the covariate DTH to model (4).

$$Y_{ij} = \mu + E_i + g_j + Eg_{ij} + x_{ij}^{(DTH)} \beta_{DTH} + \epsilon_{ij} \quad (6)$$

where all the components are the same as those defined previously for (4), and  $\beta_{DTH}$  is the coefficient effect of the covariate DTH,  $x_{ij}^{(DTH)}$ . The model E + G + DTH is the same as model (6) but without the random interaction  $Eg_{ij}$ .

### 2.6.3 | Models E + G + GE + Phys and E + G + Phys

The model E + G + GE + Phys is similar to model (6), except that now instead of considering the DTH covariate, it incorporates physiological variables such as  $g_s$ , transpiration, and  $\Phi_{PSII}$ .

$$Y_{ij} = \mu + E_i + g_j + Eg_{ij} + \sum_{l=1}^3 x_{ijl}^{(Phys)} \beta_l^{(Phys)} + \epsilon_{ij} \quad (7)$$

where all terms except  $\sum_{l=1}^3 x_{ijl}^{(Phys)} \beta_l^{(Phys)}$  are the same as those defined in model (4), and  $x_{ijl}^{(Phys)}$ ,  $l = 1, 2, 3$ , represents the physiological (Phys) covariates listed earlier and  $\beta_l^{(Phys)}$ ,  $l = 1, 2, 3$  represents its corresponding effects. The model E + G + Phys is the same as model (7) but without the random interaction effects,  $Eg_{ij}$ .

### 2.6.4 | Models E + G + GE + DerivHTP<sub>VegStage</sub> and E + G + DerivHTP<sub>VegStage</sub>

These models are like the previous models, but now the vegetation indices and PROSAIL-derived covariates in the vegetative stages (DerivHTP<sub>VegStage</sub>: CI<sub>red-edge</sub>, NDVI,

LAI\_PROSAIL, FCOVER\_PROSAIL, ALA\_PROSAIL, SR, FIPAR\_PROSAIL, CI<sub>green</sub>, C<sub>ab</sub>\_PROSAIL, LAIC<sub>ab</sub>\_PROSAIL, and MCARI) are incorporated. Because many vegetation indices share the same spectral features (e.g., reflectance in the green, red, NIR spectra) multicollinearity among covariates should be expected. We retained all vegetation indices in the models to simulate a realistic breeding pipeline in which HTP data are used without variable selection. However, it is recognized that this could affect coefficient stability, therefore future studies should explore dimensionality reduction techniques such as principal component analysis, partial least squares, or variable selection to improve model parsimony. The model E + G + GE + DerivHTP<sub>VegStage</sub> assumes the following:

$$Y_{ij} = \mu + E_i + g_j + Eg_{ij} + \sum_{l=1}^{11} x_{ijl}^{(DerivHTP\_VegStage)} \beta_l^{(DerivHTP\_VegStage)} + \epsilon_{ij} \quad (8)$$

where, once more, all terms except  $\sum_{l=1}^{11} x_{ijl}^{(DerivHTP\_VegStage)} \beta_l^{(DerivHTP\_VegStage)}$  are the same as those defined in model (4), and  $x_{ijl}^{(DerivHTP\_VegStage)}$  represents the DerivHTP<sub>VegStage</sub> covariates listed, and  $\beta_l^{(DerivHTP\_VegStage)}$  as their corresponding effects. The model E + G + DerivHTP<sub>VegStage</sub> is the same as model (8) but without the random interaction effects  $Eg_{ij}$ .

### 2.6.5 | Models E + G + GE + DerivHTP<sub>GrFill</sub> and E + G + DerivHTP<sub>GrFill</sub>

Similar to model (6), but now these models incorporate vegetation indices and PROSAIL-derived traits in grain filling (DerivHTP<sub>GrFill</sub>: CI<sub>red-edge</sub>, NDVI, LAI\_PROSAIL, FCOVER\_PROSAIL, ALA\_PROSAIL, SR, FIPAR\_PROSAIL, CI<sub>green</sub>, C<sub>ab</sub>\_PROSAIL, LAIC<sub>ab</sub>\_PROSAIL, and MCARI). The model E + G + EG + DerivHTP<sub>GrFill</sub> assumes the following:

$$Y_{ij} = \mu + E_i + g_j + Eg_{ij} + \sum_{l=1}^{11} x_{ijl}^{(DerivHTP\_GrFill)} \beta_l^{(DerivHTP\_GrFill)} + \epsilon_{ij} \quad (9)$$

where, once more, all terms except  $\sum_{l=1}^{11} x_{ijl}^{(DerivHTP\_GrFill)} \beta_l^{(DerivHTP\_GrFill)}$  are the same as those defined in model (4) and  $x_{ijl}^{(DerivHTP\_GrFill)}$  represents the DerivHTP<sub>GrFill</sub> covariates listed, and  $\beta_l^{(DerivHTP\_GrFill)}$  as their corresponding effects. The model E + G + DerivHTP<sub>GrFill</sub> is the same as model (9) but without the random interaction effects  $Eg_{ij}$ .

**TABLE 2** Performance of models that consider only one group of variables under a five-fold cross-validation scheme.

Model	Accuracy	NRMSE
E + G	0.57	0.149
E + G + DTH	0.57	0.150
E + G + Phys	0.58	0.148
E + G + DerivHTP <sub>VegStage</sub>	0.62	0.131
E + G + DerivHTP <sub>GrFill</sub>	0.71	0.120
E + G + SpRef	0.68	0.115
E + G + GE	0.56	0.121
E + G + GE + DTH	0.57	0.122
E + G + GE + Phys	0.55	0.121
E + G + GE + DerivHTP <sub>VegStage</sub>	0.61	0.116
E + G + GE + DerivHTP <sub>GrFill</sub>	0.70	0.111
E + G + GE + SpRef	0.68	0.105

Abbreviations: DerivHTP<sub>GrFill</sub>, vegetation indices and PROSAIL-derived traits in grain filling; DerivHTP<sub>VegStage</sub>, vegetation indices and PROSAIL-derived traits in vegetative stage; DTH, days to heading; E, environment; G, genotype; GE, genotype by environment interaction; NRMSE, normalized root mean square error; Phys, physiological variables (stomatal conductance, transpiration, and  $\Phi_{PSII}$ ); SpRef, spectral reflectance.

### 2.6.6 | Models E + G + EG + SpRef and E + G + SpRef

Building upon model (6), these models now incorporate SpRef information (SpRef: R\_Blue, R\_Green, R\_Red, R\_RedEdge, and R\_NIR). The model E + G + EG + SpRef assumes the following:

$$Y_{ij} = \mu + E_i + g_j + Eg_{ij} + \sum_{l=1}^5 x_{ijl}^{(\text{SpRef})} \beta_l^{(\text{SpRef})} + \epsilon_{ij} \quad (10)$$

where all terms except  $\sum_{l=1}^5 x_{ijl}^{(\text{SpRef})} \beta_l^{(\text{SpRef})}$  are the same as those defined in model (4) and  $x_{ijl}^{(\text{SpRef})}$  represents the spectral bands, and  $\beta_l^{(\text{SpRef})}$  as their corresponding effects. The model E + G + SpRef is the same as model (10) but without the random interaction effects  $Eg_{ij}$ .

The 12 models described above constitute the core of this study. To further evaluate their predictive performance, we conducted a systematic analysis by progressively incorporating different groups of variables (Table 2). For example, E + G + DTH was initially tested, then added physiological variables (Phys), resulting in the E + G + DTH + Phys model. Subsequently, vegetation indices from the vegetative stage (DerivHTP<sub>VegStage</sub>) were included, yielding the E + G + DTH + Phys + DerivHTP<sub>VegStage</sub> model. This was followed by the addition of vegetation indices collected during the grain-filling stage (DerivHTP<sub>GrFill</sub>), leading to the comprehensive E + G + DTH + Phys + DerivHTP<sub>VegStage</sub> + DerivHTP<sub>GrFill</sub>

model. This stepwise approach enabled the evaluation of all possible variable combinations, ultimately resulting in 40 additional models, 52 possible combinations in total.

The Bayesian estimation of all GBLUP models was performed using the BGLR R package (Pérez & De Los Campos, 2014), using 20,000 iterations, with the first 5000 discarded as burn-in, and a default thinning of five. Flat prior distributions were used for the environment effects ( $E_i$ ) and the effect of the covariate DTH ( $\beta_{\text{DTH}}$ ); these were specified using model = "FIXED" in BGLR, which treats these predictors as fixed (nonrandom) effects. For the remaining fixed effects ( $\beta_l^{(\text{Phys})}$ ,  $l = 1, 2, 3$ ;  $\beta_i^{(\text{DerivHTP\_VegStage})}$ ,  $i = 1, \dots, 11$ ;  $\beta_i^{(\text{DerivHTP\_GrFill})}$ ,  $i = 1, \dots, 11$ ;  $\beta_i^{(\text{SpRef})}$ ,  $i = 1, \dots, 5$ ), independent normal distributions with mean 0 and  $\sigma_g^2$  were assigned for each group of covariates  $g \in \{\text{Phys}, \text{DerivHTP}_{\text{vegStage}}, \text{DerivHTP}_{\text{GrFill}}, \text{SpRef}\}$ . These priors correspond to a Bayesian ridge regression (BRR) specification in BGLR (model = BRR). Additionally, a scaled inverse chi-squared distribution was used for each  $\sigma_g^2$ .

The five-fold CV scheme involved dividing the dataset into training and testing sets. In each iteration, the models were trained using four of the five folds (80% of the data) and evaluated on the remaining fold (20%), ensuring that individuals in the training set were not included in the testing set. This process was repeated five times, with each fold serving as the testing set once.

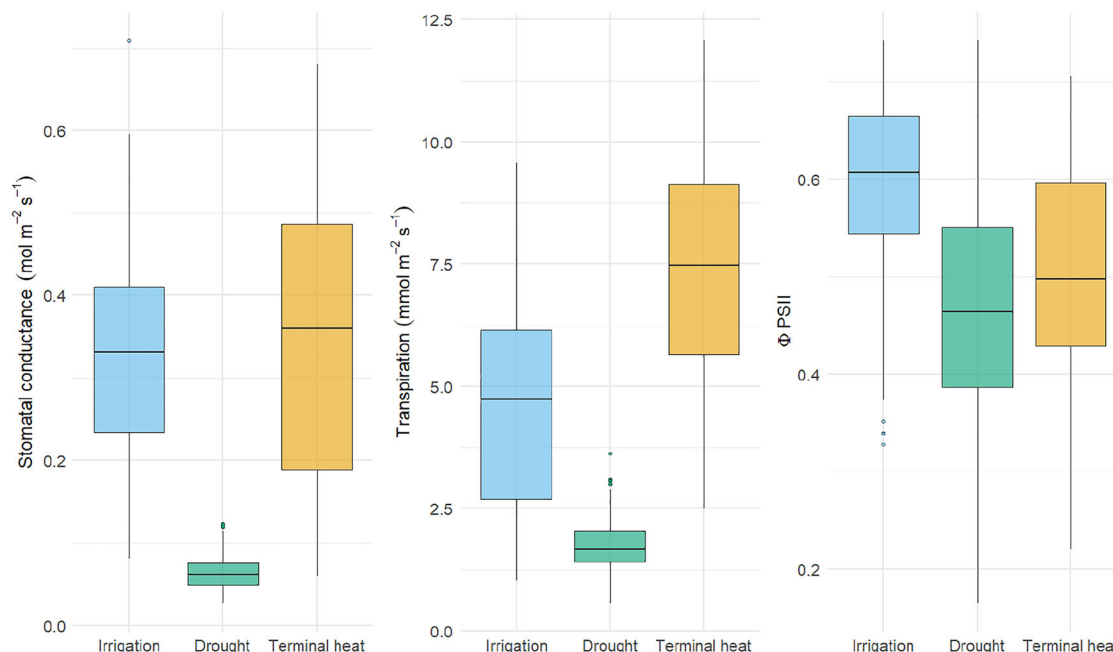
In each CV fold, the normalized root mean square error (NRMSE) and prediction accuracy were evaluated. Prediction accuracy was defined as the Pearson correlation coefficient between the predicted breeding values and observed phenotypes.

These same 40 models were also used in the evaluation of their prediction accuracy under the LOEO validation using phenotypic information from two environments to predict a third (e.g., Ortiz et al., 2023). Graphical visualization of the results was performed using the R packages reshape2 (Wickham, 2007) and ggplot2 (Gómez-Rubio, 2017).

## 3 | RESULTS

### 3.1 | Agronomic, phenological, and physiological traits

The MOLPAN bread wheat population was evaluated under three environments: irrigation, drought stress, and terminal heat stress. Under irrigation, the average grain yield was 898.48 g m<sup>-2</sup>, with values ranging from 204.5 to 1148 g m<sup>-2</sup>. In drought stress, the average yield decreased to 396.69 g m<sup>-2</sup>, ranging from 239.1 to 494.91 g m<sup>-2</sup> and in the terminal heat stress environment, mean yield was 511.65 g m<sup>-2</sup>, ranging from 229 to 665 g m<sup>-2</sup> (Table S1). Associations between



**FIGURE 1** Distribution of stomatal conductance ( $g_s$ ), transpiration, and quantum yield ( $\Phi_{PSII}$ ) measured at grain filling across three environments. Blue, green, and yellow boxplots represent irrigation, drought, and terminal heat environments, respectively.

environments showed a high correlation between irrigation and terminal heat stress and low associations of these two environments with drought (Figure S1). Under irrigation, heading occurred at an average of 84 days, dropping to 74 days under drought, and 52 days under terminal heat stress (Table S1). Stomatal conductance ( $g_s$ ) had a mean of 0.33, 0.06, and 0.34  $\text{mol m}^{-2} \text{s}^{-1}$  in irrigation, drought, and terminal heat conditions, respectively (Figure 1).

Transpiration rates were the lowest at drought environment ( $1.73 \text{ mmol m}^{-2} \text{s}^{-1}$ ), followed by the irrigation environment ( $4.48 \text{ mmol m}^{-2} \text{s}^{-1}$ ) and the highest rates were measured at terminal heat stress ( $7.38 \text{ mmol m}^{-2} \text{s}^{-1}$ ).  $\Phi_{PSII}$  had an average of 0.59 under irrigation, 0.47 under drought, and 0.50 under heat stress environments, respectively (Figure 1; Table S1).

### 3.2 | High-throughput traits

Five spectral bands were measured in all environments at the vegetative and grain-filling stages. In general, reflectance values in the green, NIR, and red-edge bands were higher at the vegetative phase compared to grain-filling, in contrast, reflectance in the red and blue band showed higher values during the grain-filling stage (Figure 2).

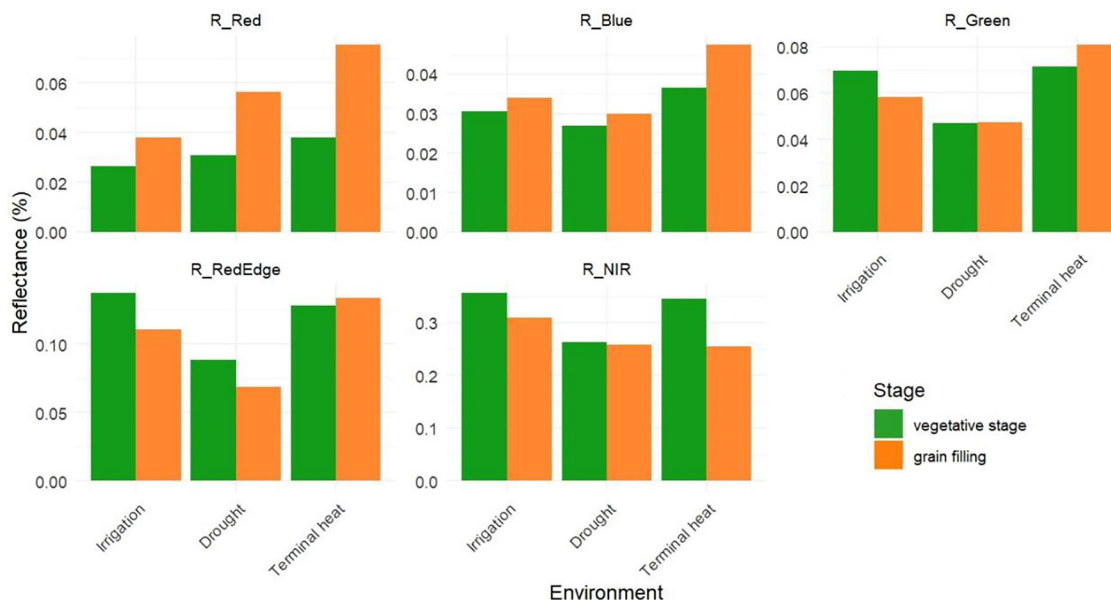
Under irrigation conditions, values for  $CI_{\text{red-edge}}$ , NDVI, LAI\_PROSAIL,  $F_{\text{cover\_PROSAIL}}$ , SR, FIPAR\_PROSAIL,  $CI_{\text{green}}$ , LAIC<sub>ab</sub>\_PROSAIL, and MCARI were the highest among the three evaluated environments (Figure 3).

During the vegetative phase, LAI\_PROSAIL,  $F_{\text{cover\_PROSAIL}}$ , SR, and FIPAR\_PROSAIL reached their highest values compared to the grain-filling stage across all environments. In contrast, ALA\_PROSAIL and  $C_{\text{ab\_PROSAIL}}$  exhibited higher values during the grain-filling phase (Figure 3).

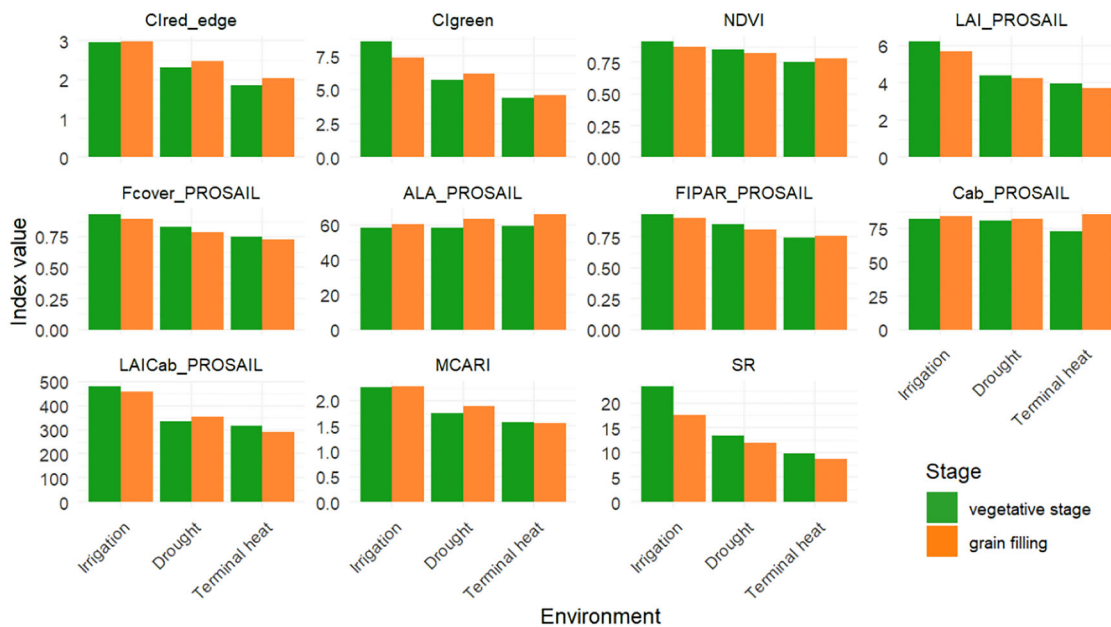
### 3.3 | Evaluating genomic prediction accuracy from five-fold CV

In the five-fold CV, among the models that included only a single group of covariates, the best-performing was E + G + DerivHTP<sub>GrFill</sub>, with a prediction accuracy of 0.71 and an NRMSE of 0.12. In contrast, the worst-performing model was E + G + DTH, with a prediction accuracy of 0.57 and an NRMSE of 0.15. Within this set of models, the most accurate model (E + G + DerivHTP<sub>GrFill</sub>) outperformed the least accurate model (E + G + DTH) by a margin of 24.5% in accuracy (Table 2).

Among the models that included combinations of variables, three achieved a prediction accuracy of 0.75: E + G + DTH + DerivHTP<sub>GrFill</sub> + SpRef, E + G + GE + DTH + Phys + DerivHTP<sub>GrFill</sub>, and E + G + GE + DTH + DerivHTP<sub>GrFill</sub> + SpRef. From these, the last model performed the best in terms of NRMSE (0.099). Additionally, models incorporating vegetation indices from the grain-filling stage outperformed those using indices from the vegetative phase by 10.9% (Table S2).



**FIGURE 2** Mean reflectance values (%) measured in five spectral bands; red (R\_Red), blue (R\_Blue), green (R\_Green), red edge (R\_RedEdge), and near-infrared (R\_NIR), for wheat grown under three environments: optimal irrigation, drought, and terminal heat. Green and orange bars represent traits measured at the vegetative and grain filling stages, respectively.

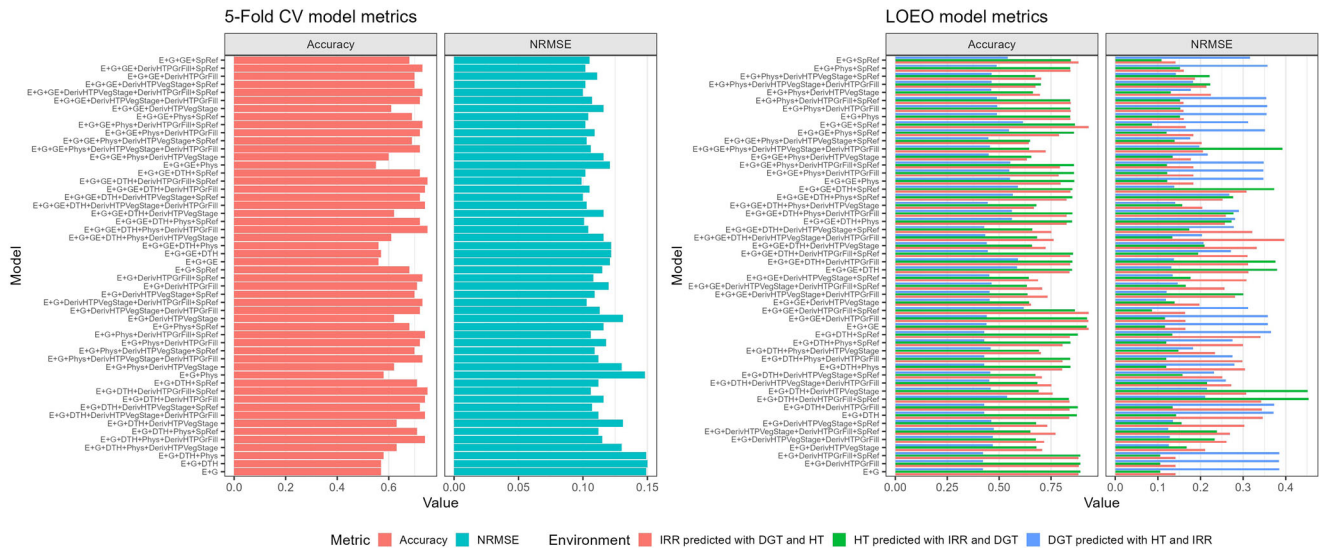


**FIGURE 3** Vegetation indices in three environments (optimal irrigation, drought, and terminal heat), green and orange bars represent traits measured at the vegetative and grain-filling stages, respectively. ALA, average leaf angle;  $C_{ab}$ , chlorophyll a + b content;  $CI_{green}$ , green chlorophyll index;  $CI_{red-edge}$ , red-edge chlorophyll index;  $F_{cover}$ , fraction of vegetation cover; FIPAR, fraction of intercepted photosynthetically active radiation; LAI, leaf area index;  $LAIC_{ab}$ , leaf area index based on chlorophyll content; MCARI, modified chlorophyll absorbance reflectance index; NDVI, normalized differenced vegetation index; SR, simple ratio.

### 3.4 | Evaluating genomic prediction accuracy using LOEO validation

In the LOEO validation, phenotypic data from two environments (training) were used to predict grain yield in a third, unobserved environment (validation). For example,

grain yield data from irrigation and drought stress environments were used to predict yield under terminal heat stress. The results showed differences in prediction accuracy across the evaluated environments. Across the 52 models tested, the average prediction accuracy in the irrigation environment was 0.78, with an NRMSE of 0.52 (Figure 4; Table S3). The lowest



**FIGURE 4** Model metrics for the five-fold cross-validation (CV) (left panel) and leave-one-environment-out (LOEO; right panel) schemes. Red bars represent the model accuracy, cyan bars the normalized root mean square error (NRMSE). For the right panel, orange represents the irrigated (IRR) environment, green represents the terminal heat stress (HT) environment, and blue represents the drought (DGT) environment.

**TABLE 3** Performance of models that consider only one group of variables using the leave-one-environment-out (LOEO) validation approach.

Model	Irrigation		Drought stress		Terminal heat stress	
	Accuracy	NRMSE	Accuracy	NRMSE	Accuracy	NRMSE
E + G	0.79	0.465	0.42	0.385	0.76	0.764
E + G + DTH	0.82	0.515	0.42	0.371	0.75	1.309
E + G + Phys	0.75	0.450	0.42	0.499	0.76	0.640
E + G + DerivHTP <sub>VegStage</sub>	0.73	0.571	0.57	0.470	0.77	1.027
E + G + DerivHTP <sub>GrFill</sub>	0.79	0.465	0.42	0.384	0.76	0.764
E + G + SpRef	0.79	0.465	0.42	0.384	0.76	0.763
E + G + GE	0.80	0.468	0.43	0.358	0.79	0.764
E + G + GE + DTH	0.82	0.549	0.44	0.277	0.74	1.221
E + G + GE + Phys	0.74	0.454	0.44	0.455	0.72	0.649
E + G + GE + DerivHTP <sub>VegStage</sub>	0.79	0.578	0.57	0.393	0.74	0.975
E + G + GE + DerivHTP <sub>GrFill</sub>	0.79	0.468	0.44	0.357	0.79	0.763
E + G + GE + SpRef	0.80	0.468	0.44	0.357	0.79	0.763

Abbreviations: DerivHTP<sub>GrFill</sub>, vegetation indices and PROSAIL-derived traits in grain filling; DerivHTP<sub>VegStage</sub>, vegetation indices and PROSAIL-derived traits in vegetative stage; DTH, days to heading; E, environment; G, genotype; GE, genotype by environment interaction; NRMSE, normalized root mean square error; Phys, physiological variables (stomatal conductance, transpiration, and  $\Phi_{PSII}$ ); SpRef, spectral reflectance.

prediction accuracy was observed in the drought stress environment, with an average accuracy of 0.50 and an NRMSE of 0.42. In contrast, under terminal heat stress, the average prediction accuracy was 0.76, with an NRMSE of 1.01 (Figure 4; Table S3).

Under irrigated conditions, all models achieved high prediction accuracy. Among the models that included only a single group of covariates, E + G + DTH and E + G + GE + DTH performed best, each reaching an accuracy of 0.82 (Table 3). The highest prediction accuracy among models incorporating multiple covariate groups was 0.83, observed in both E + G + GE + DTH + Phys + DerivHTP<sub>VegStage</sub> and

E + G + GE + DTH + DerivHTP<sub>VegStage</sub> + DerivHTP<sub>GrFill</sub> (Table S3). However, these extended models outperformed the best single covariate model by only 1.2%, indicating a marginal improvement. Overall, models including the DTH variable consistently showed strong predictive performance (Table S3).

In the drought stress environment, among the models that included only a single group of variables, the best-performing model was E + G + GE + DerivHTP<sub>VegStage</sub>, achieving a prediction accuracy of 0.57 and an NRMSE of 0.39 (Table 3). Among the models incorporating multiple variable combinations, the highest accuracy was observed in

E + G + DTH + Phys + DerivHTP<sub>VegStage</sub>, with a prediction accuracy of 0.62 and an NRMSE of 0.245. Notably, in this environment, models integrating vegetation indices from the vegetative phase exhibited strong predictive performance (Table S3).

Under terminal heat stress, among the models with a single covariate group, the best performing were E + G + GE, E + G + GE + DerivHTP<sub>GrFill</sub>, and E + G + GE + SpRef, each achieving a prediction accuracy of 0.79 (Table 3). Among the models incorporating multiple variable combinations, the highest accuracy (0.80) was observed in E + G + DTH + DerivHTP<sub>VegStage</sub> and E + G + DerivHTP<sub>VegStage</sub> + DerivHTP<sub>GrFill</sub> + SpRef, with NRMSE values of 1.245 and 1.169, respectively (Table S3). Notably, in this environment, models that included physiological covariates exhibited the lowest NRMSE values (Table S3).

## 4 | DISCUSSION

### 4.1 | Effects of stress conditions on physiological traits and canopy reflectance

In this study, the inclusion of biologically relevant covariates can be seen as a pathway to guide breeders to better genomic prediction results. For example, if a trait such as DTH is known to affect yield formation and source-sink balance, including it in the models helps reducing the environmental noise and improve the genetic values prediction. Similarly, remote sensing datasets that include NDVI or PROSAIL-derived traits can capture light interception, canopy development, and biomass accumulation dynamics over the growth cycle providing proxies of yield potential.

Late heat stress in cultivars led to higher reflectance values in the visible region (400–700 nm), likely due to reduced chlorophyll content and decreased light absorption (Sareen et al., 2024). This reduction might explain why chlorophyll-related vegetation indices ( $CI_{red-edge}$ ,  $CI_{green}$ ,  $LAIC_{ab}$ , and MCARI) were lower under terminal heat, higher under drought, and peaked with irrigation (Figure 3).

Stomatal conductance ( $g_s$ ), which measures the rate of  $CO_2$  into the leaf and water vapor out of the leaf through the stomata, decreases under drought conditions due to partial stomatal closure. This physiological mechanism reduces water loss through transpiration but also limits  $CO_2$  assimilation and the passive uptake and transport of nutrients (Flexas & Medrano, 2002; Liu et al., 2025). In the present study, comparable values of  $g_s$  were observed under irrigation and terminal heat stress (Figure 1). However, under drought stress, the average  $g_s$  was  $0.06 \text{ mol m}^{-2} \text{ s}^{-1}$ , highlighting the plants' response to reduced stomatal aperture and the canopy strategy to conserve water under stress conditions.

In certain scenarios, drought avoidance is accomplished by reducing the water-use rate, thereby extending the time required to deplete soil reserves. This approach is associated with transpiration efficiency, as higher efficiency results in greater biomass production with limited water supply (Ribaut, 2006). In this study, the average transpiration rates were 4.48, 1.73, and  $7.38 \text{ mmol m}^{-2} \text{ s}^{-1}$  under irrigation, drought, and terminal heat conditions, respectively. However, no consistent correlation was found between transpiration and grain yield across environments. Although a weak but statistically significant association ( $p < 0.05$ ) was observed under terminal heat ( $r = 0.16$ ; Figure S2), this minimal relationship suggests that transpiration rates were not a major determinant of yield in this study.

The lack of correlation between transpiration and yield could be attributed to several factors. First, transpiration rates were measured at a specific growth stage, which may not represent the behavior throughout the entire growing season. Second, soil water availability may have influenced the results, as different soil types have varying water retention capacities. A soil with high retention capacity could have provided a more consistent water supply, thereby masking its relationship with yield (Andrade et al., 2009; Ribaut, 2006).

In this study,  $\Phi_{PSII}$  values were 0.59 under irrigation, 0.47 under drought, and 0.50 under terminal heat conditions. These results support the hypothesis that stress conditions in plants induce a reduction in electron transfer activity, leading to a decrease in photochemical efficiency and an increase in regulated heat dissipation (Haque et al., 2014; Liu et al., 2025; Luo et al., 2022; Molina-Salazar et al., 2024).

### 4.2 | Prediction accuracy in the five-fold CV scheme

In our models, we have assumed additive linear effects for covariates for simplicity and model parsimony, however we acknowledge that in future studies nonlinear modeling frameworks such as splines, kernels, or deep learning should be tested to capture more complex trait-response relationships. In the five-fold CV scheme, three models incorporating combinations of variables achieved a prediction accuracy of 0.75. These models outperformed the best model with a single covariate group by 5.6% in accuracy. Notably, all four models utilized vegetation index data collected during the grain-filling stage.

These results are consistent with previous findings by Guo et al. (2020) and Krause et al. (2019), who reported that models incorporating information collected in late-season growth stages, such as heading and grain-filling, exhibited higher prediction accuracy compared to models based on vegetative phase.

The three best models also included the DTH covariate. Regarding this variable, it has been established that late-maturing genotypes tend to have higher yields compared to early maturing ones, as they benefit from a longer period to accumulate biomass (Ibrahim et al., 2024). However, some researchers have pointed out that maturity may introduce confounding effects in yield estimates.

Guo et al. (2020) confirm that correcting grain yield for DTH prevents prediction from being influenced by maturity differences among lines, which is particularly important when evaluating physiological traits that are correlated with phenology. In soybean breeding programs, yield phenotypes are often adjusted to more accurately estimate the intrinsic irrigation, avoiding indirect selection for late maturity (Escamilla et al., 2023; Moreira et al., 2020). Moreno-Amores et al. (2020) addressed the challenge posed by negative correlations between Fusarium head blight resistance and agronomic traits such DTH and plant height in durum wheat. To facilitate effective selection, they corrected Fusarium severity phenotypes through linear regressions using DTH and plant height as covariates. Additionally, they proposed the use of restriction indices to adjust genomic predictions, offering a strategy to better manage unfavorable trait correlations.

Including GE interaction in our prediction models did not increase the Pearson correlations between observed and predicted phenotypes. However, in all cases, it reduced the NRMSE. These results are aligned with previous reports of increased prediction stability by incorporation of GE interactions (Montesinos-López et al., 2022; Ortiz et al., 2023; Song et al., 2024).

### 4.3 | Prediction accuracy in the LOEO validation scheme

The LOEO validation scheme was used to assess whether the models are robust and capable of predicting performance in new environments (not observed during training).

The best-predicted environment was irrigation, in contrast, the worst predicted environment was drought. Our results showed that irrigated and heat environment were closely associated ( $r = 0.792$ ) as water is not a limiting factor, in contrast the drought environment has smaller yields and therefore lower correlation to irrigated environments (both optimal and terminal heat stress irrigation) (Figure S1). No model consistently performed well across all three environments. For instance, the E + G + SpRef model performed well under irrigation, achieving a strong Pearson correlation (0.79) and a low NRMSE (0.465) (Table S3). However, this same model showed considerably lower performance under drought conditions, with a Pearson correlation of only 0.42 and an NRMSE of 0.384 (Table S3). This could be attributed to differences in GE interaction and the specific characteristics of each environment.

It has been reported that GE interactions vary significantly across environments (McBreen et al., 2024; Sørensen et al., 2023) and are influenced by both the specific environmental conditions and the genetic responses of the genotypes. Additionally, in many cases, GE interaction patterns are complex and nonlinear, posing a considerable challenge for traditional genomic prediction models.

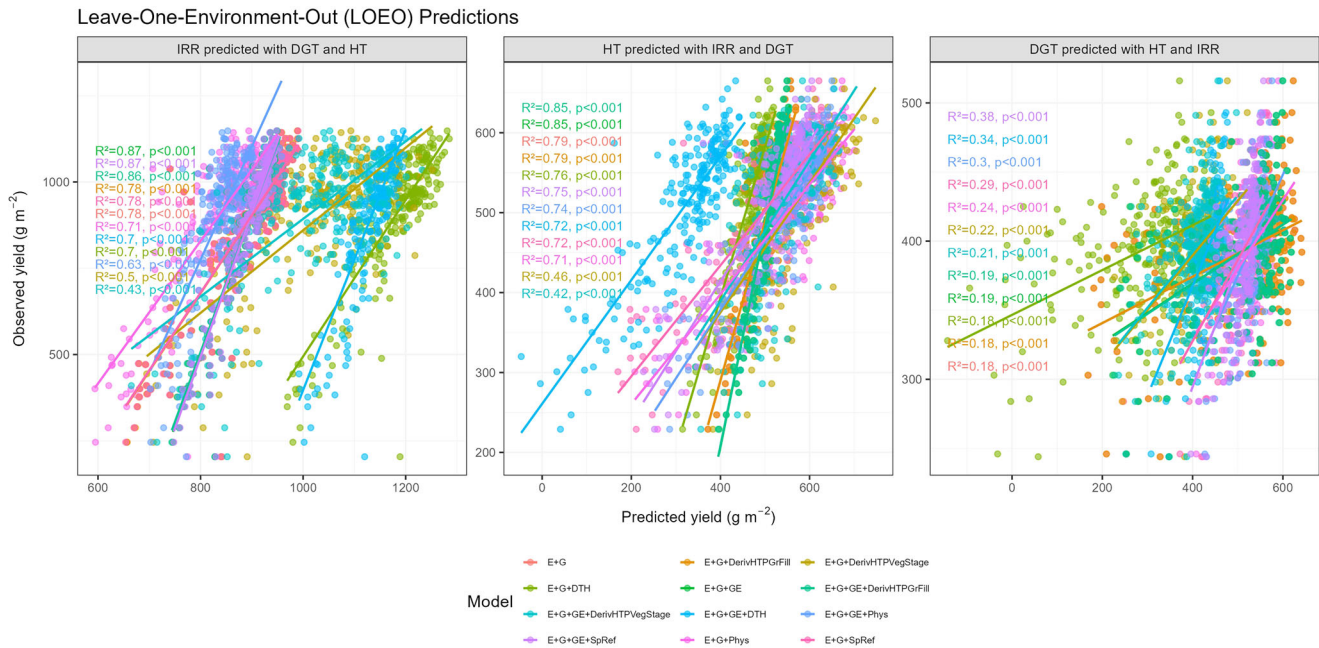
In the LOEO CV scheme, evaluating the impact of covariates can be problematic because this scheme focuses on extrapolation to unobserved environments that are omitted from the training set. In our particular case, this meant predicting terminal heat stress with the irrigation environment, irrigation with the drought environment, and drought with the terminal heat stress environment (Figure 5). This makes it challenging to determine how covariates (phenological, SpRef, vegetation indices, and physiological variables) contribute to improving prediction in other environments, especially when the conditions of unobserved sites fall outside the range of conditions in the sites used for training (McBreen et al., 2024).

For example, grain yield values across environments ranged from 204.5 to 1148 g m<sup>-2</sup> under irrigation, 239.1 to 494.91 g m<sup>-2</sup> under drought, and 229 to 665 g m<sup>-2</sup> under terminal heat stress (Figure 5; Figure S3); similarly, transpiration rates under irrigation ranged from 1.03 to 9.56 mmol m<sup>-2</sup> s<sup>-1</sup>; under drought, from 0.56 to 3.63 mmol m<sup>-2</sup> s<sup>-1</sup>; and under terminal heat, from 2.49 to 12.08 mmol m<sup>-2</sup> s<sup>-1</sup>. This demonstrates that the test environments may present yield, physiological, or remote sensing trait values not represented in the training set, potentially compromising the model's ability to make accurate predictions in these scenarios as showcased by the yield predictions under drought (Figure 5).

### 4.4 | Covariates that increase prediction accuracy

For ease of the discussion in this section, results from the LOEO validation scheme were excluded for the reasons outlined in Section 4.3.

If a breeding program could integrate only one group of covariates, the best option would be vegetation indices and PROSAIL traits derived from spectral datasets collected during the grain-filling stage, as demonstrated by the performance of both the E + G + DerivHTP<sub>GrFill</sub> model and its equivalent E + G + GE + DerivHTP<sub>GrFill</sub>, which includes GE interactions. Prediction accuracy can be further improved by combining multiple variables. For example, the models E + G + DTH + DerivHTP<sub>GrFill</sub> + SpRef, E + G + GE + DTH + DerivHTP<sub>GrFill</sub> + SpRef, and E + G + GE + DTH + Phys + DerivHTP<sub>GrFill</sub> achieving a prediction accuracy of 0.75 and NRMSE values ranging from 0.099 to 0.106.



**FIGURE 5** Scatterplot of observed versus predicted grain yield in the leave-one-environment-out (LOEO) scheme with the 12 models that consider only one group of variables. The left panel is the irrigated (IRR) environment predicted with the drought (SQ) and terminal heat stress (HT) environments, the middle panel is the terminal heat stress (HT) environment predicted with the irrigated (IRR) and drought (DGT) environments, and the left panel is the drought (DGT) environment predicted with the terminal heat stress (HT) and irrigated (IRR) environments.

However, since models incorporating multiple groups of variables outperformed the best model with a single covariate group (E + G + DerivHTP<sub>GrFill</sub>) by only 5.6%, breeders must determine whether this marginal improvement in prediction accuracy justifies the additional costs, time, and complexity required for data collection, processing, and analysis pipelines.

Finally, to optimize the selection process, it is recommended that breeders use a selection index for the simultaneous improvement of multiple traits. This approach can help balance selection pressure across different traits, minimizing the risk of indirect selection for undesirable traits, such as late maturity.

## 5 | CONCLUSIONS

Our results showed that under drought, there was a significant decline in  $g_s$  and  $\Phi_{PSII}$  rates; however, the terminal heat environment had comparable rates to the irrigated one, confirming that water limitation, rather than heat stress per se, was the primary driver of physiological constraints under abiotic stress.

The models that combined vegetation indices with PROSAIL-derived traits measured at grain filling achieved the highest yield prediction accuracy. This emphasizes the importance of targeting relevant phenological stages for data collection to improve the model prediction accuracies. Moreover, the consistency of certain covariates such as DTH or LAI

(in combination with other traits) to be important across environments suggests that a core set of traits could be routinely phenotyped in breeding programs with low-cost methods or HTP tools.

While models trained within a certain environment can perform well under similar environmental conditions, they often fail to generalize across novel environments. Therefore, future studies should prioritize modeling frameworks such as non-linear algorithms or multi-environment calibration to better capture GE interactions.

## AUTHOR CONTRIBUTIONS

**Guillermo García-Barríos:** Formal analysis; investigation; methodology; visualization; writing—original draft; writing—review and editing. **Carlos A. Robles-Zazueta:** Conceptualization; data curation; formal analysis; investigation; methodology; supervision; visualization; writing—original draft; writing—review and editing. **Abelardo Montesinos-López:** Formal analysis; investigation; methodology; writing—review and editing. **Oswal A. Montesinos-López:** Formal analysis; investigation; methodology; writing—review and editing. **Matthew P. Reynolds:** Conceptualization; funding acquisition; writing—review and editing. **Susanne Dreisigacker:** Funding acquisition; methodology; writing—review and editing. **José A. Carrillo-Salazar:** Formal analysis; methodology. **Liana G. Acevedo-Siaca:** Investigation; methodology; writing—review and editing. **Margarita Guerra-Lugo:** Data curation; formal analysis; methodology. **Gilberto**

**Thompson:** Data curation; formal analysis; methodology.  
**José A. Pecina-Martínez:** Data curation; formal analysis; methodology.  
**José Crossa:** Conceptualization; funding acquisition; investigation; methodology; writing—review and editing.

## ACKNOWLEDGMENTS

The authors acknowledge funding from the International Wheat Yield Partnership (IWYP), the Heat and Drought Wheat Improvement Consortium (HeDWIC), the Foundation for Food and Agricultural Research (FFAR), and the Accelerating Genetic Gains in Maize and Wheat (AGG) project. The authors are grateful to Jazmin Rodriguez, Mariana Aleman, and Javier Padilla from Wheat Physiology Lab at CIMMYT for their support with field data collection.

Open access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT


Agronomic, physiologic, and high-throughput traits raw and BLUEs datasets are fully available in the Dryad repository (10.5061/dryad.rbnzs7hqj). Additionally, BLUEs for agronomic and physiological traits are available as Supporting Information. For the genotypic raw dataset of this article, the data will be made available without reservations upon contacting the corresponding authors.


## ORCID

Guillermo García-Barríos  <https://orcid.org/0000-0002-7571-1769>

Carlos A. Robles-Zazueta  <https://orcid.org/0000-0001-8384-4708>

Abelardo Montesinos-López  <https://orcid.org/0000-0002-0525-9973>

Osva A. Montesinos-López  <https://orcid.org/0000-0002-3973-6547>

Matthew P. Reynolds  <https://orcid.org/0000-0002-4291-4316>

Susanne Dreisigacker  <https://orcid.org/0000-0002-3546-5989>

Liana G. Acevedo-Siaca  <https://orcid.org/0000-0003-3903-0402>

José Crossa  <https://orcid.org/0000-0001-9429-5855>

## REFERENCES

- Acquaah, G. (2007). *Principles of plant genetics and breeding* (2nd ed.). Blackwell Publishing.
- Alahmad, S., Rambla, C., Voss-Fels, K. P., & Hickey, L. T. (2022). Accelerating breeding cycles. In M. P. Reynolds & H.-J. Braun (Eds.), *Wheat improvement: Food security in a changing climate* (pp. 557–571). Springer International Publishing. [https://doi.org/10.1007/978-3-030-90673-3\\_30](https://doi.org/10.1007/978-3-030-90673-3_30)
- Alemu, A., Åstrand, J., Montesinos-López, O. A., Isidro y Sánchez, J., Fernández-González, J., Tadesse, W., Vetukuri, R. R., Carlsson, A. S., Ceplitis, A., Crossa, J., Ortiz, R., & Chawade, A. (2024). Genomic selection in plant breeding: Key factors shaping two decades of progress. *Molecular Plant*, 17(4), 552–578. <https://doi.org/10.1016/j.molp.2024.03.007>
- Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., Pérez-Rodríguez, P., & Lopez-Cruz, M. A. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *The Crop Journal*, 8(5), 745–756. <https://doi.org/10.1016/j.cj.2020.03.010>
- Andrade, F. H., Sala, R. G., Pontaroli, A. C., & León, A. (2009). Integration of biotechnology, plant breeding and crop physiology: Dealing with complex interactions from a physiological perspective. In V. Sadras & D. Calderini (Eds.), *Crop physiology* (pp. 267–276). Academic Press. <https://doi.org/10.1016/B978-0-12-374431-9.00011-6>
- Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, 19(1), 52–61. <https://doi.org/10.1016/j.tplants.2013.09.008>
- Bančič, J., Ovenden, B., Gorjanc, G., & Tolhurst, D. J. (2023). Genomic selection for genotype performance and stability using information on multiple traits and multiple environments. *Theoretical and Applied Genetics*, 136(5), 104. <https://doi.org/10.1007/s00122-023-04305-1>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2018). Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding. *The Plant Genome*, 11(1), 170043. <https://doi.org/10.3835/plantgenome2017.05.0043>
- Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., De Los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends in Plant Science*, 22(11), 961–975. <https://doi.org/10.1016/j.tplants.2017.08.011>
- Da Silva, A. V., Famoso, A., Linscombe, S., & Fritsche-Neto, R. (2025). 110 years of rice breeding at LSU: Realized genetic gains and future optimization. *Theoretical and Applied Genetics*, 138(7), 142. <https://doi.org/10.1007/s00122-025-04913-z>
- Duan, S.-B., Li, Z.-L., Wu, H., Tang, B.-H., Ma, L., Zhao, E., & Li, C. (2014). Inversion of the PROSAIL model to estimate leaf area index of maize, potato, and sunflower fields from unmanned aerial vehicle hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, 26, 12–20. <https://doi.org/10.1016/j.jag.2013.05.007>
- Escamilla, D. M., Huang, M., McHale, L., Wang, D., Diers, B., Xavier, A., & Rainey, K. M. (2023). Canopy coverage phenotyping and field spatial variability adjustment as an efficient selection tool in soybean breeding. *Crop Science*, 63(6), 3277–3291. <https://doi.org/10.1002/csc2.21084>
- Escamilla, D. M., Li, D., Negus, K. L., Kappelmann, K. L., Kusmec, A., Vanous, A. E., Schnable, P. S., Li, X., & Yu, J. (2025). Genomic selection: Essence, applications, and prospects. *The Plant Genome*, 18(2), e70053. <https://doi.org/10.1002/tpg2.70053>

- Esposito, S., Vitale, P., Taranto, F., Saia, S., Pecorella, I., D'Agostino, N., Rodriguez, M., Natoli, V., & De Vita, P. (2023). Simultaneous improvement of grain yield and grain protein concentration in durum wheat by using association tests and weighted GBLUP. *Theoretical and Applied Genetics*, *136*(12), 242. <https://doi.org/10.1007/s00122-023-04487-8>
- Féret, J.-B., & De Boissieu, F. (2024). prospect: An R package to link leaf optical properties with their chemical and structural properties with the leaf model PROSPECT. *Journal of Open Source Software*, *9*(94), 6027. <https://doi.org/10.21105/joss.06027>
- Féret, J.-B., Gitelson, A. A., Noble, S. D., & Jacquemoud, S. (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment*, *193*, 204–215. <https://doi.org/10.1016/j.rse.2017.03.004>
- Flexas, J., & Medrano, H. (2002). Drought-inhibition of photosynthesis in C3 plants: Stomatal and non-stomatal limitations revisited. *Annals of Botany*, *89*(2), 183–189. <https://doi.org/10.1093/aob/mcf027>
- Fradgley, N. S., Gardner, K. A., Kerton, M., Swarbreck, S. M., & Bentley, A. R. (2024). Balancing quality with quantity: A case study of UK bread wheat. *Plants, People, Planet*, *6*(5), 1000–1013. <https://doi.org/10.1002/ppp3.10462>
- Gómez-Rubio, V. (2017). ggplot2—Elegant graphics for data analysis (2nd edition). *Journal of Statistical Software*, *77*(2), 1–3. <https://doi.org/10.18637/jss.v077.b02>
- Guo, J., Pradhan, S., Shahi, D., Khan, J., Mcbreen, J., Bai, G., Murphy, J. P., & Babar, M. A. (2020). Increased prediction accuracy using combined genomic information and physiological traits in a soft wheat panel evaluated in multi-environments. *Scientific Reports*, *10*(1), 7023. <https://doi.org/10.1038/s41598-020-63919-3>
- Haboudane, D. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modelling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Haque, M. S., Kjaer, K. H., Rosenqvist, E., Sharma, D. K., & Ottosen, C.-O. (2014). Heat stress and recovery of photosystem II efficiency in wheat (*Triticum aestivum* L.) cultivars acclimated to different growth temperatures. *Environmental and Experimental Botany*, *99*, 1–8. <https://doi.org/10.1016/j.envexpbot.2013.10.017>
- Herr, A. W., Schmuker, P., & Carter, A. H. (2024). Large-scale breeding applications of unoccupied aircraft systems enabled genomic prediction. *The Plant Phenome Journal*, *7*(1), e20101. <https://doi.org/10.1002/ppj2.20101>
- Hunt, E. R., Daughtry, C. S. T., Eitel, J. U. H., & Long, D. S. (2011). Remote sensing leaf chlorophyll content using a visible band index. *Agronomy Journal*, *103*(4), 1090–1099. <https://doi.org/10.2134/agronj2010.0395>
- Ibrahim, A., Stuerz, S., Manneh, B., Rebolledo, M. C., & Saito, K. (2024). Consistent yield performance of rice genotypes grown under irrigated conditions in wet and dry seasons in West Africa. *Field Crops Research*, *306*, 109231. <https://doi.org/10.1016/j.fcr.2023.109231>
- Jackson, R. D., & Huete, A. R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*, *11*(3–4), 185–200. [https://doi.org/10.1016/s0167-5877\(05\)80004-2](https://doi.org/10.1016/s0167-5877(05)80004-2)
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., & Ustin, S. L. (2009). PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, *113*, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>
- Jung, M., Quesada-Traver, C., Roth, M., Aranzana, M. J., Muranty, H., Rymenants, M., Guerra, W., Holzknicht, E., Pradas, N., Lozano, L., Didelot, F., Laurens, F., Yates, S., Studer, B., Broggin, G. A. L., & Patocchi, A. (2024). Integrative multi-environmental genomic prediction in apple. *Horticulture Research*, *12*(2), uhae319. <https://doi.org/10.1093/hr/uhae319>
- Krause, M. R., González-Pérez, L., Crossa, J., Pérez-Rodríguez, P., Montesinos-López, O., Singh, R. P., Dreisigacker, S., Poland, J., Rutkoski, J., Sorrells, M., Gore, M. A., & Mondal, S. (2019). Hyperspectral reflectance-derived relationship matrices for genomic prediction of grain yield in wheat. *G3 Genes|Genomes|Genetics*, *9*(4), 1231–1247. <https://doi.org/10.1534/g3.118.200856>
- Liu, D., Lu, S., Tian, R., Zhang, X., Dong, Q., Ren, H., Chen, L., & Hu, Y.-G. (2025). Mining genomic regions associated with stomatal traits and their candidate genes in bread wheat through genome-wide association study (GWAS). *Theoretical and Applied Genetics*, *138*(1), 20. <https://doi.org/10.1007/s00122-024-04814-7>
- Luo, Y., Sun, M., Gao, Y., Lang, S., & Wang, Y. (2022). Exogenous trehalose protects photosystem II in heat-stressed wheat. *Botanical Sciences*, *101*(1), 186–196. <https://doi.org/10.17129/botsci.3038>
- McBreen, J., Babar, M. A., Jarquin, D., Khan, N., Harrison, S., DeWitt, N., Mergoum, M., Lopez, B., Boyles, R., Lyerly, J., Murphy, J. P., Shakiba, E., Sutton, R., Ibrahim, A., Howell, K., Smith, J. H., Brown-Guedira, G., Tiwari, V., Santantonio, N., & Van Sanford, D. A. (2024). Enhancing prediction accuracy of grain yield in wheat lines adapted to the southeastern United States through multivariate and multi-environment genomic prediction models incorporating spectral and thermal information. *The Plant Genome*, *18*(1), e20532. <https://doi.org/10.1002/tpg2.20532>
- Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>
- Molina-Salazar, C. Y., Morales-Nieto, C. R., Álvarez-Holguín, A., Corrales-Lerma, R., & Santellano-Estrada, E. (2024). Relación de la eficiencia fotosintética, características estomáticas y producción forrajera en poblaciones de pasto banderita. *Revista Fitotecnia Mexicana*, *47*(2), 147. <https://doi.org/10.35196/rfm.2024.2.147>
- Montesinos-López, O. A., Franco-Pérez, E., Luna-Vázquez, F. J., Salinas-Ruiz, J., Sandoval-Carrillo, S., Valenzo Jiménez, M. A., Cuervas, J., & Santana-Mancilla, P. C. (2020). Benchmarking between item based collaborative filtering algorithm and genomic best linear unbiased prediction (GBLUP) model in terms of prediction accuracy for wheat and maize. *Biotechnia*, *22*(2), 136–146. <https://doi.org/10.18633/biotechnia.v22i2.1255>
- Montesinos-López, O. A., Montesinos-López, A., Bernal Sandoval, D. A., Mosqueda-Gonzalez, B. A., Valenzo-Jiménez, M. A., & Crossa, J. (2022). Multi-trait genome prediction of new environments with partial least squares. *Frontiers in Genetics*, *13*, 966775. <https://doi.org/10.3389/fgene.2022.966775>
- Moreira, F. F., Oliveira, H. R., Volenec, J. J., Rainey, K. M., & Brito, L. F. (2020). Integrating high-throughput phenotyping and statistical genomic methods to genetically improve longitudinal traits in crops. *Frontiers in Plant Science*, *11*, 681. <https://doi.org/10.3389/fpls.2020.00681>
- Moreno-Amores, J., Michel, S., Miedaner, T., Longin, C. F. H., & Buerstmayr, H. (2020). Genomic predictions for Fusarium head blight resistance in a diverse durum wheat panel: An effective incorporation of plant height and heading date as covariates. *Euphytica*, *216*(2), 22. <https://doi.org/10.1007/s10681-019-2551-x>

- Mueller, M. L., & Van Eenennaam, A. L. (2022). Synergistic power of genomic selection, assisted reproductive technologies, and gene editing to drive genetic improvement of cattle. *CABI Agriculture and Bioscience*, 3(1), 13. <https://doi.org/10.1186/s43170-022-00080-z>
- Ortiz, R., Reslow, F., Montesinos-López, A., Huicho, J., Pérez-Rodríguez, P., Montesinos-López, O. A., & Crossa, J. (2023). Partial least squares enhance multi-trait genomic prediction of potato cultivars in new environments. *Scientific Reports*, 13(1), 9947. <https://doi.org/10.1038/s41598-023-37169-y>
- Pérez, P., & de los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, 198(2), 483–495. <https://doi.org/10.1534/genetics.114.164442>
- Pietragalla, J., & Pask, A. (2013). Grain yield and yield components. In A. Pask, J. Pietragalla, D. Mullan, P. Chávez, & M. P. Reynolds (Eds.), *Physiological breeding II: A field guide to wheat phenotyping* (pp. 95–103). CIMMYT. <https://repository.cimmyt.org/handle/10883/1288>
- Reynolds, M. P., & Braun, H.-J. (2022). Wheat improvement. In M. P. Reynolds & H.-J. Braun (Eds.), *Wheat improvement: Food security in a changing climate* (pp. 3–15). Springer International Publishing. [https://doi.org/10.1007/978-3-030-90673-3\\_1](https://doi.org/10.1007/978-3-030-90673-3_1)
- Ribaut, J.-M. (Ed.). (2006). *Drought adaptation in cereals*. Food Products Press.
- Rincint, R., Charpentier, J.-P., 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|Genetics*, 8(12), 3961–3972. <https://doi.org/10.1534/g3.118.200760>
- Robles-Zazueta, C. A., Molero, G., Pinto, F., Foulkes, M. J., Reynolds, M. P., & Murchie, E. H. (2021). Field-based remote sensing models predict radiation use efficiency in wheat. *Journal of Experimental Botany*, 72(10), 3756–3773. <https://doi.org/10.1093/jxb/erab115>
- Sandro, P., Bhatta, M., Bower, A., Carlson, S., Jannink, J.-L., Waring, D. J., Birkett, C., Smith, K., Wiersma, J., Caffè, M., Kleinjan, J., McMullen, M. S., English, L., & Gutierrez, L. (2024). Genomic prediction for targeted populations of environments in oat (*Avena sativa*). *Crop & Pasture Science*, 75(5), CP23126. <https://doi.org/10.1071/CP23126>
- Sareen, S., Meena, B. K., Sarial, A., & Kumar, S. (2024). Dissecting physiological traits for drought and heat tolerance in wheat. *Cereal Research Communications*, 52(4), 1373–1384. <https://doi.org/10.1007/s42976-023-00463-6>
- Schrauf, M. F., Martini, J. W. R., Simianer, H., de los Campos, G., Cantet, R., Freudenthal, J., Korte, A., & Munilla, S. (2020). Phantom epistasis in genomic selection: On the predictive ability of epistatic models. *G3 Genes|Genomes|Genetics*, 10(9), 3137–3145. <https://doi.org/10.1534/g3.120.401300>
- Sehgal, D., Rosyara, U., Mondal, S., Singh, R., Poland, J., & Dreisigacker, S. (2020). Incorporating genome-wide association mapping results into genomic prediction models for grain yield and yield stability in CIMMYT spring bread wheat. *Frontiers in Plant Science*, 11, 197. <https://doi.org/10.3389/fpls.2020.00197>
- Senapati, N., Semenov, M. A., Halford, N. G., Hawkesford, M. J., Asseng, S., Cooper, M., Ewert, F., van Ittersum, M. K., Martre, P., Olesen, J. E., Reynolds, M., Rötter, R. P., & Webber, H. (2022). Global wheat production could benefit from closing the genetic yield gap. *Nature Food*, 3(7), 532–541. <https://doi.org/10.1038/s43016-022-00540-9>
- Song, J., Pacheco, A., Alakonya, A., Cruz-Morales, A. S., Muñoz-Zavala, C., Qu, J., Wang, C., Zhang, X., San Vicente, F., & Dhliwayo, T. (2024). Genome-wide association mapping and genomic prediction of stalk rot in two mid-altitude tropical maize populations. *The Crop Journal*, 12(2), 558–568. <https://doi.org/10.1016/j.cj.2024.02.004>
- Sørensen, E. S., Jansen, C., Windju, S., Crossa, J., Sonesson, A. K., Lillemo, M., & Alsheikh, M. (2023). Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa*). *Plant Breeding*, 142(1), 41–53. <https://doi.org/10.1111/pbr.13061>
- van Eeuwijk, F. A., Bustos-Korts, D., Millet, E. J., Boer, M. P., Kruijer, W., Thompson, A., Malosetti, M., Iwata, H., Quiroz, R., Kuppe, C., Muller, O., Blazakis, K. N., Yu, K., Tardieu, F., & Chapman, S. C. (2019). Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. *Plant Science*, 282, 23–39. <https://doi.org/10.1016/j.plantsci.2018.06.018>
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Walter, J., Edwards, J., Cai, J., McDonald, G., Miklavcic, S. J., & Kuchel, H. (2019). High-throughput field imaging and basic image analysis in a wheat breeding programme. *Frontiers in Plant Science*, 10, 449. <https://doi.org/10.3389/fpls.2019.00449>
- Wang, F., Feldmann, M. J., & Runcie, D. E. (2025). *Don't benchmark phenomic prediction against genomic prediction accuracy*. BioRxiv. <https://doi.org/10.1101/2025.01.09.632209>
- Wang, H. (Ed.). (2023). *Case studies of breeding strategies in major plant species*. IntechOpen. <https://doi.org/10.5772/intechopen.100882>
- Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12), 1–20. <https://doi.org/10.18637/jss.v021.i12>
- Xu, Y., Ma, K., Zhao, Y., Wang, X., Zhou, K., Yu, G., Li, C., Li, P., Yang, Z., Xu, C., & Xu, S. (2021). Genomic selection: A breakthrough technology in rice breeding. *The Crop Journal*, 9(3), 669–677. <https://doi.org/10.1016/j.cj.2021.03.008>
- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 2017, 1–17. <https://doi.org/10.1155/2017/1353691>
- Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, 14(6), 415–421. <https://doi.org/10.1111/j.1365-3180.1974.tb01084.x>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** García-Barríos, G., Robles-Zazueta, C. A., Montesinos-López, A., Montesinos-López, O. A., Reynolds, M. P., Dreisigacker, S., Carrillo-Salazar, J. A., Acevedo-Siaca, L. G., Guerra-Lugo, M., Thompson, G., Pecina-Martínez, J. A., & Crossa, J. (2025). Integration of physiological and remote sensing traits for improved genomic prediction of wheat yield. *The Plant Genome*, 18, e70110. <https://doi.org/10.1002/tpg2.70110>