Received: 28 December 2023 | Revised: 8 April 2024 | Accepted: 22 April 2024

DOI: 10.1002/fes3.544

Food and Energy Security Open Access WILEY

ORIGINAL ARTICLE

Harnessing enviromics to predict climate-impacted high-profile traits to assist informed decisions in agriculture

Bosen Zhang[1] | **Amber L. Hauvermale**[1] | **Zhiwu Zhang**[1] | **Alison Thompson**[2] | **Clark Neely**[1] | **Aaron Esser**[1] | **Michael Pumphrey**[1] | **Kimberly Garland-Campbell**[1,2] | **Jianming Yu**[3] | **Camille Steber**[1,2] | **Xianran Li**[1,2]

[1]Department of Crop and Soil Sciences, Washington State University, Pullman, Washington, USA

[2]USDA-ARS, Wheat Health, Genetics, and Quality Research Unit, Pullman, Washington, USA

[3]Department of Agronomy, Iowa State University, Ames, Iowa, USA

Correspondence

Camille Steber and Xianran Li, USDA-ARS, Wheat Health, Genetics, and Quality Research Unit, Pullman, WA 99164, USA.

Email: camille.steber@usda.gov and xianran.li@usda.gov

Funding information

Foundation for Food and Agriculture Research, Grant/Award Number: 22-000210; National Institute of Food and Agriculture, Grant/Award Number: 2021-67013-33833; Agricultural Research Service, Grant/Award Number: 2090-21000-039-00D

Abstract

Modern agriculture is a complex system that demands real-time and large-scale quantification of trait values for evidence-based decisions. However, high-profile traits determining market values often lack high-throughput phenotyping technologies to achieve this objective; therefore, risks of undermining crop values through arbitrary decisions are high. Because environmental conditions are major contributors to performance fluctuation, with the contemporary informatics infrastructures, we proposed enviromic prediction as a potential strategy to assess traits for informed decisions. We demonstrated this concept with wheat falling number (FN), a critical end-use quality trait that significantly impacts wheat market values but is measured using a low-throughput technology. Using 8 years of FN records from elite variety testing trials, we developed a predictive model capturing the general trend of FN based on biologically meaningful environmental conditions. An explicit environmental index that was highly correlated ($r = 0.646$) with the FN trend observed from variety testing trials was identified. An independent validation experiment verified the biological relevance of this index. An enviromic prediction model based on this index achieved accurate and on-target predictions for the FN trend in new growing seasons. Two applications designed for production fields illustrated how such enviromic prediction models could assist informed decision along the food supply chain. We envision that enviromic prediction would have a vital role in sustaining food security amidst rapidly changing climate. As conducting variety testing trials is a standard component in modern agricultural industry, the strategy of leveraging historical trial data is widely applicable for other high-profile traits in various crops.

KEYWORDS

climate change, enviromic prediction, falling number, food supply chain

1 | INTRODUCTION

Modern agricultural industry is a sophisticated system encompassing research and breeding programs, crop field production, trade and insurance organizations, processing factories, and end users (Council, 2015; Motes, 2010). Together, these sectors compose the food supply chain. For commodity crops, the chain begins with developing elite varieties through research and breeding programs. Seeds of elite varieties are then sold and planted in growers' production fields. Harvested grain from different production fields are transported and sold to elevators, which serve as central hubs in the supply chain. Grain stored in elevators are distributed to manufacturing factories for processing. Final processed products are then distributed to end users through wholesale and retail.

Sustaining a complex system depends on timely and objective decisions (Lempert, 2002). To make informed decisions maximizing profits, accurate assessment of traits, especially high-profile ones determining crop market values, is of paramount importance. Crop performance, jointly determined by internal genetic constitutes, external environmental conditions, and their interactions, fluctuates because of varied weather conditions and management practices across production fields. High-throughput, efficient, and cost-effective technologies to measure and monitor traits are ideal to assist decision-making (Watt et al., 2020; Yang et al., 2020). However, current assessments of most high-profile traits remain low-throughput, time-consuming, and expensive. Without critically needed trait information being delivered in time, arbitrarily called decisions are inevitable, which poses significant risks of undermining crop value. Until high-throughput phenotyping methods are developed, developing predictive models capturing performance trends could be a potential strategy to assist the decision-making process.

The term "envirome" is coined to define the total sets of environmental conditions affecting performance or state of an organism (Anthony et al., 1995). As rapidly changing climatic conditions threaten food security, enviromics is emerging as an essential component to be integrated into breeding strategies for crop improvement (Cooper & Messina, 2021, 2023; Crossa et al., 2021; Guo & Li, 2023; Resende et al., 2021; Xu, 2016; Xu et al., 2022). Identification of key environmental indices from the envirome allows for the explaination of complex patterns across environments (Guo et al., 2020; Li et al., 2018, 2021; Mu et al., 2021). Genomic prediction integrated with enviromic prediction could assist in breeding elite varieties capable of adapting to varied environmental conditions (Millet et al., 2019). Meanwhile, incorporating enviromic prediction for other sections along the food supply chain remains to be explored.

Wheat (*Triticum aestivum* L.) is the most traded food-use commodity. End-use quality traits are of great importance in determining usage and trading price. Falling number (FN), quantifying the degree of starch degradation by endogenous α-amylases in wheat grain, is one such high-profile trait (Steber, 2017). Grain with elevated α-amylase activities tend to have lower FN readings. Milled flour from low FN wheat grain produces low-quality baked goods and foods (Figure S1). Varieties and management practices that reduce the risk of low FN are desired by growers to avoid the industry discount that reduces the sale price. Furthermore, because of the catalytic nature of α-amylases, a small amount of low FN grain will ruin a large quantity of high FN grain if mixed. Separate storage of low and high FN grain at elevators is critical, but these decisions must be made quickly when tons of grain are received from different production fields in a short period during harvest season.

Low FN is triggered by either a pre-harvest sprouting (PHS) or late maturity α-amylase (LMA) event (Hu et al., 2022; Mares & Mrva, 2014). PHS could be potentially identified via grain morphological changes upon careful inspections, while LMA has not been associated with visible indicative grain feature changes. Without reliable visible diagnostic features, samples of harvested wheat grain from each field must be quantified by the standard Hagberg-Perten method for grading. This method, established in the 1960s, is time consuming and can only be done in laboratories equipped with expensive instruments operated by trained specialists (Figure S1). During harvest season, massive quantities of grain samples are sent to certified laboratories for quantifying FN. The overwhelming workload and complexity of the test procedure prolongs the final sale price determination and, most importantly, fails to provide storage or blending guidance for elevators. Forecast models capturing trait trends would be beneficial to overcome such informatics bottlenecks.

Here, we proposed and demonstrated the enviromic prediction strategy for informed decisions by using wheat FN as an example (Figure 1). Leveraging long-term variety testing trial records, we trained an enviromic prediction model capturing the general FN trend from production fields. Two applications were further developed to depict how such a predictive model can guide informed decisions in the food supply chain.

2 | MATERIALS AND METHODS

2.1 | Contemporary low throughput falling number (FN) test procedure

FN is a critical measurement for grain quality, which not only determines the final sale price but also guides

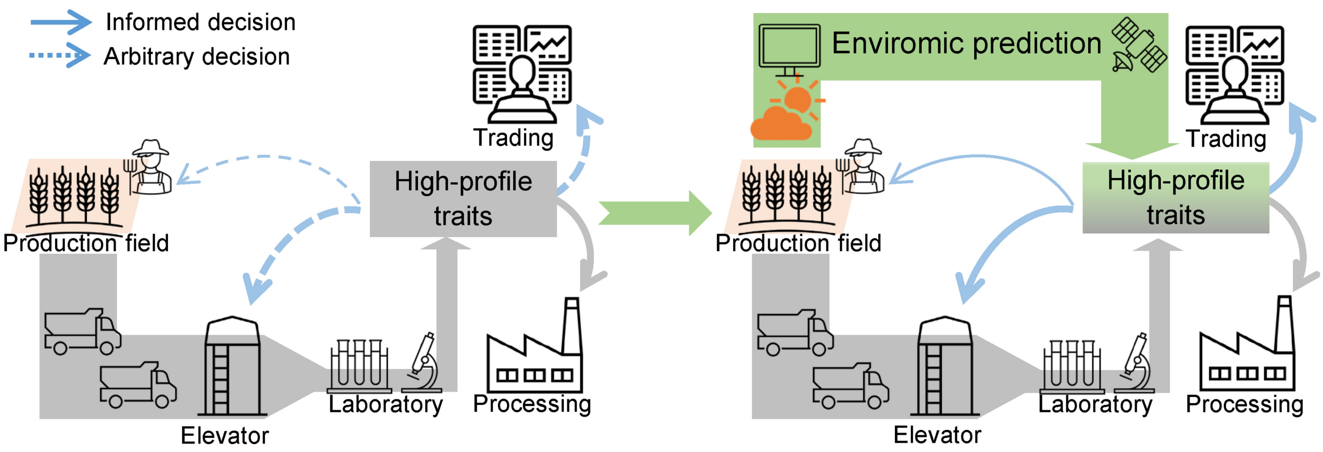

FIGURE 1 Enviromic prediction overcomes the bottleneck for informed decisions along the food supply chain.

the blending practices (Perten, 1964). The FN test for the harvested grain was conducted based on the established protocol according to the AACC Method 56-81.03A (1999) from 2012 to 2017, and then according to AACC Method 56-81.04 (2018) from 2018 to 2021 (Hagberg, 1960; Perten, 1964). The difference between the two methods is the change in FN correction for barometric pressure. Briefly, 25 g of wheat grain was milled into 2 oz airtight glass jars using a UDY Cyclone Mill (Hauvermale et al., 2023; Sjoberg et al., 2020). The sample weight for the FN test was adjusted to the equivalent of 7 g at 14% moisture, placed into a glass FN tube, and mixed with 25 mL of distilled water. There were two technical replicates for each sample. Samples were mixed on a shaker for 5 s and then transferred into the Hagberg-Perten FN Apparatus (Perten Instruments) to measure FN.

2.2 | Plant material and cultural data

Each year, the Washington State University Cereal Variety Testing Program grows a set of hard spring wheat varieties (the majority are hard red) to evaluate yield and other critical agronomic traits. These are elite and advanced experimental varieties from both public and private breeding programs. The same set of varieties were planted in multiple sites across the main wheat-growing region in Washington State. The testing sites were managed according to local standard farming practices. The cultural information, including planting date, harvesting date, field location (latitude and longitude coordinates) documented in the publicly accessible annual reports (https://smallgrains.wsu.edu/variety/archives/), was recompiled into a database. Not all tested varieties from all locations were used for the FN measurement.

A set of 13 spring wheat varieties were planted in the same field (Pullman, WA) in 2021 twice: April 1st and May 3rd and managed as in Liu et al. (2021). None of these varieties had been evaluated in the Cereal Variety Testing trials. Grain from each variety after maturity were harvested and measured for FN to estimate the average mean of FN from each planting.

2.3 | Modeling FN with the environmental index identified by CERIS

The number of replications for each tested variety from each environment varied from 1 to 2. FN readings of each variety from each environment were deposited at the PNW Falling Numbers website (http://steberlab.org/project7599.php). R *vca* package was used to partition the FN variance into environment, genotype, and their interaction with replicated observations in individual environments.

To identify environmental indices strongly correlated with the expected FN readings from each field, a whole-season enviromic variable matrix including potential environmental factors was compiled. One environmental factor was defined as an environmental parameter from a period of growth window, such as the average temperature from 10 to 20 days-after-planting (DAP). Four primary daily environmental parameters, day length (DL, h), daily minimal temperature (T_{min}, °F), daily maximal temperature (T_{max}, °F), and precipitation (PR, mm), were used to derive other composite environments, including GDD = $(T_{max} + T_{min})/2-32$, DTR = $T_{max} - T_{min}$, PTT = GDD × DL, PTR = GDD/DL, PRDTR = PR/DTR. Day length of each location was calculated with R *geosphere* package, while T_{min}, T_{max}, and PR were retrieved from the NASA Langley Research Center Power Project (https://power.larc.nasa.gov/) database using R *nasapower* package based on latitude and longitude coordinates (Sparks, 2018). Critical environmental regressor through informed search (CERIS, https://github.com/jmyu/CERIS_JGRA) was applied to uncover the environmental index strongly correlated with the mean FN readings (Li et al., 2021). The environmental parameter with the strongest correlation with FN was defined as the environmental index, which was used to model and predict the FN. To test the robustness of environmental mean estimates, three sets of the FN mean from each environment were estimated: the first set was calculated from all the tested varieties, the second set was calculated from the commonly tested varieties with FN from more than 44 environments, and the third was calculated from the sporadically tested varieties with FN from less than 21 environments.

2.4 | Predicting FN observed in the variety testing trials

Replacing the environmental mean, the explanatory variable used in the Finlay-Wilkinson regression model, with the identified environmental index by CERIS enables trait prediction in new environments (Eberhart & Russell, 1966; Finlay & Wilkinson, 1963; Li et al., 2021). Specifically, FN at the variety level was modeled as $y_{ij}=a_i+b_i\times \mathrm{EI}_j+e_{ij}$, where y_{ij} is the measured FN reading for the ith variety at the jth environment, a_i is the intercept for the ith variety, b_i is the slope for the variety, EI_j is the identified environmental index by CERIS from the jth environment, and e_{ij} is the residual.

To evaluate this predictive model, the k-fold cross-validation scheme was conducted. All 51 environments were randomly split into k chunks with equal number of environments ($51/k$). Sequentially, each chunk was treated as testing data, while the remaining chunks were training data to build the model and predict the performance of varieties in the testing data. Three folds, 5, 10, and 51, were tested, where 51-fold equivalates to leave-one-environment-out cross-validation. The prediction accuracy, Pearson correlation coefficient, was calculated after all k chunks were predicted, i.e., the Pearson correlation coefficient was reported based on all 51 environments. For 5- and 10-fold, 10 iterations of randomly splitting were conducted. Three different thresholds (5, 10, and 20) of minimum tested environments were applied to filter the varieties in the training dataset. For the cross-validation scheme, the same environmental index identified with all 51 environments was used to simplify the process.

The forecasting models were developed by separating environments based on the growing season into training and testing environments. The first model was trained with environments from 2012 to 2017 to predict the FN since 2018. The second model was trained using data from 2012 to 2018 to predict the FN since 2019. The corresponding environmental index was identified with respective training datasets only by CERIS. The model for each variety was trained and used to forecast FN in new growing seasons at different sites. Besides the prediction accuracy, the ratio between measured and predicted FN for each variety in the new environment was also calculated. A ratio of 1 indicates the predicted value is the same as the observed value.

2.5 | Predicting FN from production fields

The average FN for each environment was modeled as $y_j=a+b\times \mathrm{EI}_j+e_j$, where y_j is the expected FN from any wheat growing at the jth environment, a is the intercept, b is the slope, EI_j is the environmental index from the jth environment, and e_j is the residual. This model was trained with the 51 environments from the Washington State Cereal Variety Testing trial and used to predict the expected FN from the two-plantings experiment, the simulated planting date at a specific location across 10 years, and the overview of the FN trend from production fields across a large geographic region in 2021. For each tested environment, the planting date, either known, simulated, or estimated based on Normalized Difference Vegetation Index (NDVI) dynamics, and the GPS coordinates were fed into CERIS to calculate EI_j and predict the FN.

Two satellite imagery databases were combined to estimate the planting date of each spring wheat production field in 2021. The coordinates and the total acreages of spring wheat fields were obtained from the USDA CroplandCROS explorer (Boryan et al., 2011). The weekly NDVI dynamics of 2022 from the same regions were obtained from the Vegetation Condition Explorer VegScape (Yang et al., 2013), which aggregates NDVI values from the National Aeronautics Space Administration's MODIS satellite. Both CroplandCROS and NDVI geospatial data (weekly) were downloaded in the GeoTIFF format with the resolution of 30 m per pixel and 250 m per pixel, respectively. The NDVI values, which ranged from −1 to 1, of spring wheat fields defined by CroplandCROS were retrieved by overlaying the GeoTIFF images through the R package *raster*. The time-series dynamics of NDVI, which is associated with crop photosynthesis activity, has been widely used to predict crop phenology and growth (Moriondo et al., 2007; Wang et al., 2021). NDVI values of 49 weeks of each pixel field were smoothed with the *loess* function. Pixels with the highest fitted NDVI value less than

0.5 were removed. The planting date of each field was broadly defined as 7 weeks before the week with the highest fitted NDVI. The actual planting dates in 2021 from 13 Washington State University Program locations, which generally cover the wheat-growing area, were used to verify the accuracy of estimated planting dates based on NDVI dynamics.

3 | RESULTS

3.1 | Enviromics-based approach to predict high-profile traits for making informed decisions

Low-throughput phenotyping technologies for high-profile traits impede the timely delivery of critical trait values needed for informed decisions throughout the food supply chain. Given that environmental conditions are the primary driver for crop performance fluctuation in production fields, we reasoned that enviromic prediction models, which predict trait values based on weather conditions, could effectively address this bottleneck and enhance decision support (Figure 1). Enviromic prediction models can be standalone for the resolution at the environment level, or seamlessly integrated with genomic prediction models to reach the variety level resolution. With existing infrastructures, including weather profile databases and satellite imagery databases, building enviromic prediction models to forecast and monitor crop performance across large geographic areas is possible.

3.2 | Developing enviromic prediction model from long-term historical variety testing trials

We have measured FN for elite hard spring wheat varieties harvested from Washington State University Cereal Variety Performance trials. These trials were conducted across major wheat growth areas across Washington State to provide growers unbiased information for variety performance. From 2012 to 2020, a total of 3384 FN readings were recorded for 133 varieties from 51 environments (unique combinations of site and season, Table S1 and Figure S2). On average, 33 varieties (ranging from 22 to 42) were tested in each environment. The readings at the variety level ranged from 75 to 576 s (Figure 2a). For varieties with FN records in more than 10 environments, the coefficient of variation ranged from 4.7% to 25.6%, indicating some varieties were stable across environments, while others were susceptible to environmental variations. A portion of trials had FN records from replicated plots, which enabled partitioning of the total variance into the environment (40.8%), genotype (19.1%), and their interaction (14.7%) (Figure 2b). These results agreed with the consensus that FN fluctuations were mainly attributed to environment (Sjoberg et al., 2020).

To train a predictive model capturing the environmental contribution to FN, we performed Critical Environmental Regressor through Informed Search (CERIS), which was developed to identify the desired explicit environmental index that is strongly correlated with environmental mean, the average trait value from each environment (Li et al., 2021). Despite the unbalanced nature of the dataset due to yearly replacement of a portion of old varieties with newly developed ones (Figure S2A), subsampling analyses supported that the environmental mean was a robust estimate (Figure S2). Based on the planting date of each environment, we compiled a whole-season enviromic variant matrix with a total of 73,080 potential explanatory variables (Li, Guo, et al., 2022). Among these enviromic variables, CERIS identified that photothermal time (PTT) from 109 to 123 days-after-planting (DAP), or $PTT_{(109–123)}$, had the strongest correlation with the mean FN readings (Figure 2c and Figure S3). The positive correlation ($r = 0.646$, $p = 3.027 \times 10^{-7}$) indicated that hard spring wheat experienced higher PTT within this 2-week growth window tended to have higher FN readings (Figure 2d). For every unit increase of $PTT_{(109–123)}$, an increase of 0.37 s FN readings was expected.

We conducted a validation experiment to test the biological relevance of $PTT_{(109–123)}$ to FN. A new set of wheat varieties was planted twice (April 1st and May 3rd) at the same location (Pullman, WA) in 2021 (Table S2). $PTT_{(109–123)}$ from the first planting had a value of 716, and 450 from the second planting (Figure 2e and Figure S4). Based on the relationship between FN and $PTT_{(109–123)}$, grain harvested from the first planting was predicted to have higher FN readings (98 seconds on average) than the one from the second planting. The observed mean FN difference between the two plantings was 92 seconds, similar to the predicted (Figure 2f). This experiment indicated that this environmental index identified by CERIS was biologically relevant to FN.

3.3 | Enviromic prediction forecasted FN

$PTT_{(109–123)}$ was derived from a growth window prior to harvesting among 80% of trials (Figure S5). Therefore, a predictive model based on this explanatory variable could forecast the FN trend before harvest to facilitate the decision-making process. We first tested cross-validation scenarios by randomly splitting the 51 environments as training or testing. The leave-one-environment-out

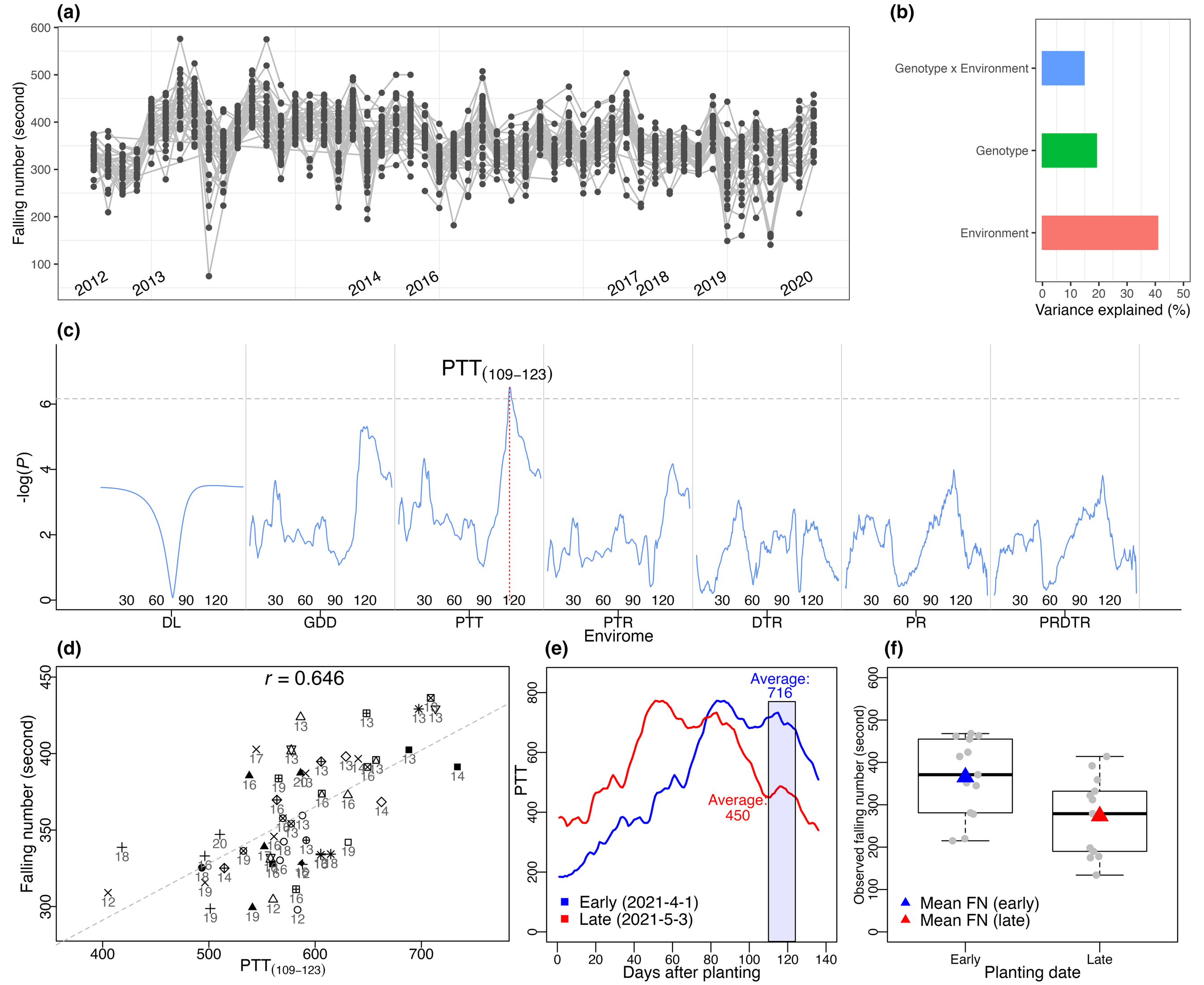

FIGURE 2 Enviromic prediction for FN. (a) Fluctuation of FN readings from testing trials spanning a large temporospatial scale. Environments are positioned on the *x*-axis based on growing season and alphabetic order of testing sites. Each line connects the FN readings of the same variety in different environments. (b) The environmental term was a major contributor to FN variation. (c) $PTT_{(109–123)}$ had the strongest correlation with the average FN. (d) Environments with higher $PTT_{(109–123)}$ generally had higher FN readings. Dot shapes denote testing sites, while the gray numbers indicate the growing seasons (such as 12 for 2012). (e) PTT profiles (2-week average) from two plantings at the same location in 2021. Blue line represents the first planting in April, while red line represents the second planting in May. The shaded area indicates the growth period of 109–123 days after planting. (f) $PTT_{(109–123)}$ explained the FN difference between two planting dates. Gray dots are the observed FN reading for tested varieties. Triangles show the average FN readings.

cross-validation had a prediction accuracy of 0.624 and the average ratio between measured and predicted FN was 1.002 (with a standard deviation of 0.122) at the variety level (Figure S6A). The similar level of prediction accuracy from the 5- and 10-fold cross-validation (with less training samples) supported the robustness of this enviromic prediction model capturing the FN trend (Figure S6B–H).

We then tested a series of forecasting scenarios based on $PTT_{(109–123)}$ to predict FN by splitting the training and testing dataset based on growing season. The first forecast model was trained with data from 2012 to 2017 to predict the FN trend for growing seasons after 2018. The second forecast model was trained with data from 2012 to 2018, which reflected the practice of updating forecast models with new available data. Because only two sites in 2020 had FN records, we didn't update the forecasting model by including 2019 data to predict 2020. The prediction accuracies ranged from 0.42 to 0.58 (Figure 3a). Meanwhile, the average ratio between measured and predicted values for each variety was 0.97 ± 0.13 (Figure 3b). These statistics suggested these predictive models captured the general trend of FN in new environments.

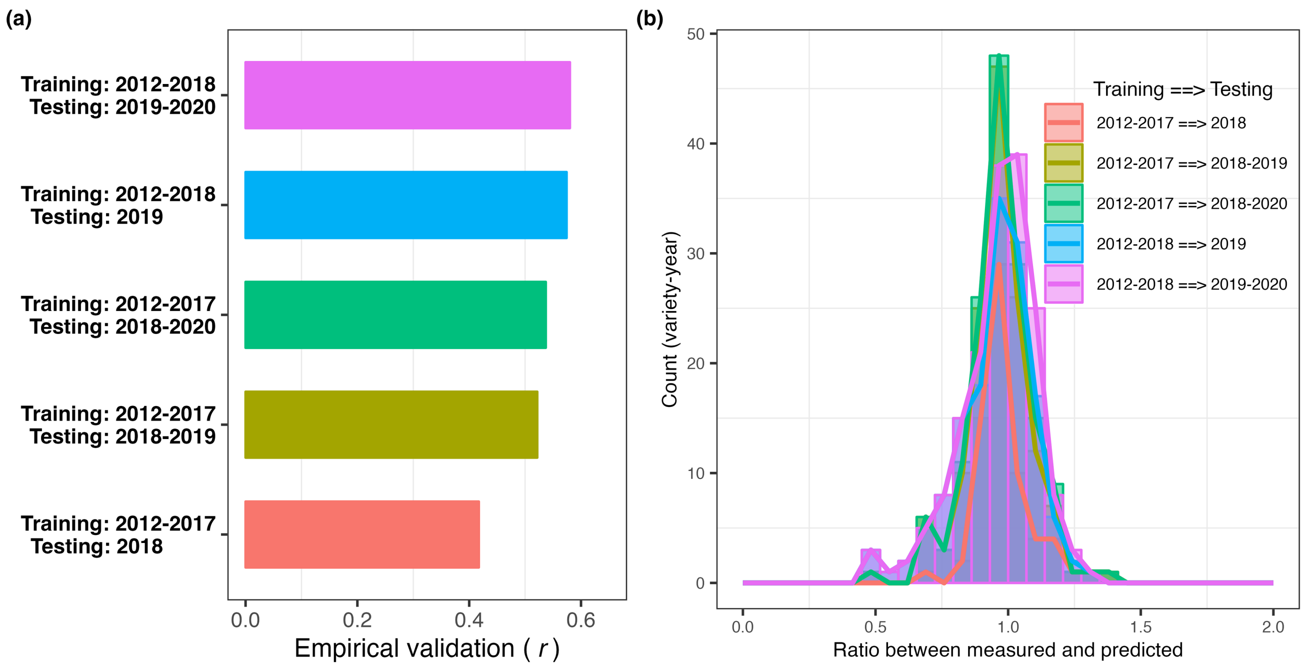

FIGURE 3 Enviromic prediction models to forecast trait trend in new season. (a) Prediction accuracy of FN in new growing seasons. (b) Distribution of the ratio between measured and predicted FN at the variety level. A ratio equal to one indicates the predicted FN is the same as the measured value.

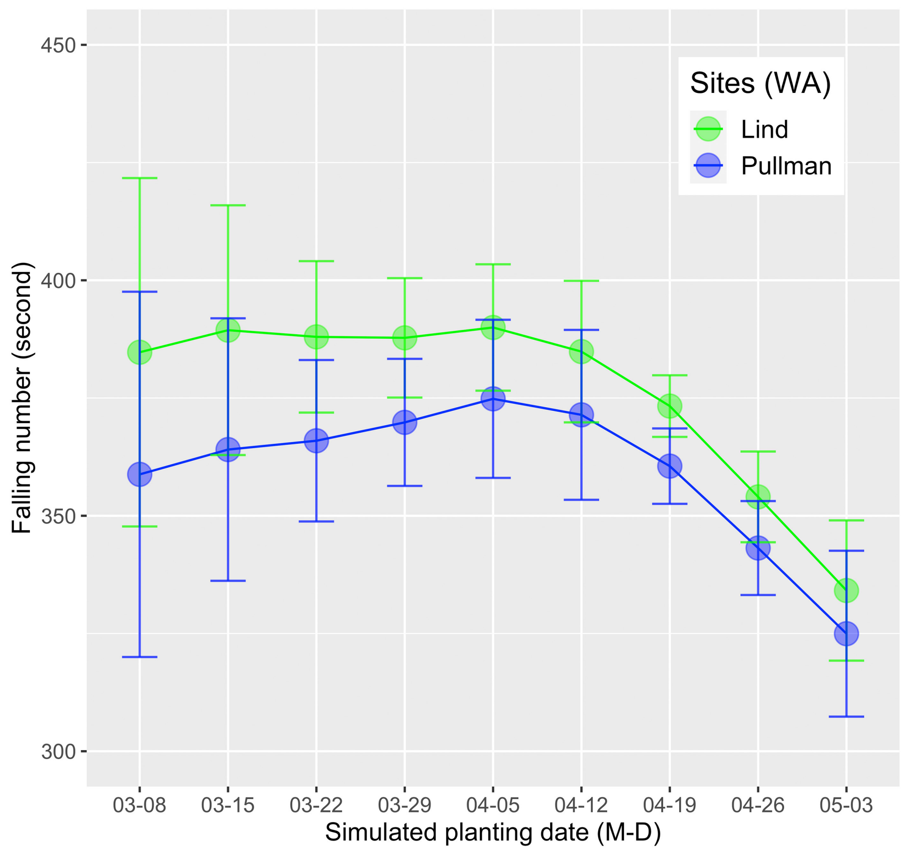

FIGURE 4 A potential application for the predictive model is using historical weather profiles to recommend the optimal planting dates at a specific location with the lowest risk of FN issues.

3.4 | Applications of enviromic prediction for production fields

We developed two applications to illustrate how such predictive models might facilitate decision-making regarding high-profile traits. The first application predicted the FN trends from specific locations for potential planting dates in spring (Guo et al., 2023). For each field, $PTT_{(109–123)}$ was calculated with the potential planting dates in each year from 2012 to 2021 and used to predict FN. Variation of predicted values based on historical weather profiles over a decade was used to recommend the optimal planting dates with the lowest risk for FN. For instance, the optimal planting date at Lind would be the week of March 29 (Figure 4). Predicted results from other locations also revealed a trend that planting earlier would more likely result in higher FN values than planting later (Figure S7).

The second application predicted FN from production fields across large geographic regions by integrating publicly accessible agricultural satellite imagery databases: USDA CroplandCROS (Figure S8) and Vegetation Condition Explorer VegScape (Figure S9). In 2021, spring wheat was grown on about 626,613 acres (2535.81 km^2) in Washington State. To predict FN from each field, we first estimated the most likely planting date. Using the documented variety testing trial in 2021 (13 sites), we found the dynamics of Normalized Difference Vegetation Index (NDVI) across the season could be used to infer the planting dates (Figures S10 and S11). We were able to infer planting dates for ~50% of spring wheat fields, which allowed us to calculate $PTT_{(109–123)}$ values to predict FN for these fields (Figure 5 and Figure S12). This prediction showed that, in 2021, FN from ~8% of spring wheat fields would be potentially lower than 300 seconds. Such an in-season, site-specific, and real-time trait trend overview might help grain elevator companies and FN laboratories to prepare logistic plans.

4 | DISCUSSION

Sustaining the complex agricultural industry depends on timely and objective decisions based on trait values for crops harvested from production fields. Because of varied trait values across fields, high-throughput phenotyping methods efficiently measuring traits across large production fields are highly desired for timely informed decisions for the agricultural industry. While such methods are

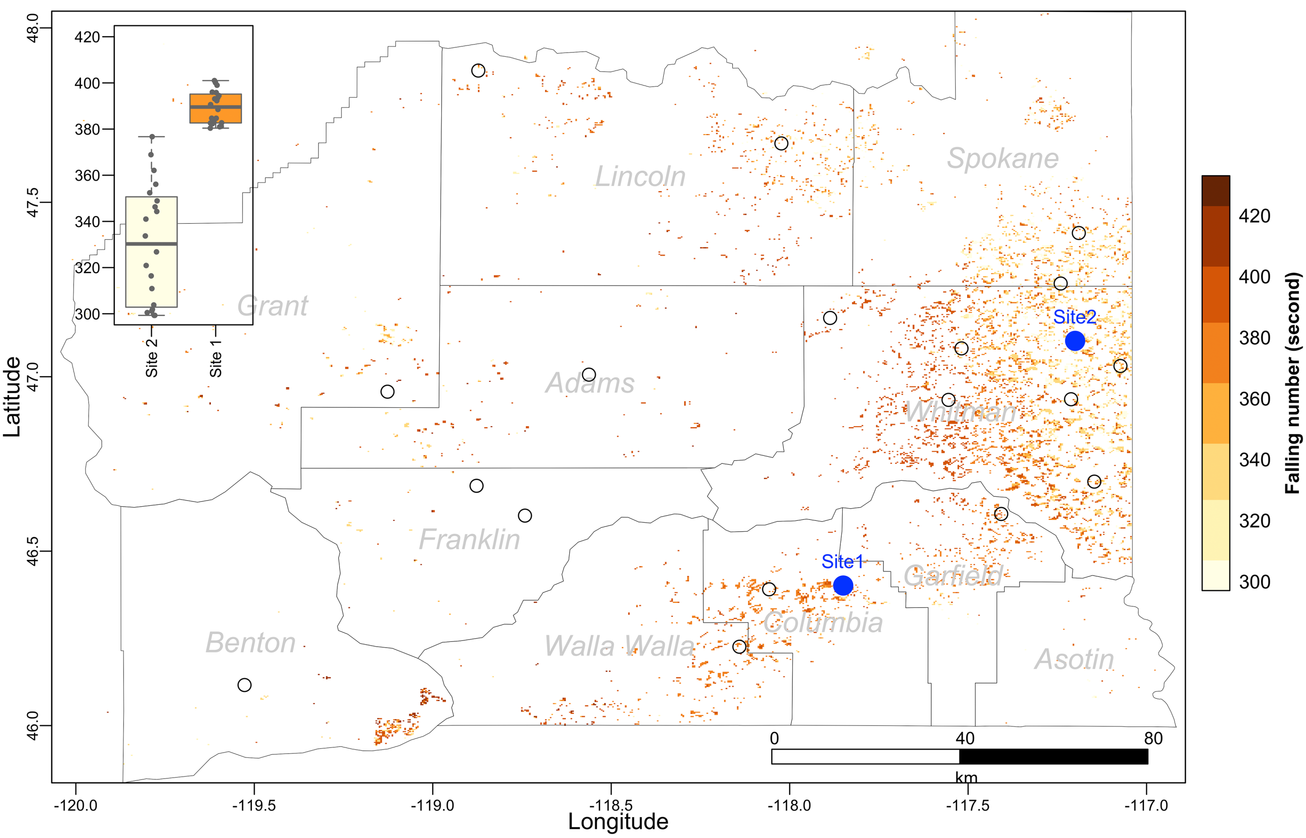

FIGURE 5 A potential application of predicted FN overview for spring wheat production fields. The fields were color-coded based on the predicted FN. The inset showed the predicted FN for potential planting dates (±10 days) around the estimated planting date for two randomly selected fields. The black circles indicate the variety of testing sites used in Figure S10.

under active development, enviromics-based approaches to predict trait trends could be an alternative because environmental conditions are a major contributor to trait fluctuations.

Models and simulations using historical or forecasted weather data have been built and conducted to predict crop yield in a large-scale geographic region (Gardner et al., 2021; Li, Wang, et al., 2022). Most of these studies trained models with a large number of independent weather variables. In contrast, the prediction model presented here was trained with a single independent weather variable, or an environmental index, $PTT_{(109-123)}$. The biological relevance of this variable to the trait of interest was further verified by a two-planting date experiment in a new growing season. Therefore, our predictive model was not only able to capture trait trends but also shed new insights on how environmental conditions contribute to traits of interest in finer detail.

Previous studies suggested that including multiple environmental indices might further increase prediction accuracy (Hardwick & Wood, 1972; Li, Guo, et al., 2022; Piepho & Blancon, 2023). In this study, among all the tested environmental variables, only $PTT_{(109-123)}$ passed the significance threshold. Therefore, we did not include other environmental variables in the prediction model. As variety performance is also influenced by genetic constituents, if genotype information is available, integration of genetic marker and environmental index would be able to predict performance of untested varieties in new environments. One approach is training genomic prediction models to predict the property of phenotypic plasticity of untested varieties, such as intercept and slope related to the identified environmental index, then integrating this property with enviromic prediction (Guo et al., 2020; Li et al., 2018, 2021; Mu et al., 2021). However, in this study, genotype data were not available, especially for proprietary varieties (almost half the tested varieties).

For large-scale production fields, growers typically only plant seeds once for each growing season. The timing of planting is critical and should be optimized (McDonald et al., 2022; Qiao et al., 2023). Decisions of individual growers are made based on multiple factors, including recommendation from seed labels, field conditions, and forecasted near-term weather conditions. With enviromic prediction models, it is possible to predict multiple traits of interest through simulating all possible planting dates at each field based on historical weather conditions. A composite index based on predicted traits could offer another layer of information to jointly decide the best timing of planting. Furthermore, the computational efficiency of these types of explainable enviromic prediction models permits the integration of other dataset, such as satellite imagery databases (Resende et al., 2024), to develop applications for large-scale production fields.

Developing a robust predictive model requires a large training dataset. This proof-of-concept case leveraged well-documented long-term historical records, which have been used to unveil the intricate relationship between crop performance and climatic conditions (Bonecke et al., 2020; Laidig et al., 2017; Li, Guo, et al., 2022). As conducting variety testing trials has been an essential component in breeding programs, the proposed enviromics concept of integrating long-term crop performance records and weather databases can be widely applied to other traits and crops. For high-profile traits lacking available historical records, acquiring a large training dataset is achievable through coordinated trials across large locations within a few years.

5 | CONCLUSION

The threats from rapidly changing climatic conditions to global food security are multi-faceted. Informed decisions based on crop performance, which fluctuates among fields, locations, and seasons, are essential for maximizing profit margins. For critical traits awaiting high-throughput phenotyping technologies, we proposed and demonstrated enviromic prediction as a potential strategy to capture the general trends of high-profile traits to assist decision-making. The availability of infrastructure, such as weather stations, satellite imagery databases, Internet of Things, and artificial intelligence analytics (Negus et al., 2024), makes it possible to generate forecasts for other high-profile traits and provide practical guidance for informed decisions for the agricultural industry.

AUTHOR CONTRIBUTIONS

X.L. and C.S. designed the research. B.Z. and X.L. conducted the research. A.H., Z.Z., A.T., C.N., M.P., and K.G.-C. contributed materials and analysis tools and interpreted results. X.L., B.Z., and C.S. drafted the manuscript with input from all authors.

ACKNOWLEDGMENTS

The authors thank Stephen Guy and Ryan Higginbotham for managing the trials and Galina Mikhaylenko, Tracy Harris, John Kelly, and Rehana Parveen for measuring FN.

FUNDING INFORMATION

This work was supported by USDA-ARS In-House Project 2090-21000-039-00D, the Agriculture and Food Research Initiative competitive grant 2021-67013-33833, and the Foundation for Food & Agriculture Research grant 22-000210.

CONFLICT OF INTEREST STATEMENT

The authors have stated explicitly that there are no conflicts of interest in connection with this article.

DATA AVAILABILITY STATEMENT

All data are available in supplementary materials.

ORCID

Xianran Li https://orcid.org/0000-0002-4252-6911

REFERENCES

Anthony, J. C., Eaton, W. W., & Henderson, A. S. (1995). Looking to the future in psychiatric epidemiology. *Epidemiologic Reviews*, *17*(1), 240–242. https://doi.org/10.1093/oxfordjournals.epirev.a036182

Bonecke, E., Breitsameter, L., Bruggemann, N., Chen, T. W., Feike, T., Kage, H., Kersebaum, K. C., Piepho, H. P., & Stutzel, H. (2020). Decoupling of impact factors reveals the response of German winter wheat yields to climatic changes. *Global Change Biology*, *26*(6), 3601–3626. https://doi.org/10.1111/gcb.15073

Boryan, C., Yang, Z., Mueller, R., & Craig, M. (2011). Monitoring US agriculture: The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto International*, *26*(5), 341–358.

Cooper, M., & Messina, C. D. (2021). Can we harness "enviromics" to accelerate crop improvement by integrating breeding and agronomy? *Frontiers in Plant Science*, *12*, 735143. https://doi.org/10.3389/fpls.2021.735143

Cooper, M., & Messina, C. D. (2023). Breeding crops for drought-affected environments and improved climate resilience. *Plant Cell*, *35*(1), 162–186. https://doi.org/10.1093/plcell/koac321

Council, N. R. (2015). *A framework for assessing effects of the food system*. The National Academies Press.

Crossa, J., Fritsche-Neto, R., Montesinos-Lopez, O. A., Costa-Neto, G., Dreisigacker, S., Montesinos-Lopez, A., & Bentley, A. R. (2021). The modern plant breeding triangle: Optimizing the use of genomics, phenomics, and enviromics data. *Frontiers in Plant Science*, *12*, 651480. https://doi.org/10.3389/fpls.2021.651480

Eberhart, S., & Russell, W. (1966). Stability parameters for comparing varieties. *Crop Science*, *6*(1), 36–40.

Finlay, K., & Wilkinson, G. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research*, *14*(6), 742–754.

Gardner, A., Maclean, I., Gaston, K., & Bütikofer, L. (2021). Forecasting future crop suitability with microclimate data. *Agricultural Systems*, *190*, 103084.

Guo, T., & Li, X. (2023). Machine learning for predicting phenotype from genotype and environment. *Current Opinion in Biotechnology*, *79*, 102853. https://doi.org/10.1016/j.copbio.2022.102853

Guo, T., Mu, Q., Wang, J., Vanous, A. E., Onogi, A., Iwata, H., Li, X., & Yu, J. (2020). Dynamic effects of interacting genes underlying rice flowering-time phenotypic plasticity and global adaptation. *Genome Research*, *30*(5), 673–683. https://doi.org/10.1101/gr.255703.119

Guo, T., Wei, J., Li, X., & Yu, J. (2023). Environmental context of phenotypic plasticity in flowering time in sorghum and rice. *Journal of Experimental Botany*, *75*, 1004–1015. https://doi.org/10.1093/jxb/erad398

Hagberg, S. (1960). A rapid method for determining alpha-amylase activity. *Cereal Chemistry*, *37*(2), 218–222.

Hardwick, R. C., & Wood, J. T. (1972). Regression methods for studying genotype-environment interactions. *Heredity (Edinb)*, *28*(2), 209–222. https://doi.org/10.1038/hdy.1972.26

Hauvermale, A. L., Parveen, R. S., Harris, T. J., Tuttle, K. M., Mikhaylenko, G., Nair, S., McCubbin, A. G., Pumphrey, M. O., & Steber, C. M. (2023). Streamlined alpha-amylase assays for wheat preharvest sprouting and late maturity alpha-amylase detection. *Agrosystems, Geosciences & Environment*, *6*(1), e20327. https://doi.org/10.1002/agg2.20327

Hu, Y., Sjoberg, S. M., Chen, C. J., Hauvermale, A. L., Morris, C. F., Delwiche, S. R., Cannon, A. E., Steber, C. M., & Zhang, Z. (2022). As the number falls, alternatives to the Hagberg-Perten falling number method: A review. *Comprehensive Reviews in Food Science and Food Safety*, *21*(3), 2105–2117. https://doi.org/10.1111/1541-4337.12959

Laidig, F., Piepho, H. P., Rentel, D., Drobek, T., Meyer, U., & Huesken, A. (2017). Breeding progress, environmental variation and correlation of winter wheat yield and quality traits in German official variety trials and on-farm during 1983-2014. *Theoretical and Applied Genetics*, *130*(1), 223–245. https://doi.org/10.1007/s00122-016-2810-3

Lempert, R. J. (2002). A new decision sciences for complex systems. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(Suppl 3), 7309–7313. https://doi.org/10.1073/pnas.082081699

Li, S., Wang, B., Feng, P., Li Liu, D., Li, L., Shi, L., & Yu, Q. (2022). Assessing climate vulnerability of historical wheat yield in south-eastern Australia's wheat belt. *Agricultural Systems*, *196*, 103340.

Li, X., Guo, T., Bai, G., Zhang, Z., See, D., Marshall, J., Garland-Campbell, K. A., & Yu, J. (2022). Genetics-inspired data-driven approaches explain and predict crop performance fluctuations attributed to changing climatic conditions. *Molecular Plant*, *15*(2), 203–206. https://doi.org/10.1016/j.molp.2022.01.001

Li, X., Guo, T., Mu, Q., Li, X., & Yu, J. (2018). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(26), 6679–6684. https://doi.org/10.1073/pnas.1718326115

Li, X., Guo, T., Wang, J., Bekele, W. A., Sukumaran, S., Vanous, A. E., McNellie, J. P., Cortes, L. T., Lopes, M. S., Lamkey, K. R., Westgate, M. E., McKay, J. K., Archontoulis, S. V., Reynolds, M. P., Tinker, N. A., Schnable, P. S., & Yu, J. (2021). An integrated framework reinstating the environmental dimension for GWAS and genomic selection in crops. *Molecular Plant*, *14*(6), 874–887. https://doi.org/10.1016/j.molp.2021.03.010

Liu, C., Parveen, R. S., Revolinski, S. R., Campbell, K. A. G., Pumphrey, M. O., & Steber, C. M. (2021). The genetics of late maturity alpha-amylase (LMA) in North American spring wheat (*Triticum aestivum* L.). *Seed Science Research*, *31*(3), 159–168. https://doi.org/10.1017/S0960258521000064

Mares, D. J., & Mrva, K. (2014). Wheat grain preharvest sprouting and late maturity alpha-amylase. *Planta*, *240*(6), 1167–1178. https://doi.org/10.1007/s00425-014-2172-5

McDonald, A. J., Balwinder, S., Keil, A., Srivastava, A., Craufurd, P., Kishore, A., Kumar, V., Paudel, G., Singh, S., Singh, A. K., Sohane, R. K., & Malik, R. K. (2022). Time management governs climate resilience and productivity in the coupled rice-wheat cropping systems of eastern India. *Nature Food*, *3*(7), 542–551. https://doi.org/10.1038/s43016-022-00549-0

Millet, E. J., Kruijer, W., Coupel-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., Charcosset, A., Welcker, C., van Eeuwijk, F., & Tardieu, F. (2019). Genomic prediction of maize yield across European environmental conditions. *Nature Genetics*, *51*(6), 952–956. https://doi.org/10.1038/s41588-019-0414-y

Moriondo, M., Maselli, F., & Bindi, M. (2007). A simple model of regional wheat yield based on NDVI data. *European Journal of Agronomy*, *26*(3), 266–274. https://doi.org/10.1016/j.eja.2006.10.007

Motes, W. C. (2010). *Modern agriculture and its benefits – Trends, implications and outlook*. Global Harvest Initiative.

Mu, Q., Guo, T., Li, X., & Yu, J. (2021). Phenotypic plasticity in plant height shaped by interaction between genetic loci and diurnal temperature range. *The New Phytologist*, *233*, 1768–1779. https://doi.org/10.1111/nph.17904

Negus, K. L., Li, X., Welch, S. M., & Yu, J. (2024). The role of artificial intelligence in crop improvement. *Advances in Agronomy*, *184*, 1–66. https://doi.org/10.1016/bs.agron.2023.11.001

Perten, H. (1964). Application of the falling number method for evaluating alpha-amylase activity. *Cereal Chemistry*, *41*(3), 127–140.

Piepho, H. P., & Blancon, J. (2023). Extending Finlay–Wilkinson regression with environmental covariates. *Plant Breeding*, *142*(5), 621–631.

Qiao, S., Harrison, S. P., Prentice, I. C., & Wang, H. (2023). Optimality-based modelling of wheat sowing dates globally. *Agricultural Systems*, *206*, 103608.

Resende, R. T., Hickey, L., Amaral, C. H., Peixoto, L. L., Marcatti, G. E., & Xu, Y. (2024). Satellite-enabled enviromics to enhance crop improvement. *Molecular Plant*. https://doi.org/10.1016/j.molp.2024.04.005

Resende, R. T., Piepho, H. P., Rosa, G. J. M., Silva-Junior, O. B., E Silva, F. F., de Resende, M. D. V., & Grattapaglia, D. (2021). Enviromics in breeding: Applications and perspectives on envirotypic-assisted selection. *Theoretical and Applied Genetics*, *134*(1), 95–112. https://doi.org/10.1007/s00122-020-03684-z

Sjoberg, S. M., Carter, A. H., Steber, C. M., & Garland-Campbell, K. A. (2020). Unraveling complex traits in wheat: Approaches for analyzing genotype × environment interactions in a multienvironment study of falling numbers. *Crop Science*, *60*(6), 3013–3026.

Sparks, A. H. (2018). Nasapower: A NASA POWER global meteorology, surface solar energy and climatology data client for R. *Journal of Open Source Software*, *3*(30), 1035.

Steber, C. M. (2017). Avoiding problems in wheat with low falling numbers. *Crops & Soils*, *50*(2), 22–25.

Wang, J., Li, X., Guo, T., Dzievit, M. J., Yu, X., Liu, P., Price, K. P., & Yu, J. (2021). Genetic dissection of seasonal vegetation index dynamics in maize through aerial based high-throughput phenotyping. *Plant Genome*, *14*(3), e20155. https://doi.org/10.1002/tpg2.20155

Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O., & Schurr, U. (2020). Phenotyping: New windows into the plant for breeders. *Annual Review of Plant Biology*, *71*, 689–712. https://doi.org/10.1146/annurev-arplant-042916-041124

Xu, Y. (2016). Envirotyping for deciphering environmental impacts on crop plants. *Theoretical and Applied Genetics*, *129*(4), 653–673. https://doi.org/10.1007/s00122-016-2691-5

Xu, Y., Zhang, X., Li, H., Zheng, H., Zhang, J., Olsen, M. S., Varshney, R. K., Prasanna, B. M., & Qian, Q. (2022). Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. *Molecular Plant*, *15*(11), 1664–1695. https://doi.org/10.1016/j.molp.2022.09.001

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., & Yan, J. (2020). Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Molecular Plant*, *13*(2), 187–214. https://doi.org/10.1016/j.molp.2020.01.008

Yang, Z., Yu, G., Di, L., Zhang, B., Han, W., & Mueller, R. (2013). Web service-based vegetation condition monitoring system-VegScape. In: 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, 2013. IEEE, pp. 3638–3641. https://doi.org/10.1109/igarss.2013.6723618

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: Zhang, B., Hauvermale, A. L., Zhang, Z., Thompson, A., Neely, C., Esser, A., Pumphrey, M., Garland-Campbell, K., Yu, J., Steber, C., & Li, X. (2024). Harnessing enviromics to predict climate-impacted high-profile traits to assist informed decisions in agriculture. *Food and Energy Security*, *13*, e544. https://doi.org/10.1002/fes3.544