Extending genomic prediction to future climates through crop modelling. A case study on heading time in barley.

Livia Paleari[1,*], Alessandro Tondelli[2], Luigi Cattivelli[2], Ernesto Igartua[3], Ana M. Casas[3], Andrea Visoni[4], Alan H. Schulman[5,6], Laura Rossini[7], Robbie Waugh[8], Joanne Russell[8], Roberto Confalonieri[1]

[1] Università degli Studi di Milano, ESP, Cassandra lab, via Celoria 2, 20133 Milan, Italy

[2] CREA, Research Centre for Genomics and Bioinformatics, via S. Protaso 302, 29017 San Protaso, Italy

[3] Spanish National Research Council, Department of Genetics and Plant Production, E-50080-Zaragoza, Spain

[4] International Center for Agricultural Research in the Dry Areas, Rabat 10100, Morocco.

[5] University of Helsinki, Helsinki, Finland.

[6] Production Systems, Natural Resources Institute Finland (LUKE), Helsinki, Finland

[7] Università degli Studi di Milano, DISAA, via Celoria 2, 20133 Milan, Italy

[8] The James Hutton Inst., Invergowrie, Dundee, DD2 5DA UK

*Corresponding author: livia.paleari@unimi.it

Abstract

Integrating genomic prediction (GP) and biophysical crop models has the potential to support plant breeding in defining adapting strategies to climate change. However, whether this integrated approach can actually broaden the prediction domain to unexplored environments is still unclear. We showed how crop models can extend GP to new environments and capture genotype-specific response to future climate conditions.

Days to heading (HD) in spring barley was used as a case study, given the primary role of phenology for crop adaptation. Phenotypic and genomic (50K Illumina SNP array) data for 151 two-row genotypes were used, with HD determined on 17 site × season combinations spread from the Mediterranean basin to Northern Europe. A dedicated modelling solution was developed by integrating approaches from the WARM and WOFOST crop models, with optimization algorithms used to derive accession-specific values of most relevant model parameters. GP was carried out on the model parameters with the R package rr-BLUP, and the resulting markers-derived parameters were used to simulate heading date with the crop model explicitly reproducing G×E×M interactions. Prediction accuracy for unknown genotypes was evaluated through ten-fold cross validation, whereas the capability of crop models to extend GP to unexplored environments was evaluated with a leave-one site-out cross validation.

Results were encouraging, with average R^2, RMSE and Nash-Sutcliffe efficiency for unknown genotypes in unexplored environments equal to 0.97, 9.27, and 0.94, respectively. Accession-specific HD under future climates (20 seasons centred on 2060 generated for the IPSL-CM6A-LR realization of SSP3-7.0) highlighted large G×E×M interactions, confirming the potential of integrating crop modelling and GP to supporting breeding programs targeting adaptation to climate change.

Keywords

Genomic prediction; spring barley; phenology; climate change; photoperiod; genotype × environment interactions.

Primary research article

1. Introduction

The development of new cultivars more suited to the climate conditions expected for the coming decades is one of the most promising adaptation strategies to preserve crop productivity and satisfy the projected global food demand in the mid- and long-term (Reynolds et al., 2017; Snowdon et al., 2021). In this context, understanding and predicting relationships between genotype features and phenotypic expression of traits is crucial to accelerate the rate of genetic gain in plant breeding (Hammer et al., 2006; Messina et al., 2011; Bernardo, 2020). Alongside huge progresses in genotyping techniques and phenotyping tools, statistical approaches to understand gene-to-phenotype relationships have seen crucial developments in recent decades (van Etten et al., 2023).

Genome-wide prediction (GP hereafter, Meuwissen et al., 2001) defines relationships between whole-genome information (SNP markers) and phenotypic data for a training pool of genotypes, to predict the phenotype of new accessions based on their genomic makeup. Applications to plant breeding have proved successful for predicting quantitative traits in different species (Bernardo 2008; Heffner et al., 2009; Cooper et al., 2014). However, such genotype-to-phenotype relationships are purely based on statistics, so there are no guarantees – by definition – they can be applied to conditions different from those used for their development, since (G) × environment (E) × management (M) interactions can be strongly non-linear (Cooper et al., 2021). Moreover, the lack of any ecophysiological insight in the GP-framework can lead to low prediction accuracy when analyzing traits affected by large G×E×M interactions, such as yield under drought conditions (Heslot et al., 2014; Messina et al., 2018).

Process-based crop simulation models (CSMs) reproduce growth and development of genotypes as a function of environmental variables (climate, soil, management) and of model parameters (also called "genetic coefficients", Boote et al., 2013) used to describe their morphological and physiological

features. As such, they represent a unique tool to quantitatively analyze G×E×M interactions by explicitly accounting for plant physiological processes (Messina et al., 2006; Boote et al., 2013; Paleari et al., 2022). Deriving relationships between crop model parameters – in turn representing functional traits (i.e., traits whose expression is lightly influenced by environmental conditions like, e.g., sensitivity to photoperiod; Litrico and Violle, 2015) – and genomic data would allow to introduce a biological component in the evaluation of gene to phenotype relations, dynamically capturing variation in performance traits (e.g., heading time) in response to the surrounding environment (Diepenbrock et al., 2021; Yang et al., 2022). One of the greatest potentials of this approach is the possibility of predicting the phenotype of new or existing accessions under environmental conditions different from those explored for phenotypic data collection. The prediction domain could thus be drastically extended to encompass other geographic areas, management and climate scenarios, as well as a spatially distributed evaluation of phenotypic performance (Paleari et al., 2015; Messina et al., 2018; van Etten et al., 2023).

Crop development is a critical feature for adaptation to different agro-climatic contexts, with phenological traits explaining large part of yield variability under current and future climate conditions (Paleari et al., 2022). Accuracy of GP for phenological traits for new accessions is thus crucial to evaluate their potential adaptation to different growing conditions and to future climate scenarios.

Given the capability of crop models to explicitly account for G×E×M interactions while predicting crop development, many attempts have been recently carried out to link crop model parameters to the underpinning genetic base. As shown by Onogi et al. (2016) for rice, integrating GP techniques and crop modelling would improve the prediction of heading date as compared to genomic or modelling approaches alone. However, studies testing the capability of CSMs-GP integration to extend the prediction to completely new environments are lacking. Attempts have been conducted (Onogi et al.,

2016; Kawakita et al., 2024) but the set of environments explored was always limited to the same country (Japan), and the validation against unexplored conditions was carried out on a single location. In this study, the capability of crop models to extend genomic prediction to new environments is validated on a wide range of conditions, referring to an extremely diverse pool of climates, from Mediterranean conditions (e.g., Morocco) to far North ones (e.g., Finland), and including different sowing times (spring/autumn) within the same country. This provides a unique framework to analyze the limits of crop model-aided GP for effectively reproducing G×E×M interactions when growing conditions completely different from those evaluated with training datasets are explored.

By using heading date in barley as a case study, this work was focused on predicting the phenotype of new, untested genotypes in a range of completely unexplored environments – including future climates – by extending GP with crop modelling. This integrated approach would allow to expand the prediction domain to untested agro-climatic contexts and projected climate scenarios, thus supporting breeding efforts targeting adaptation via G × E× M optimization.

2. Materials and methods

2.1. Genotypic and phenotypic data

Phenotypic and genotypic data for 151 two-rows, spring barley genotypes (*Hordeum vulgare convar. distichon* (L.)) from the ClimBAR and ExBarDiv projects were used (Tondelli et al., 2013; Bretani et al., 2022; Montardit-Tarda et al. 2024). The collection is a representative set of the 213 genotypes described by Tondelli et al. (2013), and it includes cultivars released in different EU countries between 1924 and 2006, thus representing over 80 years of spring barley breeding in Europe (Tondelli et al. 2013). As reported by Tondelli et al. (2013), both STRUCTURE and PCO analyses did not result in a clear population structure, rather a gradient of differentiation between older and more modern barley cultivars as a result of directional selection on specific breeding targets. Principal

Component Analysis on genomic markers for the 151 genotypes used in this study confirmed these results (Fig. S1). Genomic data consist of 40639 single-nucleotide polymorphisms (SNPs) from the Illumina 50K chip – the latter previously validated through comparison with legacy datasets and exome capture SNPs showing rates of agreement higher than 93% (Bayer et al., 2017) – reduced to 28988 markers after filtering for minor alleles frequency above 0.05.

The phenological trait used in this study refers to the heading date, defined as Zadoks stage Z55 (Zadoks et al., 1974), i.e., days from sowing to 50% of tillers with mid-head exertion (HD, hereafter) evaluated with a multi-environment trial involving 17 combinations of location × season (Table 1). To fully explore crop response to environmental factors, field trials were conducted in a wide range of conditions, as provided by seven locations spread from the Mediterranean basins to Northern Europe and by both autumn and spring sowing times (Table 1). The result is a wide range of temperature and daylength conditions (Table 1), which reflects on a large variability in days to heading (Fig. S2), with average values of all genotypes ranging from 53 days (LUKE16 dataset) to 175 days (CREA17 dataset). Sowing dates and field management were defined according to local practices, to ensure typical growing conditions for barley at the different sites. Plots in all trials followed an alpha-lattice design with two replicates. Each plot was 2-3 m long and 1-1.5 m width, and included eight rows. After data quality checks, best linear unbiased estimators (BLUEs) of HD were calculated with Genstat 20 (VSN 129 International, 2022). Full details about the experimental scheme, data collection and data analysis can be found in Montardit-Tarda et al. (2024).

138 *Table 1. Multi-environment trials used to evaluate days to heading of the 151 spring barley genotypes. All genotypes were grown in all*

139 *environments. A synthetic agro-climatic characterization is reported for each dataset.*

Dataset ID	Country	Location	Latitude, Longitude	Sowing date	T_{avg} (°C)[a]	Daylength (h)[b]
CREA09	Italy	Fiorenzuola d'Arda	44°55'52.5"N, 9°53'41.0"E	2/17/2009	12.4	13.1
CREA10	Italy	Fiorenzuola d'Arda	44°55'58.3"N, 9°53'57.5"E	3/1/2010	12.4	13.6
CREA16	Italy	Fiorenzuola d'Arda	44°55'31.8"N, 9°53'39.3"E	11/5/2015	7.2	10.6
CREA17	Italy	Fiorenzuola d'Arda	44°55'36.1"N, 9°53'36.3"E	11/8/2016	7.1	10.8
CSIC16	Spain	El Vedado	41°51'36"N, 0°39'14"W	11/11/2015	8.2	10.8
CSIC17	Spain	El Vedado	41°51'30.2"N, 00°39'15.3"W	11/15/2016	8.1	10.7
ICARDA16	Morocco	Marchouch	33°37'10"N, 6°43'00"W	1/6/2016	13.8	11.5
IPK09	Germany	Gatersleben	51°49'38.3"N, 11°16'10.0"E	4/3/2009	13.6	15.2
IPK10	Germany	Gatersleben	51°49'38.3"N, 11°16'10.0"E	4/8/2010	12.1	15.5
JHI09	Scotland	Dundee	56°29'04.4"N, 3°08'16.7"W	3/25/2009	9.3	15.8
JHI10	Scotland	Dundee	56°29'04.9"N, 3°07'45.6"W	4/1/2010	9.2	16.0
JHI16	Scotland	Balruddery	56°28'50.6"N, 3°06'35.3"W	3/16/2016	8.2	15.4
JHI17	Scotland	Balruddery	56°29'00.7"N, 3°07'08.7"W	3/29/2017	9.7	15.8
LUKE16	Finland	Museopelto	60°13'22"N, 25°00'43"E	5/11/2016	14.1	18.4
LUKE17	Finland	Museopelto	60°13'22"N, 25°00'43"E	5/19/2017	12.4	18.5
UHal09	Germany	Halle	51°29'46.05"N, 11°59'29.58"E	3/31/2009	13.6	14.9
UHal10	Germany	Halle	51°29'46.1"N, 11°59'29.6"E	4/6/2010	12.2	15.3

140 [a] mean daily air temperature and [b] daylength for the period between sowing and heading date (HD) of the accession with the longest cycle.

2.2. Crop modelling approach

A daily time step modelling solution was developed, including the phenology component of the crop model WARM (Confalonieri et al., 2009) – selected because it provides a good compromise between accuracy and robustness (i.e., reliability under a wide range of conditions; Confalonieri et al. 2010), the latter being a key feature in the context of extending the prediction to new environments – further extended to account for photoperiodic response of barley with algorithms derived from WOFOST (van Diepen et al., 1989). As reported by Yang et al. (2022) this phenological model demonstrated its reliability for being integrated with GP techniques, also for its parsimony in terms of parameterization requirements.

According to this approach, which involves six genotype-specific parameters (Table S1), daily thermal time accumulation – expressed as growing degree days (*GDD*, °C day^{-1}) – is derived as a function of the average daily air temperature (T_{avg}, °C) and the two parameters base temperature (T_{base}, °C) and cut-off temperature (T_{max}, °C):

$$GDD = \begin{cases} 0 & T_{avg} \leq T_b \\ T_{avg} - T_b & T_b < T_{avg} < T_{max} \\ T_{max} - T_b & T_{avg} \geq T_{max} \end{cases} \quad \text{(Eq. 1)}$$

Daily cumulated thermal time is then corrected for the effect of photoperiod (F_{pr}, photoperiod reduction factor, from 0 to 1, unitless), which is derived as a function of daylength (*DL*, hr), and of the three parameters optimum ($Photo_{opt}$, hr) and critical ($Photo_{crit}$, hr) daylengths, and photoperiod sensitivity ($Photo_{sens}$, from 0 to 1, unitless):

$$F_{pr} = \frac{DL - Photo_{crit}}{Photo_{opt} - Photo_{crit}} \cdot Photo_{sens} \quad \text{(Eq. 2)}$$

F_{pr} is set to 0 and 1 in case of values below or above these limits, respectively. Daily cumulated thermal time is then integrated over time to derive the cumulated thermal time since sowing (GDD_{cum}, °C day^{-1}). Heading stage is achieved when GDD_{cum} is equal to (or higher than) the thermal time requirements to reach heading (parameter GDD_{HD}, °C day^{-1}).

2.3. Integration of crop modelling and genomic prediction

Genomic information was linked to model parameters with a two-step approach, as in Yang et al. (2022). According to this approach, the observed phenotypic trait (heading date in this study) is firstly decomposed into functional traits (Litrico and Violle, 2015), which represent the genotype-specific response to environmental factors, such as temperature and photoperiod in case of phenology. According to Yang et al. (2022), trait decomposition was carried out by deriving accession-specific values of model parameters representing functional traits by using the evolutionary shuffled simplex optimization algorithm (Duan et al., 1992) and the phenotypic data collected in multi-environment trials.

To avoid equifinality issues, leading to uncertain parameter estimation due to multiple combinations of values providing the same model output, sensitivity analysis (SA) was carried out to identify the parameters with the largest impact on HD. The variance-based global sensitivity analysis method of Sobol (Sobol', 1993) was used. The method was parameterized as in Paleari et al., (2017), with the number of simulations carried out in each experiment (i.e., each location × season combination) equal to the lowest value of $M|M > (\gamma \cdot n)$, with $M = 2^{q+3}(2n + 2)$, $q = \{1, 2, 3, \dots, Q\}$, γ is the number of model runs for each parameter (Confalonieri et al., 2010), and n is the number of parameters under SA, in this study equal to six. Parameter distributions were derived from literature (Table S1). According to SA results, the most influential parameters for HD across all environments

were selected, and their genotype-specific value derived by minimizing the root mean squared error (RMSE) between observed and simulated HD through the optimization algorithm.

In the second step, accessions were divided into training and test sets. Relationships between the optimized model parameters and markers were defined for the training set through GP approaches and used to predict model parameters for the test set based on markers data. Genomic-predicted model parameters and the crop model were then used to predict HD for test accessions in each environment, leaving to the crop model the responsibility to explicitly reproduce the interaction between model parameters (representing functional traits) and environmental and management conditions. Genomic prediction of model parameters was carried out with the R (R core Team, 2022) rr-BLUP package (Endelman et al., 2011), which allows to control for both kinship and population structure. The Ridge Regression-BLUP approach was chosen given the good results achieved in similar studies carried out on rice (Yang et al., 2022) and in studies where different methods were compared (Kawakita et al., 2024). As in Yang et al. (2022), kinship matrix was derived by using the additive relationship matrix function of rr-BLUP. Population structure was evaluated through Principal Component Analysis on genomic markers (Fig. S1), but no subpopulations were identified in line with Tondelli et al. (2013). A ten-fold cross validation scheme was used, randomly selecting 10% of genotypes as test set and using the remaining 90% as training set (Endelman, 2011; Schrauf et al., 2021). The process was iterated multiple times to ensure that each genotype was included in the test set at least once. Genomic-predicted parameters were then used to simulate HD for test genotypes in all environments, and prediction accuracy was evaluated by comparing simulation outputs with observed data. Agreement between observed and simulated HD was evaluated for each accession separately by using the following metrics: root mean square error (RMSE; Jørgensen et al., 1986), Nash and Sutcliffe modelling efficiency (EF; Nash and Sutcliffe, 1970), coefficient of residual mass (CRM; Loague and Green, 1991), and the R^2 of the linear regression between observed and simulated data.

Given that the objective of this study was to evaluate the potential of using crop models to extend GP to unexplored environments, the cross-validation scheme described above – which quantifies prediction accuracy for new (untested) genotypes in known environments (blue box in Fig. 1) – was nested in a second cross validation cycle (green box Fig. 1) to evaluate the prediction accuracy for new genotypes in unexplored environments. To this aim, one location at a time was excluded from the entire process (steps 1 to 3 in Fig. 1), and cross validation was conducted to derive genomic-predicted model parameters for test (unobserved) genotypes. After all locations were used once as validation environments, genomic predicted parameters for the test set were provided to the crop model to predict HD of unknown genotypes in test (unobserved) locations (step 4, Fig. 1). The two sites in Scotland were regarded as a single location in the cross-validation scheme, given they were very close to each other (Table 1). Through this cross-validation scheme, prediction capabilities of CSMs-GP are tested under conditions completely different from those explored in the training, thanks to the wide range of agro-climatic contexts explored (Table 1). This allows overcoming the limitations of leave-one site-out cross-validation approaches, which can lead to overestimating prediction accuracy for validation environments in case of similarities with those used for the training.

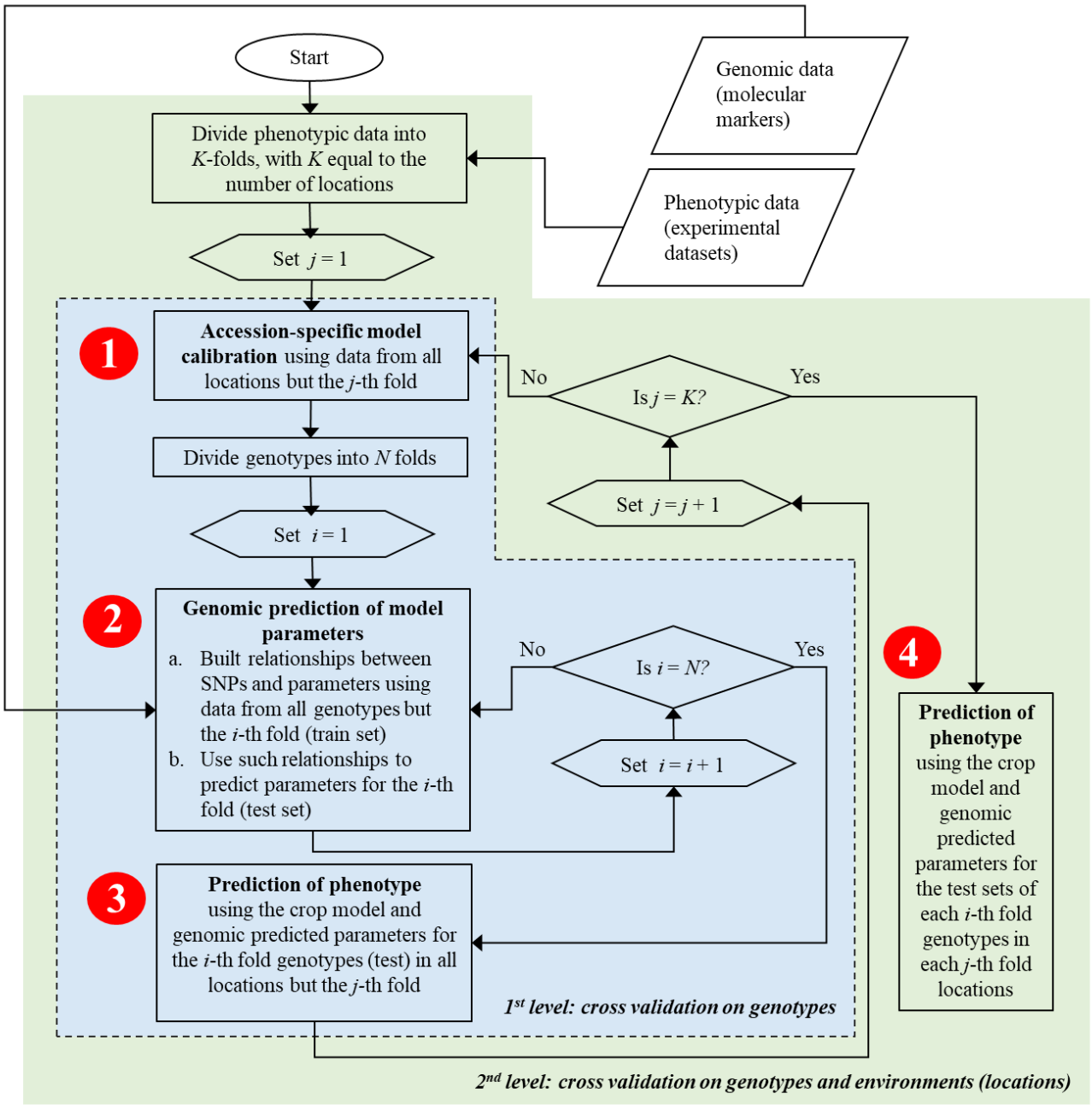

Figure 1. Flow-chart of the cross-validation scheme adopted to evaluate prediction accuracy. The cross-validation to evaluate accuracy for unknown (test) genotypes in known environments (steps 1-3) was nested in a leave-one site-out cross validation (green box), to evaluate accuracy when extending prediction of unknown genotypes in new (untested) environments.

To quantify improvement in prediction accuracy due to dynamic simulations of plant response to environmental factors, we compared results of CSMs-GP with those obtained with conventional GP (i.e., GP without crop models). The reaction norm approach (in particular, the EGW-G×WG×E model) proposed by Jarquin et al. (2014) was used, which includes environmental covariates in the GP statistical model. Average temperature and photoperiod were used as environmental factors, and prediction accuracy was evaluated through the same cross validation scheme adopted for the model-aided GP. The EnvRtype R package (Costa-Neto et al., 2021) was used for deriving environmental covariate matrix and building environmental relationship kernels. The GP model was fitted by using the R package BGLR (de los Campos and Perez-Rodriguez, 2013).

2.4. Extension to future climates

Once the capability of the approach developed to predict HD of new genotypes in new environments was assessed through the cross-validations schemes described in Fig. 1, the prediction domain was extended to future climate projections, for quantifying the genotype-specific response to climate variations. To this end, genomic-predicted model parameters for test accessions were derived by using data from all locations (steps 1 and 2, Fig. 1), and their HD was predicted through the crop model for each location under current (baseline) and future climate conditions (20-year weather series centered on 2004 and 2060, respectively). The climate scenarios were derived from the realization of the shared socio-economic pathway SSP3-7.0 (Meinshausen et al., 2020) provided by the general circulation model IPSL-CM6A-LR (Boucher et al., 2020). The SSP3-7.0 can be regarded as a midway scenario in terms of projected temperature increase as compared to other more optimistic (e.g., SSP1-2.6) or pessimistic (SSP5-8.5) ones (Meinshausen et al., 2020), thus allowing to provide a general picture of genotype-specific impact of climate change on heading date in barley.

Data were statistically downscaled using the stochastic weather generator LARS-WG6 (Semenov and Barrow, 1997) and historical weather series (1995-2014) retrieved from the ERA-5 dataset (Hersbach et al., 2020). This allowed generating 20 seasons of weather data accounting for the year-to-year variability typical of the different sites. To better highlight the effect of climate variations only, no management-based adaptation was considered, with current sowing dates used also for future climate conditions (average sowing date for each site × year combinations; Table S2). For the site in Italy, both autumn and spring sowing times were considered according to current management practices (Table 1), whereas for the sites in Scotland, future climate projections were derived only for Dundee given the closeness with the other site (about 15 km).

3. Results

3.1. Accession-specific model calibration

Sensitivity analysis results (Fig. 2) highlighted a clear North-South gradient in the role of parameters involved with photoperiodic response ($Photo_{sens}$, $Photo_{opt,}$ and $Photo_{crit}$), which explained most of the variability in HD in southern environments. Among those three parameters, $Photo_{sens}$ played the key role, being always ranked first regardless of the environment (Fig. 2). Complementarily, parameters involved with response to temperatures (GDD_{HD}, T_{base}, and T_{max}) increased their relevance while moving towards the North. On average, T_{base} was the parameter with the highest influence on HD among those involved with thermal response, followed by GDD_{HD}. No effect was detected for T_{max}.

According to SA results, the optimization-based decomposition of HD was carried out by focusing on the parameters T_{base} and $Photo_{sens}$, being those with the highest relevance on HD across all environments (mean value of Sobol' total order effect equal to 26% and 33%, respectively; Fig. 2). To avoid a loss of model robustness when moving from the training to the test sets of genotypes due to equifinality (Yang et al., 2022), only the two most relevant parameters were analyzed.

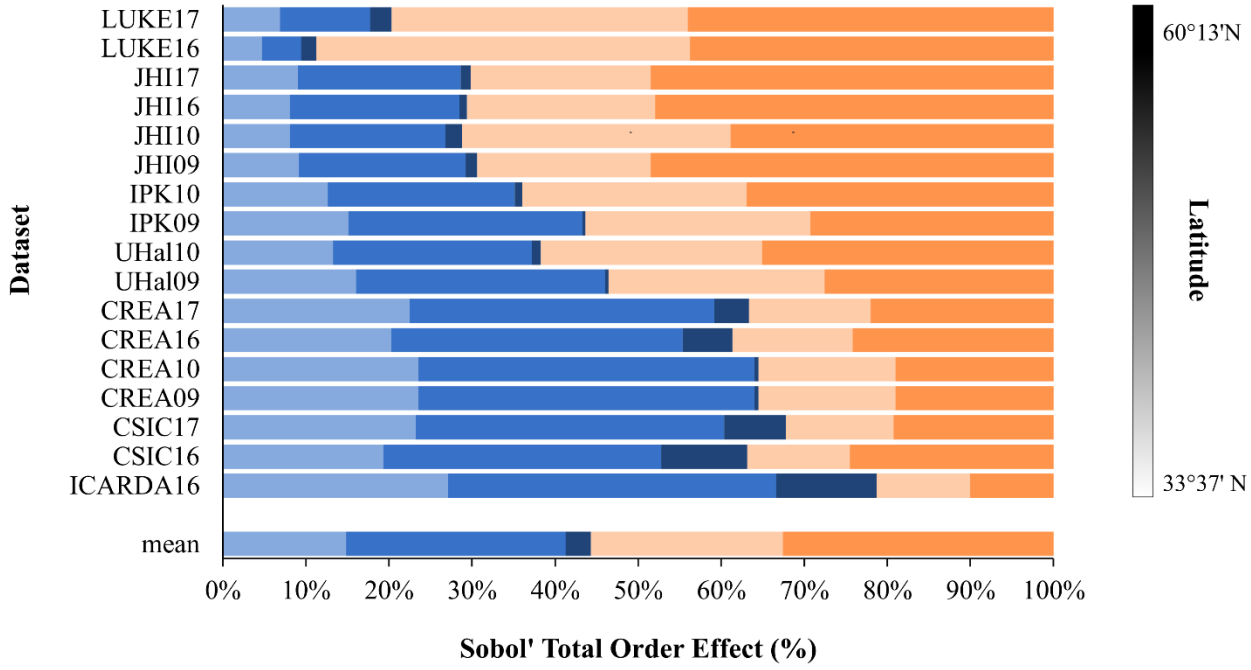

Figure 2. Stacked bar chart showing sensitivity analysis results (Sobol' total order effect, reported as percentage) for days to heading for all datasets (17 combinations of location by season; see Table 1 for details). Datasets are ordered according to the latitude of the site and, within each site, in chronological order. Bars in shades of blue refer to parameters involved with photoperiodic response (light, medium, and dark blue for, respectively, Optimal daylength, Photoperiod sensitivity, and Critical daylength). Bars in shades of orange refer to parameters for response to temperatures (light and medium orange for, respectively, Thermal time requirements to reach heading and Base temperature).

Despite focusing on two parameters only, the decomposition of HD allowed to successfully reproduce G×E×M interactions for heading time across environments, with good agreement between observed and simulated HD for all accessions (Fig. S3). The average R^2 and RMSE for all genotypes

were indeed equal to 0.97 (range 0.95-0.98) and 7.4 days (range between 5.7 to 11.3), respectively, whereas EF ranged between 0.89 to 0.98 (average value of 0.96). Model performance was satisfactory also for the leave-one site-out cross validation, with small variability in the metrics achieved at each cycle. These results are better than – or consistent with – those achieved by other authors comparing different phenological models for wheat (e.g., Wallach et al., 2021) and barley (Kumar et al., 2021), or different optimization methods (Gao et al., 2020). Analysis of residuals (Fig. S4) did not highlight any specific pattern of variability in errors.

3.2. Prediction accuracy in known environments

The simulation of HD in known environments using genomic-predicted model parameters was satisfactory for both the training and the test sets of accessions. Agreement metrics are reported in Table 2, whereas Fig. 3 shows the distribution of two of them (EF and RMSE).

Results obtained for the training set were close to those achieved in the accession-specific optimization, suggesting that prediction errors were partly explained by the incomplete ability of the crop model to fully capture phenological response to environmental factors in some of the combination genotype × environment. Modelling efficiency and R^2 were both close to their optimum, with average values equal to 0.96 and 0.97, respectively, whereas average RMSE was equal to 7.7 days.

Results for the test set of accessions showed a similar prediction accuracy, with an average increase of RMSE of about +0.2 days. Average values of R^2 and EF were almost unchanged, with slight worsening of the latter for few genotypes only (minimum value of test set equal to 0.76 instead of the 0.82 obtained for the training one). Such worsening was mainly due to an overestimation of HD in southern environments with autumn sowing (sites with the largest days to heading, Fig. S2).

Table 2. Agreement between observed and predicted days to heading for training and test sets in known and unknown environments. The mean and the standard deviation (the later between brackets) of results obtained for all accessions are reported.

Dataset	RMSE[a] (days)	EF [b] (-)	CRM [c] (-)	R^2 (-)
Training	7.70 (0.95)	0.96 (0.02)	0.01 (0.02)	0.97 (0.004)
Test in known environments	7.89 (1.2)	0.96 (0.02)	0.01 (0.03)	0.97 (0.01)
Test in unknown environments	9.27 (1.1)	0.94 (0.02)	0.01 (0.03)	0.97 (0.01)

[a] Root mean square error; minimum and optimum = 0, maximum = +∞ (Jørgensen et al., 1986).
[b] Modelling efficiency; maximum and optimum = 1, minimum = -∞ (Nash and Sutcliffe, 1970).
[c] Coefficient of residual mass; minimum = -∞, maximum= +∞, optimum = 0, (Loague and Green, 1991).

3.3. Extending genomic prediction to unknown environments

Results of the leave-one site-out cross validation highlighted the capability of the crop model to correctly simulate HD of unknown genotypes (test set) in unknown environments by using genomic-predicted parameters (Fig. 3).

Prediction accuracy showed just a slight worsening as compared to the model performance under known environments, with an average R^2 for all accessions equal to 0.98. Nash-Sutcliffe EF was always largely positive, with a mean value equal to 0.94 and a small variability (values ranged from 0.74 to 0.97). No systematic over- or under-estimation was observed in general, as highlighted by the mean CRM close to 0 (Table 2), although a slight overestimation was detected for some accessions (minimum value of CRM equal to -0.14). Average RMSE was equal to 9.27 days, with most values ranging between 8.7 days and 9.6 days (25^{th} and 75^{th} percentile, respectively).

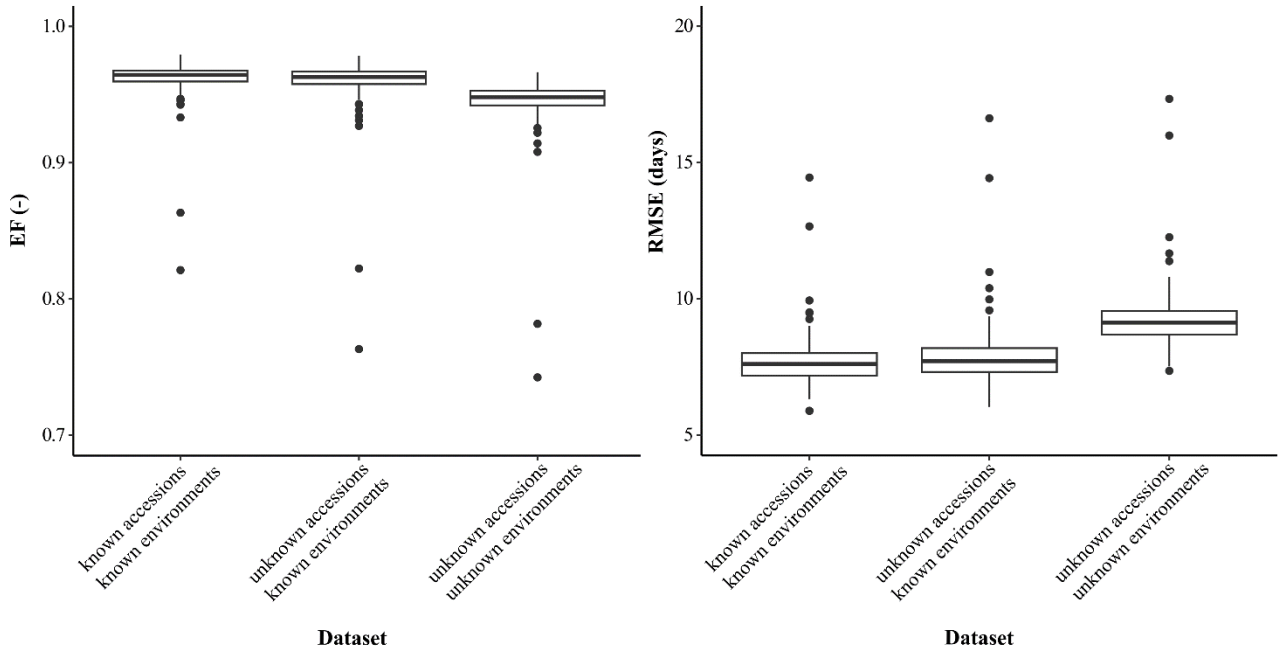

Figure 3. Evaluation of prediction accuracy through cross-validation. Agreement between observed days to heading and those simulated with the crop model and genomic-predicted model parameters for known accessions (training set) and known environment (training environments), and for unknown accessions (test set) under known (training) and unknown (test) environments. Agreement metrics are Nash and Sutcliffe modelling efficiency (EF; Nash and Sutcliffe, 1970; from -∞ to 1, optimum 1) and root mean square error (RMSE, from 0 to +∞, optimum 0) which represents a measure of the agreement between observed and simulated data which weighs more large deviations.

Results obtained without integrating the crop model in GP showed that the model does not increase – with the dataset used in this study – the prediction accuracy during the training and the test of new genotypes in known environments. Actually, the agreement metrics (e.g., EF = 0.99 for test in known environments) achieved were slightly better compared to the CSMs-GP. On the contrary, the crop model demonstrated to be fundamental to properly capture G × E × M interactions when the system is

tested to predict HD of new genotypes in unknown environments. In this case, indeed, excluding the crop model did not allow to explain the variability in observations, to the extent that negative correlations and modelling efficiencies were achieved.

3.4. Extending prediction to climate change scenarios

Figure 4 shows the variation of HD under future climate conditions as compared to the baseline for each site and accession. Given that sowing dates under current and future climate were the same, temperature is the only environmental factor that changes between the two timeframes. As expected, the projected warmer climate led to earlier heading, especially in southern sites with autumn sowing (Italy and Spain) where the average shortening of the HD was around eight days (20-year mean value of all genotypes; Fig. 4). Smaller variation of HD was observed in other environments, ranging from two (German sites) to six days (Finland). The within-site variability in genotype response to future climates was generally similar, with differences of up to seven days among generated seasons within the same site×genotype combination. The smallest variation was observed for the site in Finland, where – regardless of the genotype – the vegetative phase is already very short with the baseline climate (mean HD is 54 days). For this reason, the differences among genotypes in the shortening of the period sowing-HD due to climate change are, in absolute terms (days), proportionally small.

Despite the low complexity of the model, the estimate of the impact of the climate change scenario can be considered as reliable. Indeed, although the model does not consider the negative effect of temperatures above the optimal range, mean air daily temperatures for the climate scenario in the period from sowing to heading exceeded 25°C only in 1 day out of more than 1300 (20 seasons × days from sowing to heading) for the two German sites and in 16 days out of more than 1900 for the spring sowings in the Italian site. This allow considering high temperatures as non-limiting crop development under the selected climate change scenario (Pessotto et al., 2023).

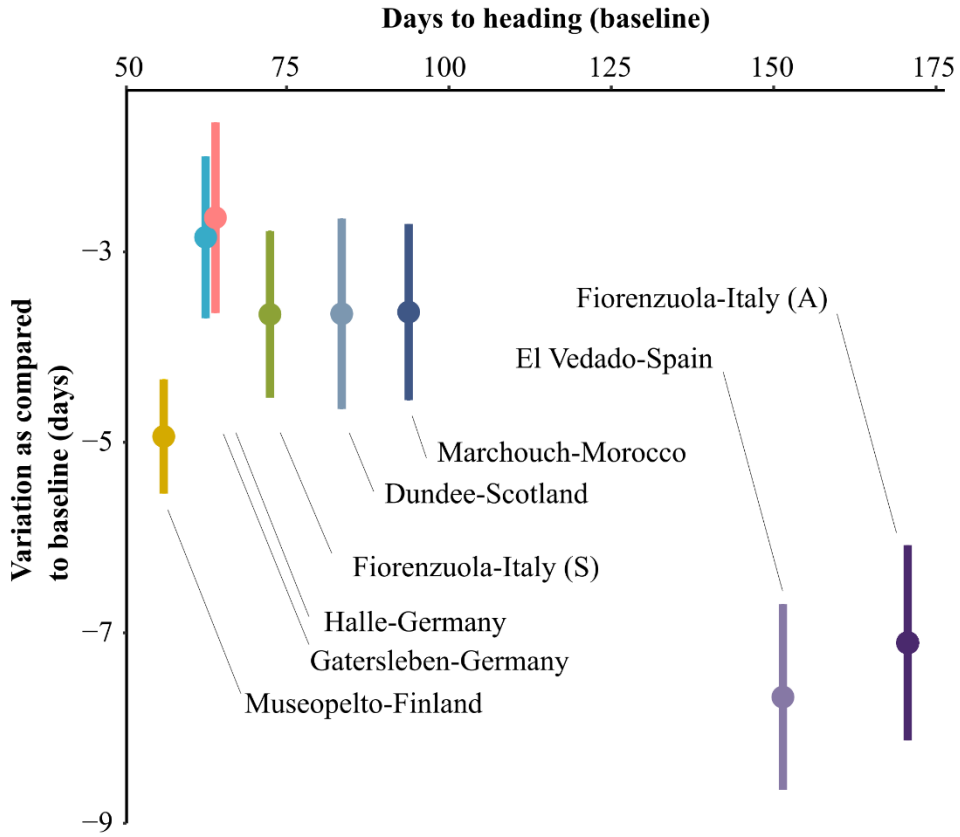

Figure 4. Extending prediction of heading date for unknown genotypes (test set) to future climates. Variation of days to heading (y-axis) predicted under future climate conditions as compared to the baseline (x-axis). For the Italian site, both autumn (A) and spring (S) sowing times were considered according to current management practices. For each site, the mean value (dot) and the interquartile range (line) of heading date simulated for the 151 genotypes in the 20 seasons centred on 2060 are reported. Climate change scenario refers to the shared socio-economic pathway SSP3-7.0 as provided by the general circulation model (GCM) IPSL-CM6A-LR. The 20 seasons were obtained using the LARS-WG6 weather generator to downscale the GCM outputs.

4. Discussion

We investigated the capability of crop models to extend genomic prediction for heading date in barley to environments unexplored during the phenotyping trials. To reach this aim, a dedicated cross validation scheme was developed, following a leave-one site-out strategy with widely heterogeneous sites. According to our knowledge, this is the first time this kind of evaluation has been carried out, given recent studies have only extended the prediction domain to environments close to those explored during training. Kawakita et al. (2024) used data from a single independent site to evaluate prediction accuracy of the integrated CSMs-GP approach in rice. However, all sites considered were in Japan, with a limited variability in environmental conditions as compared to that explored in this study.

The good results achieved for the test sets (unobserved genotypes in untested environments) are in line with those obtained for rice by the same authors (Kawakita et al., 2024), who achieved an average RMSE, MAE, and EF equal to 8.9, 6.9, and 0.82, respectively. They concluded that the use of a process-based crop model to integrate statistical gene-to-phenotype relationships can – at least to some extent – successfully extrapolate G×E×M interactions to new conditions. This would provide new tools to optimize manageable factors (G, M) in the challenge of improving crop productivity under future climate conditions (Cooper et al., 2021).

The integration of GP and crop modelling was carried out by following the methodology used in the multi-model analysis carried out by Yang et al. (2022) on rice. However, that study only focused on known environments, without testing the capability of CSMs-GP approaches to extend the prediction to new conditions. Moreover, the agro-climatic variability explored in that study was limited as compared to the dataset used here, given it referred to three sites in two countries (Madagascar and Senegal) for a total of 8-10 observations for each genotype. Nevertheless, one of the main limits of this study as compared to multi-model analyses is the lack of quantification of the impact of model structure on

prediction accuracy, which can largely vary according to the model used (Yang et al., 2022). Further studies are needed to compare results achieved with alternative modelling approaches, to identify potential criteria for model choice and configuration (number of predicted parameters) when targeting model-aided GP.

A key point in our results is the limited loss of prediction accuracy while moving from the training to the test set, in line with results obtained with a similar approach in rice (Yang et al., 2022; Kawakita et al. 2024). This might be related with the low ratio between number of model parameters optimized while decomposing performance traits as compared to the number of observations available for each accession, which should reduce the uncertainty in parameter estimation due to equifinality issues (Yang et al., 2022). Kawakita et al. (2024), tested four phenological models, with a large number of observations per genotype (ranging from 20 to 70) as compared to the number of parameters considered (ranging from 3 to 5). In the multi-model study of Yang et al. (2022), the model with the lowest ratio (3 parameters for 8 to 10 observations for each accession) was one of those showing the most stable performance between training and test sets. In this study we considered only two parameters (independent and with the largest overall sensitivity; Fig. 2), with a number of observations for each accession equal to 17. Besides sensitivity analysis results, this was due to the fact that tests conducted by including three parameters (3-parameter configuration) allowed just a minor increase in the system capability to predict heading dates, whereas the correlation between optimized parameters values and genomic-predicted ones worsened likely due to equifinality issues, increasing the risk of losing robustness. In particular, the correlation between optimized and predicted parameter values with the 2-parameter configuration (mean Pearson's coefficient equal to 0.93 and 0.89 for T_{base} and $Photo_{sens}$, respectively, for the training dataset; corresponding values for the test dataset were 0.45 and 0.46) was consistent with what achieved by Kawakita et al. (2024), whereas the 3-parameter configuration led to mean correlation values for the test dataset equal to 0.08 and 0.22 for the

parameters T_{base} and GDD_{HD}, which are obviously not independent. To a certain extent – however - equifinality issues also derive from the partial unsuitability of models based on the photo-thermal rate of development for properly capturing all the biological processes involved with phenological development (Kumar et al., 2009). Multi-model studies were carried out by Yang et al. (2022) and Kawakita et al. (2024), but they focused on different populations and genomic data, making new dedicated studies fundamental to quantify the impact of uncertainty in parameter estimation (function of model structure and on the number of parameters used within the same model) on prediction robustness.

Other aspects that can be further investigated deal with the effect of population size/structure and marker density on prediction accuracy. Evaluating the impact of different marker densities on prediction accuracy would be crucial to optimize costs associated with genotyping (Abed et al., 2018). Similarly, understanding the role of population size and composition would allow optimizing phenotyping costs (e.g., Heffner et al., 2011). While different authors have addressed these issues in the context of conventional genomic prediction (e.g., Lorenzana and Bernardo, 2009; Isidro et al., 2015; Wu et al., 2023) little is known about the role of these factors in case of CSMs-GP frameworks. In this study we considered an unstructured, barley population of elite cultivars genotyped with a 50k marker density. Only a gradient of differentiation between older and more modern barley cultivars was present as a result of directional selection on specific breeding targets. Spurious associations can occur when there is a correlation between the population structure and the phenotype (e.g., Larsson et al., 2013). While old and modern spring barley cultivars clearly differ for improvement traits such as plant height and grain yield, this is not the case for flowering time (Fig. S5). The absence of population structure in the panel we used was expected, given that relevant factors known to affect population structure (geographical origin, row type, and growth habit; Milner et al., 2019) were not segregating in our genotypes, being all two-rows, spring barley cultivars from Europe. Results achieved by Yang et al.

(2022) were based on a high-density array (700k SNPs), although lower marker densities or pre-selected markers were also successfully used in other studies (Onogi et al., 2016; Dingkhun et al., 2017; Kawakita et al., 2024). Tests with more structured germplasm, including cultivars with different origin, row-type, growth habit (Milner et al., 2019) or specifically developed populations would also shed light on the possibility to extend prediction to accessions with a more heterogeneous genetic background.

The causes of the unsatisfactory performance for few (less than 4%) accessions need to be further investigated, possible explanations referring to their genetic makeup or with the inability of the crop model to capture genotype-specific physiological responses to environmental factors. This is supported by the fact that the accessions showing the worst results in terms of prediction accuracy were also those affected by the largest error in the very first step, where accession-specific model parameters were derived. Moreover, the two genotypes with the worst performance (Mona and Saana) were the only two in the panel carrying a mutant allele in a major flowering time gene *HvELF3* (Faure et al., 2012), which causes early flowering regardless of the daylength, especially in Southern environments (Montardit-Tarda et al.; 2024). Such mutation induces indeed a long photoperiod response even under short daylengths. As a results, these genotypes are very early under the short days of the autumn sowings in the South and much less in the North, where any long day response is saturated for all genotypes because of late spring sowings (Montardit-Tarda et al., 2024). Given the mutated allele is rare (only Mona and Saana in this dataset), such peculiar response to photoperiod cannot be captured by the model-aided GP. This led to a slight overestimation of simulated HD in Mediterranean locations with autumn sowing, where the wide range of daylength experienced by the crop exacerbates the impact of uncertainty in photoperiod-related parameters on prediction error, with subsequent slight worsening of agreement metrics (e.g., EF). The collection of data to formalize this peculiar response to photoperiod in model algorithms will allow to extend the domain of model- aided GP to a larger pool of genotypes.

Comparison with results of prediction obtained without integrating crop models highlighted large benefits in terms of accuracy for the integrated approach when new environments were explored. This is likely due to non-linearities in plant response to environmental conditions which make the dynamic simulation of the underlying biological processes a crucial point when targeting a wide range of conditions as in this study, where major environmental effects and G×E×M interactions are expected. These results are in line with those reported by other authors, although the extent of the crop model-induced prediction gain presents a certain variability. Onogi et al. (2016) found that integrating crop models and GP for predicting heading date in rice provided more accurate results than purely statistical GP approaches, but with smaller benefits as compared to those found here (Pearson's correlation coefficient increased from 0.87 of conventional GP to 0.92 of integrated GP with the same approach used in this study). This is likely due to the fact that the environments explored by Onogi et al. (2016) – nine environments (site × year) in Japan – were less heterogeneous than those considered in this study. Marked improvements in prediction accuracy were instead found by Messina et al. (2018) for maize yield when a variety of environments were derived by applying a wide range of irrigation treatments. In that case, the large G×E×M interactions could not be captured without integrating a biophysical simulation of yield response in GP models, leading to accuracy gains similar to those achieved in this study.

5. Conclusions

We evaluated the capability of crop models to extend the domain of GP to unexplored environments through the explicit, process-based simulation of G×E×M interactions.

- A dedicated validation scheme was developed, with cross-validation on genotypes – testing prediction accuracy for unknown genotypes in known environments – nested in a leave-one site-out cross validation, to evaluate predictive ability for unknown genotypes in unexplored

environments. The large variability among sites (multi-environmental trials were spread from northern Africa to northern Europe) allowed evaluating the integrated crop modelling-GP framework under extremely different thermal and pluviometric regimes.

- The good performances achieved open new avenues for the role of crop models in supporting GP and plant breeding. This support is due to the fact that, according to the framework we used, GP is used to relate genomic information to functional traits (e.g., photoperiod sensitivity), whereas the crop model is used to interpret the effect of the interaction between functional traits and E×M on the target performance trait (HD). The use of statistics in the step of the prediction process less affected by G×E×M interactions allows extending GP to unexplored environment, thanks to the crop model capability to reproduce plant responses to environmental and management factors on an ecophysiological basis. This allows large scale screening of genotyped populations to evaluate their phenotypic response under a wide range of conditions, including future climates. On the other hand, it provides insights about the possibility of integrating molecular marker data in simulation models to provide the latter with the capability of carrying out genotype-specific assessment of crop performances. This would allow increasing the potential of model-based analyses of alternative management and climate change scenarios, improving the quantification of crop responses in light of the existing germplasms.
- The target trait in this study was HD in spring barley, but the same methodology can be applied to other crops, different processes/traits and alternative crop modelling/GP approaches. Crops for which climate change impacts are expected to be negative in most regions (e.g., maize; Rezaei et al., 2023) and traits involved with tolerance to abiotic stressors (e.g., heat stress) – which are expected to become more frequent in the coming decades but whose impacts on crops is largely affected by multiple interacting factors (e.g., crop susceptibility according to the

phenological stage) – would particularly benefit from the integration of genomic selection schemes with crop-model-aided mid-term predictions, to support crop adaptation via G×E×M optimization.

- Future efforts will be focused on applying the integrated crop modelling-GP framework to other traits (e.g., yield), and on investigating the role of factors that might affect prediction accuracy, such as population size and structure, markers density, GP approaches and characteristics of the crop model, as well as the ratio between number of model parameters involved in the analysis and number of observations for each accession.

Acknowledgments

This study was funded by the project BARISTA – carried out under the ERA-NET Cofund SusCrop (Grant N°771134) being part of the Joint Programming Initiative on Agriculture, Food Security and Climate Change (FACCE-JPI) – and the project RECOBAR, Grant PCI2022-135024-2 funded by MCIN/AEI/10.13039/501100011033 and by European Union NextGenerationEU/PRTR. The authors also gratefully acknowledge the projects ExBarDiv and ClimBar for the data provided, and the computational resources provided by the INDACO Platform (project of High Performance Computing at the University of Milan).

Conflict of interest

The authors declare they have no conflict of interest.

References

Abed, A., Pérez-Rodriguez, P., Crossa, J., Belzile, F., 2018. When less can be better: How can we make genomic selection more cost-effective and accurate in barley? Theor. Appl. Genet. 131, 1873-1890.

Bayer, M.M., Rapazote-Flores, P., Ganal, M., Hedley, P.E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P.D., Thomas, W., Waugh, R., 2017. Development and evaluation of a barley 50k iSelect SNP array. Front. Plant Sci. 8:1792.

Bernardo, R., 2008. Molecular markers and selection for complex traits in plants: Learning from the last 20 years. Crop Sci. 48, 1649–1664.

Bernardo, R., 2016. Bandwagons I, too, have known. Theor. Appl. Genet. 129, 2323–2332.

Bernardo, R., 2020. Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. Heredity, 125, 375–385.

Boote, K.J., Jones, J.W., White, J.W., Asseg, S., Lizaso, J.I., 2013. Putting mechanisms into crop production models. Plant Cell Environ. 36, 1658–1672.

Boucher, O., Servonnat, J., Albright, A.L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.Y., Guenet, B., Guez, Lionel, E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P.,

562 Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw,
563 D., Thiéblemont, R., Traore, A.K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., Vuichard,
564 N., 2020. Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. J. Adv. Model.
565 Earth Syst. 12, e2019MS002010.
566 Bretani, G., Shaaf, S., Tondelli, A., Cattivelli, L., Delbono, S., Waugh, R., Thomas, W., Russell, J.,
567 Bull, H., Igartua, E., Casas, A.M., Gracia, P., Rossi, R., Schulman, A.H., Rossini, L. 2022. Multi-
568 environment genome-wide association mapping of culm morphology traits in barley. Front. Plant
569 Sci. 13:926277.
570 Confalonieri, R., Rosenmund, A.S., Baruth, B., 2009. An improved model to simulate rice yield.
571 Agron. Sustain. Dev. 29, 463–474.
572 Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., Acutis, M., 2010. Comparison of
573 sensitivity analysis techniques: a case study with the rice model WARM. Ecol. Model. 221,
574 1897-1906.
575 Confalonieri, R., Bregaglio, S., Acutis, M., 2010. A proposal of an indicator for quantifying model
576 robustness based on the relationship between variability of errors and of explored conditions.
577 Ecol. Model. 221, 960-964.
578 Cooper, M., Gho, C., Leafgren, R., Tang, T., Messina, C.D., 2014. Breeding drought-tolerant maize
579 hybrids for the US corn-belt: Discovery to product. J. Exp. Bot. 65, 6191–6204.
580 Cooper, M., Voss-Fels, K.P., Messina, C.D., Tang, T., Hammer, G.L., 2021. Tackling G×E×M
581 interactions to close on-farm yield-gaps: creating novel pathways for crop improvement by
582 predicting contributions of genetics and management to crop productivity. Theor. Appl. Genet.
583 134, 1625–1644.
584 Costa-Neto, G., Galli, G., Carvalho, H.F., Crossa, J., Fritsche-Neto, R., 2021. EnvRtype: a software to
585 interplay enviromics and quantitative genomics in agriculture. G3, 11, jkab040.

de los Campos, G., Perez-Rodriguez, P., 2013. BGLR: Bayesian generalized linear regression. R package version. http://R-Forge.R-project.org/projects/bglr/

Diepenbrock, C.H., Tang, T., Jines, M., Technow, F., Lira, S., Podlich, D., Cooper, M., Messina, C., 2022. Can we harness digital technologies and physiology to hasten genetic gain in US maize breeding? Plant Physiol. 188, 1141–1157.

Dingkuhn, M., Pasco, R., Pasuquin, J.M., Damo, J., Soulie, J.C., Raboin, L.M., Dusserre, J., Sow, A., Manneh, B., Shrestha, S., Balde, A., Kretzschmar, T., 2017. Crop-model assisted phenomics and genome-wide association study for climate adaptation of indica rice. 1. Phenology. J. Exp. Bot. 68, 4369–4388.

Duan, Q., Sorooshian, S., Gupta, V.K., 1992. Effective and efficient global optimization for conceptual rainfall runoff models. Water Resour. Res. 28, 1015–1031.

Endelman, J. B., 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome-US, 4, 250–255.

Faure, S., Turner, A.S., Gruszka, D., Laurie, D., 2012. Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. PNAS, 109, 8328-8333.

Gao, Y., Wallach, D., Liu, B., Dingkuhn, M., Boote, K., Singh, U., Asseng, S., Kahveci, T., He, J., Zhang, R., Confalonieri, R., Hoogenboom, G., 2020. Comparison of three calibration methods for modeling rice phenology. Agr. Forest Meteorol., 280, 107785.

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., Chapman, S., Podlich, D., 2006. Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci. 11, 587–593.

Heffner, E.L., Sorrells, M.E., Jannink, J.L., 2009. Genomic selection for crop improvement. Crop Sci. 49, 1–12.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J. Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J-N., 2020. The ERA5 global reanalysis. Q. J. Roy. Meteor. Soc. 146, 1999–2049.

Heslot, N., Akdemir, D., Sorrells, M.E., 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.

Isidro, J., Jannink, J.L., Akdemir, D., Poland, J., Heslot, N., Sorrells, M.E., 2015. Training set optimization under population structure in genomic selection. Theor. Appl. Genet. 128, 145–158.

Jarquin, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Largeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., de los Campos, G., 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 127, 595-607.

Jørgensen, S.E., Kamp-Nielsen, L., Christensen, T., Windolf-Nielsen, J., Westergaard, B., 1986. Validation of a prognosis based upon a eutrophication model. Ecol. Model. 32, 165–182.

Kawakita, S., Yamasaki, M., Teratani, R., Yabe, S., Kajiya-Kanegae, H., Yoshida, H., Fushimi, E., Nakagawa, H., 2024. Dual ensemble approach to predict rice heading date by integrating multiple rice phenology models and machine learning-based genetic parameter regression models. Agr. Forest Meteorol. 344, 109821.

Kumar, S.R., Hammer, G.L., Broad, I., Harland, P., McLean, G., 2009. Modelling environmental effects on phenology and canopy development of diverse sorghum genotypes. Field Crop. Res. 111, 157-165.

Kumar, U., Morel, J., Bergkvist, G., Palosuo, T., Gustavsson, A.-M., Peake, A., Brown, H., Ahmed, M., Parsons, D., 2021. Comparative Analysis of Phenology Algorithms of the Spring Barley Model in APSIM 7.9 and APSIM Next Generation: A Case Study for High Latitudes. Plants, 10, 443.

Larsson, S.J., Lipka, A.E., Buckler E.S., 2013. Lessons from Dwarf8 on the Strengths and Weaknesses of Structured Association Mapping. PlosONE, 9, e1003246.

Litrico, I., Violle, C., 2015. Diversity in Plant Breeding: A New Conceptual Framework. Trends Plant Sci. 20, 604–613.

Loague, K.M., Green, R.E., 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. J. Contam. Hydrol. 7, 51–73.

Lorenzana, R., Bernardo, R., 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor. Appl. Genet. 120, 151-161.

Meinshausen, M., Nicholls, Z.R.J., Lewis, J., Gidden, M.J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J.G., Daniel, J.S., John, A., Krummel, P.B., Luderer, G., Meinshausen, N., Montzka, S.A., Rayner, P.J., Reimann, S., Smith, S.J., Van Den Berg, M., Velders, G.J.M., Vollmer, M.K., Wang, R.H.J., 2020. The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. Geosci. Model Dev. 13, 3571–3605.

Messina C.D., Jones, J.W., Boote, K.J., Vallejos, C.E., 2006. A gene-based model to simulate soybean development and yield response to environment. Crop. Sci. 46, 456–466.

Messina, C.D., Podlich, D., Dong, Z., Samples, M., Cooper, M., 2011. Yield-Trait performance landscapes: From theory to application in breeding maize for drought tolerance. J. Exp. Bot. 62, 855–868.

Messina, C.D., Technow, F., Tang, T., Totir, R., Gho, C., Cooper, M., 2018. Leveraging biological insight and environmental variation to improve phenotypic prediction: Integrating crop growth models (CGM) with whole genome prediction (WGP). Eur. J. Agron. 100, 151–162.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157, 1819–1829.

Milner, S.G., Jost, M., Taketa, S., Mazón, E.R., Himmelbach, A., Oppermann, M., Weise, S., Knüpffer, H., Basterrechea, M., König, P., Schüler, D., Sharma, R., Pasam, R.K., Rutten, T., Guo, G., Xu, D., Zhang, J., Herren, G., Müller, T., Krattinger, S.G., Keller, B., Jiang, Y., González, M.Y., Zhao, Y., Habekuß, A., Färber, S., Ordon, F., Lange, M., Börner, A., Graner, A., Reif, J.C., Scholz, U., Mascher, M., Stein, N., 2019. Genebank genomics highlights the diversity of a global barley collection. Nat. Genet. 51, 319–326.

Montardit-Tarda, F., Casas, A.M., Thomas, W.T.B., Schnaithmann, F., Sharma, R., Shaaf, S., Campoli, C., Russel, J., Ramsay, L., Delbono, S., Jääskeläinen, M., Maitry, P., Stoddard, F.L., Visioni, A., Flavell, A.J., Pillen, K., Kilian, B., Graner, A., Rossini, L., Waugh, R, Cattivelli, L., Shulman, A.H., Tondelli, A., Igartua, E., 2024. New loci and candidate genes in spring two-rowed barley detected through meta-analysis of a field trial European network. Submitted to Theoretical and Applied Genetics, available as preprint on https://doi.org/10.1101/2024.09.04.611234.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models Part I. A discussion of principles. J. Hydrol. 10, 282–290.

677 Onogi, A., Watanabe, M., Mochizuki, T., Hayashi, T., Nakagawa, H., Hasegawa, T., Iwata, H., 2016.
678 Toward integration of genomic selection with crop modelling: the development of an integrated
679 approach to predicting rice heading dates. Theor. Appl. Genet. 129, 805–817.
680 Paleari, L., Cappelli, G., Bregaglio, S., Acutis, M., Donatelli, M., Sacchi, G.A., Lupotto, E., Boschetti,
681 M., Manfron, G., Confalonieri, R., 2015. District specific, in silico evaluation of rice ideotypes
682 improved for resistance/tolerance traits to biotic and abiotic stressors under climate change
683 scenarios. Climatic Change, 132, 661–675.
684 Paleari, L., Movedi, E., Cappelli G., Wilson T.L., Confalonieri, R., 2017. Surfing parameter
685 hyperspaces under climate change scenarios to design future rice ideotypes. Glob. Change Biol.
686 23, 4651–4662.
687 Paleari, L., Li, T., Yang, Y., Wilson, L.T., Hasegawa, T., Boote, K.J., Buis, S., Hogenboom, G., Gao,
688 Y., Movedi, E., Ruget, F., Singh, U., Stöckle, C.O., Tang, L., Wallach, D., Zhu, Y., Confalonieri,
689 R., 2022. A trait-based model ensemble approach to design rice plant types for future climate.
690 Glob. Change Biol. 28, 2689–2710.
691 Pessotto, M.V., Roberts, T.L., Bertucci, M., dos Santos, C., Ross, J., Savin, M., 2023. Determining
692 cardinal temperatures for eight cover crop species. Agrosyst. Geosci. Environ. 6, e20393.
693 R Core Team., 2022. R: A language and environment for statistical computing. R Foundation for
694 Statistical Computing. https://www.r-project.org/
695 Reynolds, M.P., Braun H.J., Cavalieri, A.J., Chapotin, S., Davies, W.J., Ellul, P., Feuillet, C., Govaerts,
696 B., Kropff, M.J., Lucas, H., Nelson, J., Powell, W., Quilligan, E., Rosegrant, M.W., Singh, R.P.,
697 Sonder, K., Tang, H., Visscher, S., Wang, R., 2017. Improving global integration of crop
698 research. Science, 357, 359–360.
699 Rezaei, E.E., Webber, H., Asseng, S., Boote, K., Durand, J.L., Ewert, F., Martre, P., MacCarthy, D.S.,
700 2023. Climate change impacts on crop yields. Nat. Rev. Earth Environ. 4, 831–846.

Schrauf, M.F., de los Campos, G., Munilla, S., 2021. Comparing Genomic Prediction Models by Means of Cross Validation. Front. Plant Sci. 12, 734512.

Semenov, M.A., Barrow, E.M., 1997. Use of a stochastic weather generator in the development of climate change scenarios. Clim. Change 35, 397–414.

Snowdon, R.J., Wittkop, B., Chen, T.W., Stahl, A., 2020. Crop adaptation to climate change as a consequence of long-term breeding. Theor. Appl. Genet. 134, 1613–1623.

Sobol', I. M., 1993. Sensitivity estimates for non-linear mathematical models. Math. Comput. Model. 14, 407–414.

Tondelli, A., Xu, X., Moragues, M., Sharma, R., Schnaithmann, F., Ingvardsen, C., Manninen, O., Comadran, J., Russell, J., Waugh, R., Schulman, A.H., Pillen, K., Rasmussen, S.K., Kilian, B., Cattivelli, L., Thomas, W.T.B., Flavell, A.J. 2013. Structural and temporal variation in genetic diversity of European spring two-row barley cultivars and association mapping of quantitative traits. Plant Genome-US, 6, 1–14.

van Diepen, C.A.; Wolf, J.; van Keulen, H.; Rappoldt, C., 1989. WOFOST: A Simulation Model of Crop Production. Soil Use Manag. 5, 16–24.

Van Etten, J., de Sousa, K., Cairns, J.E., Dell'Acqua, M., Fadda, C., Guereña, D., van Heerwaarden, J., Assefa, T., Manners, R., Müller, A., Pé, M.E., Polar, V., Ramirez-Villegas, J., Solberg. S. Ø., Teeken, B., Tufan, H.A., 2023. Data-driven approaches can harness crop diversity to address heterogeneous needs of breeding products. P. Natl. Acad. Sci. USA 120, e2205771120.

VSN International (2022). Genstat for Windows 22nd Edition. VSN International, Hemel Hempstead, UK. Web page: Genstat.co.uk

Wallach, D., Palosuo, T., Thorburn, P., Hochman, Z., Andrianasolo, F., Asseng, S., Basso, B., Buis, S., Crout, N., Dumont, B., Ferrise, R., Gaiser, T., Gayler, S., Hiremath, S., Hoek, S., Horan, H., Hoogenboom, G., Huang, M., Jabloun, M., Jansson, P.E., Jing, Q., Justes, E., Kersebaum, K.C.,

Launay, M., Lewan, E., Luo, Q., Maestrini, B., Moriondo, M., Olesen, J.E., Padovan, G., Poyda, A., Priesack, E., Pullens, J.W.M., Qian, B., Schütze, N., Shelia, V., Souissi, A., Specka, X., Kumar Srivastava, A., Stella, T., Streck, T., Trombi, G., Wallor, E., Wang, J., Weber, T.K.D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao, C., Zhu, Y., Seidel, S.J., 2021. Multi-model evaluation of phenology prediction for wheat in Australia. Agric. For. Meteorol. 298–299, 108289.

Wu, P.Y., Ou, J.H., Liao, C.T., 2023.Sample size determination for training set optimization in genomic prediction. Theor. Appl. Genet. 136, 57.

Yang, Y., Wilson, L.T., Li, T., Paleari, L., Confalonieri, R., Zhu, Y., Tang, L., Qiu, X., Tao, F., Chen, Y., Hoogenboom, G., Boote, K.J., Gao, Y., Onogi, A., Nakagawa, H., Yoshida, H., Yabe, S., Dingkuhn, M., Lafarge, T., Hasegawa, T., Wang, J., 2022. Integration of genomics with crop modeling for predicting rice days to flowering: A multi-model analysis. Field Crop. Res. 276, 108394.

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Res. 14, 415-421.