

RESEARCH ARTICLE OPEN ACCESS

Hybrid Optimization of Wheat Yield Using Genetic Algorithms and Machine Learning With Phenotypic and Climatic Features

Juan Carlos Moreno Sánchez  | Adrián Trueba Espinosa  | Sergio Ruiz Castilla  | Farid García Lamont 

Centro Universitario UAEM Texcoco, Autonomous University of the State of Mexico, Texcoco, State of Mexico, Mexico

Correspondence: Adrián Trueba Espinosa (atruebae@uaemex.mx)

Received: 5 September 2025 | **Revised:** 1 June 2026 | **Accepted:** 29 June 2026

Academic Editor: Pramita Mishra

Keywords: bootstrapping validation | genetic algorithms | genotype-environment interaction | machine learning | soft computing | wheat yield prediction

ABSTRACT

Accurate wheat yield prediction is critical for global food security, yet existing forecasting models often struggle to balance high-dimensional genomic data with dynamic environmental variables. This study developed an automated framework based on genetic algorithms (GAs) to simultaneously optimize phenotypic selection, climatic feature engineering, and machine learning hyperparameters. The framework was evaluated across two contrasting cultivation environments: irrigated (Mexico) and nonirrigated (Middle East). For the irrigated dataset, the model achieved a peak performance of coefficient of determination (R^2) = 0.8363 and root mean squared error (RMSE) = 38.59, demonstrating that the proposed methodology is capable of predicting wheat yield with a R^2 exceeding 0.80 under irrigated conditions. Meanwhile, in the nonirrigated environment, the system maintained robust predictive power with R^2 = 0.6199 and RMSE = 721.67. To ensure the statistical reliability and reproducibility of these findings, a bootstrapping validation (1000 iterations) was performed on the top-performing individuals. This process yielded narrow 95% confidence intervals, confirming that while the GA-optimized features provide higher stability in irrigated systems, the framework effectively captures genotype-environment interactions even under water-limited conditions. This dual-environment validation, underpinned by robust resampling techniques, demonstrates the scalability of the proposed soft computing approach for precision breeding across diverse agroclimatic zones.

1 | Introduction

Grain yield (GY) in wheat is strongly influenced by genotype $\times$ environment ($G \times E$) interactions, where the performance of a genotype varies across contrasting environmental conditions such as temperature regimes, water availability, and management practices [1–3]. These interactions introduce high variability and nonlinear responses that complicate yield prediction for breeding and agronomic decision-making.

Traditional statistical approaches, including linear regression, mixed models, and analysis of variance, have been widely applied to analyze $G \times E$ effects and yield variability [4, 5]. While

statistically rigorous and interpretable, these methods struggle to handle the high-dimensional datasets now generated by field phenotyping, UAV-based measurements, and climatic databases [6].

To overcome these limitations, machine learning (ML) techniques such as support vector regression (SVR), random forest (RF), gradient boosting, and regularized regressions have demonstrated improved capacity to capture nonlinear relationships among phenotypic, climatic, and remote-sensing variables [7–9]. However, ML performance critically depends on selecting relevant input features and appropriately tuning hyperparameters.

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

Copyright © 2026 Juan Carlos Moreno Sánchez et al. *Applied Computational Intelligence and Soft Computing* published by John Wiley & Sons Ltd.

Among these approaches, evolutionary algorithms—particularly genetic algorithms (GAs)—have been increasingly used for feature selection or hyperparameter optimization in agricultural ML applications [10–13].

Although these approaches have demonstrated promising results in agricultural prediction tasks, several methodological challenges remain unresolved. Most studies apply GAs either for feature selection or for hyperparameter tuning as independent optimization steps rather than within a unified framework. In addition, previous research often focuses on specific data domains—such as climatic, phenotypic, or remote sensing variables—without systematically evaluating how their combination influences predictive performance under contrasting irrigation scenarios [14]. Furthermore, insufficient integration between evolutionary optimization, model configuration, and validation procedures continues to limit reproducibility and comparability across studies dealing with high-dimensional agricultural datasets [15, 16].

To address these challenges, the present study develops and evaluates a GA framework that simultaneously optimizes feature subsets and hyperparameter configurations for multiple ML models in wheat yield prediction under irrigated and non-irrigated environments.

This study makes three primary contributions.

First, it proposes an integrated GA framework that simultaneously performs feature selection and hyperparameter optimization across multiple ML models for wheat yield prediction.

Second, it introduces and evaluates two evolutionary search strategies (GA-V1 and GA-V2), demonstrating how model-specialized populations improve predictive performance and optimization stability under contrasting irrigation conditions.

Third, it provides a reproducible workflow that integrates heterogeneous phenotypic, UAV-derived, climatic, and ML information, offering practical insights for precision breeding and agricultural decision-making under genotype $\times$ environment variability.

2 | Materials and Methods

2.1 | Datasets and Experimental Environments

2.1.1 | Irrigated Dataset

The irrigated dataset consists of 370 bread wheat breeding lines evaluated under optimal agronomic management at the Norman E. Borlaug Experimental Station (CENEB). These experimental data were originally reported as part of the GRATA wheat breeding evaluations conducted under controlled irrigation conditions [17]. The same dataset was later employed in our

previous work on wheat yield prediction and variable selection using ML techniques [14].

Phenotypic traits include biomass, harvest index, thousand grain weight, NDVI measurements from ground sensors and UAV platforms, and canopy temperature. These measurements were collected under conditions where environmental stress is minimized and yield expression is primarily driven by genotypic and physiological differences.

2.1.2 | Nonirrigated Dataset

The nonirrigated dataset includes 85 wheat varieties evaluated across Afghanistan, Iran, and Turkey between 2015 and 2019. This dataset corresponds to the phenotypic data from the International Wheat Yield Partnership (IWYP) Wheat Exchange Set, publicly released for Genome-Wide Association Studies and multienvironment analysis [18].

The dataset contains phenotypic descriptors such as early vigor, senescence, green area from imaging, NDVI, and grains per spike, collected under water-limited and climatically heterogeneous environments.

2.1.3 | Climatic Data Source

For the nonirrigated environments, local weather station data were not available. Therefore, climatic variables were obtained from the official NASA POWER (Prediction of Worldwide Energy Resources) platform, a NASA project that provides satellite-assimilated meteorological data widely used in agroclimatic and crop-modeling applications [19].

Daily climatic records from sowing to harvest were downloaded for each location and transformed into statistical descriptors (mean, maximum, minimum, and range) to characterize environmental variability during the crop cycle.

2.1.4 | Experimental Rationale

The irrigated dataset represents a genotype-driven scenario where environmental variability is minimized, while the non-irrigated dataset represents an environment-driven scenario dominated by climatic stress. The irrigated dataset contains phenotypic evaluations of 370 distinct wheat genotypes assessed across two growing seasons under controlled irrigation conditions. In contrast, the nonirrigated dataset comprises phenotypic records from 85 distinct wheat genotypes evaluated across two seasons under water-limited environments.

Analyzing both datasets independently allows evaluating whether the proposed optimization framework adapts its feature selection and model configuration depending on the dominant source of variability. Table 1 quantifies both the dimensionality

TABLE 1 | Dimensionality of the feature space and training instances addressed by the genetic algorithm.

Dataset	Years	Training instances	Phenotypic features	Climatic descriptors	Total candidate features
Irrigated (Mexico)	2015–2017	740	18	936	954
Nonirrigated (Middle East)	2015–2019	168	40	648	688

Note: Training instances correspond to genotype–year observations.

of the feature space and the number of training instances addressed by the proposed GA.

The combination of high-dimensional phenotypic and climatic descriptors with a limited number of training instances creates a complex optimization scenario where traditional manual feature engineering becomes impractical, motivating the use of evolutionary optimization strategies.

2.2 | Data Integration, Preprocessing, and Evolutionary Optimization Pipeline

Phenotypic measurements, UAV-derived variables, and climatic descriptors were originally stored in independent tables with different temporal and spatial resolutions. These sources were integrated by aligning records according to experimental unit, year, and location, resulting in a unified feature table in which each row represents a unique experimental observation and each column corresponds to a candidate feature for optimization, following current practices in multisource crop yield modeling [20–22].

Missing values in continuous variables were handled using mean imputation when the proportion of missing data was low and variables exhibited continuous distributions [23, 24]. All features were normalized using min–max scaling to the range [0, 1] to prevent variables with larger numeric ranges from dominating the optimization process [25, 26].

For the nonirrigated dataset, daily climatic data obtained from the NASA POWER platform were aggregated into statistical descriptors (mean, maximum, minimum, and range) over the crop cycle and joined to the corresponding phenotypic records. For the irrigated dataset, where climatic variability is minimal, only phenotypic and spectral variables were considered.

After integration and preprocessing, the dataset was structured as a feature matrix X containing phenotypic, spectral, and climatic variables and a target vector y corresponding to GY. No manual feature engineering or variable transformation was applied; instead, the complete feature space was provided to the GA, allowing the evolutionary process to identify relevant variables and interactions without prior assumptions.

Within the proposed framework, the GA controls model construction rather than acting as a preprocessing step. Each

chromosome encodes (i) a subset of features extracted from X , (ii) a regression model, and (iii) the associated hyperparameters, consistent with evolutionary-based optimization strategies in ML [27]. For each candidate solution, the selected variables are used to train the encoded model, and the resulting prediction error defines the fitness value guiding the evolutionary search. This internal evaluation is used exclusively to compare candidate solutions and does not represent the final statistical validation.

Because the datasets represent contrasting agronomic conditions—genotype-driven variability under irrigation and environment-driven variability under nonirrigated conditions—the GA converges toward different feature subsets and model configurations without scenario-specific adjustments. After convergence, bootstrap validation is applied as the final statistical evaluation step, particularly appropriate for small datasets where traditional partitioning may yield unstable estimates [28].

The interaction between data integration, evolutionary search, embedded model training, and final statistical validation is summarized in Figure 1.

2.3 | GA Design and Configuration

The optimization framework is governed by two evolutionary designs: GA-V1 (baseline) and GA-V2 (proposed swarm architecture). Both versions share the same chromosome representation but differ in their selection strategy, crossover and mutation operators, and execution architecture. GA-V1 follows a conventional evolutionary structure, whereas GA-V2 introduces a swarm-based design to improve exploration of high-dimensional feature spaces with embedded hyperparameter optimization, following established principles of exploration–exploitation balance and selection pressure in evolutionary search [29, 30].

2.3.1 | Chromosome Representation

Each chromosome encodes two components: (i) a binary vector indicating feature inclusion or exclusion from the full matrix X and (ii) a numerical segment representing the hyperparameters of the selected regression model. This structure enables simultaneous feature selection and hyperparameter tuning within the

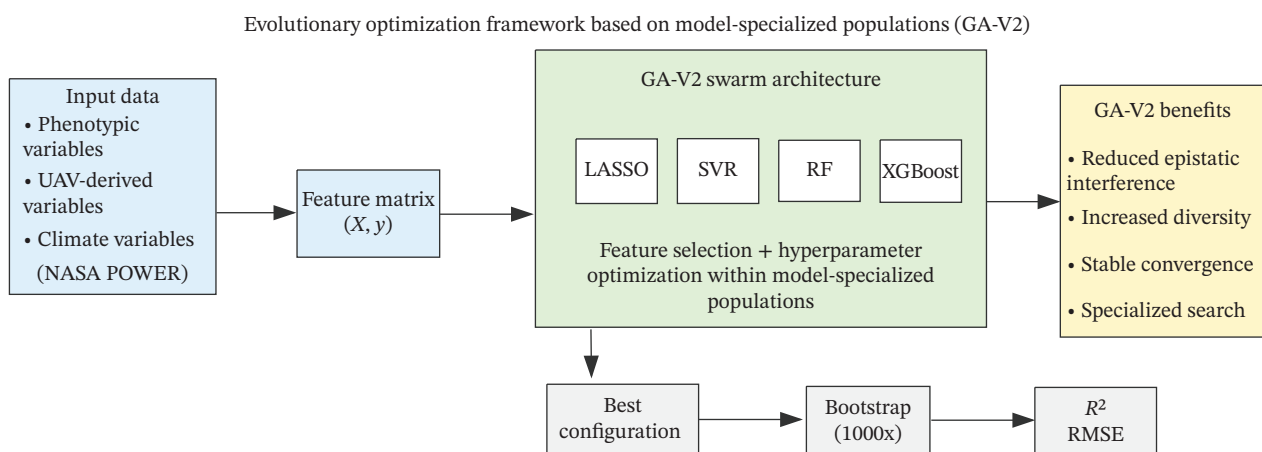


FIGURE 1 | Evolutionary optimization pipeline for feature selection, model configuration, and final validation.

same evolutionary process, consistent with evolutionary optimization strategies in ML [31].

2.3.2 | Selection Strategy

GA-V1 applies Tournament Selection based on fitness score. GA-V2 implements an alpha-beta rank-based selection in which high- and low-ranking individuals are paired to promote diversity and reduce premature convergence.

2.3.3 | Crossover and Mutation Operators

Because the chromosome combines binary and numerical segments, operators are adapted accordingly. GA-V1 employs single-point crossover with bit-flip and categorical mutation. GA-V2 applies a hybrid crossover tailored to the dual structure and a precision-oriented mutation for continuous hyperparameter refinement.

2.3.4 | Population Size and Stopping Criteria

To ensure reproducibility, parameters were fixed across experiments. GA-V1 uses a population of 40 individuals evolved for 80 generations. GA-V2 uses four parallel populations of 25 individuals each, where each population specializes in a regression model, also evolved for 80 generations. Mutation rates were set to 12% for the binary segment and 8% for the numerical segment, with a crossover probability of 85% in both versions.

The stopping criterion was defined as reaching the maximum number of generations or an early convergence condition in which the best fitness value did not improve by more than 0.001 over 12 consecutive generations.

2.3.5 | Structural Differences Between GA-V1 and GA-V2

GA-V2 constitutes a distinct evolutionary architecture rather than a parameter adjustment of GA-V1. The technical differences between both designs are summarized in Table 2.

2.4 | ML Models and Training Configuration

2.4.1 | Regression Models Embedded in the GA

Four regression models were integrated into the optimization framework: LASSO, SVR, RF, and Extreme Gradient Boosting (XGBoost). These models were selected based on their recurrent use and demonstrated robustness in wheat yield prediction studies using phenotypic, environmental, and remote sensing data [22, 32–35].

LASSO performs embedded variable selection through L1 regularization [36], SVR is effective in handling nonlinear relationships through kernel functions [37], RF captures complex interactions using ensemble decision trees [38], and XGBoost offers high predictive performance through gradient boosting with regularization mechanisms [39].

Within the proposed framework, these models are not trained independently. Instead, they are embedded into the GA as part of the chromosome structure, allowing the evolutionary process to simultaneously determine the most relevant subset of variables and the optimal hyperparameter configuration for each model under different environmental scenarios.

A baseline feature selection strategy previously reported in [14] was used as a reference point. In that approach, variables were grouped by source due to the extremely large dimensionality of the dataset. Initial experiments excluded remote sensing variables, followed by isolated inclusion of UAV-derived features, and finally, climatic variables. The results demonstrated that biomass and NDVI were the most informative predictors, highlighting the difficulty of manual feature engineering in high-dimensional agricultural datasets.

This limitation motivates the need for an automated evolutionary search capable of exploring combinations of phenotypic, environmental, and UAV-derived descriptors beyond what manual or group-based selection strategies can achieve.

2.4.2 | Hyperparameter Space Explored by the GA

The GA explores both feature subsets and hyperparameter configurations simultaneously. Table 3 describes the search space assigned to each model during the evolutionary process.

This configuration allows the GA to simultaneously explore model structure and parameter tuning under a unified evolutionary framework.

2.4.3 | Training Procedure and Fitness Evaluation

For every chromosome generated during the evolutionary process, the selected subset of features is extracted from the feature matrix, and the encoded model is trained using the associated hyperparameters.

The performance of each candidate solution is evaluated through a fitness function as follows:

- In GA-V1, fitness is computed using adjusted mean absolute percentage error ($A - MAPE$), defined as follows:

TABLE 2 | Technical comparison of genetic operators (V1 vs. V2).

Feature	Genetic algorithm V1 (baseline)	Genetic algorithm V2 (proposed swarm)
Selection strategy	Standard tournament selection	Alpha-beta selection (novel)
Crossover operator	Single-point crossover	Hybrid dual-structure crossover
Mutation process	Bit-flip + categorical model mutation	Precision-focused mutation (no model mutation)
Hyperparameter tuning	Discrete/categorical search	Continuous high-precision tuning (10^{-3})
Execution architecture	Sequential single population	Parallel swarm of model-specialized populations

TABLE 3 | Hyperparameter search space and data types for genetic optimization.

Model	Hyperparameter	Search range/space	Description
SVR	Kernel type	{Linear, Polynomial, RBF, Sigmoid}	Mapping function to higher-dimensional space
	C	[0.01, 100.0]	Margin vs. error trade-off
	Epsilon (ϵ)	[0.001, 1.0]	Width of epsilon-insensitive tube
RF	$n_estimators$	[10, 5000]	Number of trees
	max_depth	[1, 1000] or None	Tree depth control
	$min_samples_split$	[2, 10]	Minimum samples to split
XGBoost	Learning rate (η)	[0.01, 0.3]	Step size shrinkage
	Max depth	[3, 10]	Tree depth
	Subsample	[0.5, 1.0]	Fraction of samples per tree
LASSO	Alpha (α)	[0.0001, 10.0]	L1 regularization strength
	Selection method	{Cyclic, Random}	Coefficient update strategy

$$A - MAPE = 1 - \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right). \quad (1)$$

- In GA-V2, fitness is based on the coefficient of determination (R^2), promoting solutions with higher explanatory power.

This fitness evaluation is part of the internal evolutionary search and does not represent the final statistical validation of the model.

2.4.4 | Baseline Models Without GA Optimization

To evaluate the real contribution of the evolutionary optimization process, results obtained with the GA are compared against a baseline modeling strategy previously developed in [14].

In that approach, due to the very high dimensionality of the dataset, feature selection was performed manually by grouping variables into categories. Models were trained using the following:

1. Phenotypic variables without remote-sensing data,
2. Remote sensing variables independently,
3. Attempts to include climatic variables, which resulted in poor predictive performance.

The study concluded that biomass and NDVI were the most informative predictors for GY estimation [14]. Hyperparameters

were selected from predefined discrete sets, and no evolutionary optimization or LASSO-based feature selection was applied.

This baseline provides a direct reference to evaluate how the proposed GA framework improves feature selection, model configuration, and predictive performance.

2.5 | Representative Hyperparameter Configurations Selected by the GA

To illustrate the practical effect of the evolutionary optimization process, this section presents representative hyperparameter configurations identified by GA-V2 under contrasting experimental scenarios.

Unlike the baseline approach previously reported in [14], where regression models were applied using default hyperparameters and manually grouped feature subsets, GA-V2 performs simultaneous feature selection and continuous hyperparameter optimization. This allows the evolutionary process to converge toward model configurations specifically adapted to the dominant source of variability in each dataset.

Table 4 summarizes the key hyperparameters selected by the GA for two representative scenarios: the nonirrigated dataset, dominated by environmental stress, and the irrigated dataset using the full combination of phenotypic, UAV, and climatic descriptors.

The configurations reported reveal that the GA does not converge to generic or default parameter values. Instead, it adapts model complexity, tree depth, learning rate, and regularization behavior according to the agronomic context. This behavior confirms that

TABLE 4 | Representative hyperparameter configurations selected by GA-V2 across experimental scenarios.

Scenario	Model	Key hyperparameters selected by GA-V2
Nonirrigated	LASSO	$\alpha = 0.0795$, selection = cyclic
	XGBoost	$n_estimators = 1068$, $max_depth = 30$, $\eta = 0.275$, subsample = 1.0, $\gamma = 0.492$
	Random forest	$n_estimators = 601$, $max_depth = \text{None}$, $min_samples_split = 0.65$
	SVR	kernel = linear, $C = 0.202$, $\epsilon = 7.83$
Irrigated	LASSO	$\alpha = 0.0287$, selection = random
	SVR	kernel = polynomial (degree 2), $C = 0.164$, $\epsilon = 4.84$
	XGBoost	$n_estimators = 5236$, $max_depth = 220$, $\eta = 0.992$, subsample = 0.471, $\gamma = 0.418$
	Random forest	$n_estimators = 2723$, $max_depth = \text{None}$, $min_samples_split = 0.133$

the evolutionary search process plays a central role in tailoring model structure to the data characteristics, which would be difficult to obtain through manual tuning.

The ten best hyperparameter configurations identified by GA-V2 for the LASSO and XGBoost models are provided in Supporting Table S1.

2.6 | Training, Validation, and Evaluation Strategy

2.6.1 | Bootstrap Validation

During the evolutionary optimization process, candidate solutions are evaluated exclusively through an internal fitness function used to compare configurations within the GA search. This evaluation must not be interpreted as a measure of model generalization.

Statistical validation is performed only after the evolutionary process converges and the best configuration is identified. At this stage, a bootstrap resampling procedure is applied as the final evaluation step. Bootstrap is a nonparametric validation technique based on repeated random sampling with replacement from the original dataset, particularly suitable for small datasets where conventional data partitioning strategies may lead to unstable estimates [40, 41].

For each bootstrap iteration, a new training and testing subset is generated, the selected model is retrained, and its performance metrics are recorded. Repeating this process provides an empirical estimation of variability, robustness, and generalization capacity.

The distribution of performance metrics across bootstrap iterations was also analyzed to verify the stability of the GA-selected feature subsets and hyperparameter configurations. The low standard errors and narrow confidence intervals observed indicate that the evolutionary optimization converges toward statistically robust and nonrandom solutions.

2.6.2 | Performance Metrics

Model performance is evaluated using three complementary metrics:

- R^2 , measuring the proportion of variance in GY explained by the model.
- Root mean squared error (RMSE), quantifying the magnitude of prediction error.
- A-MAPE, where values closer to 1 indicate better predictive accuracy.

A-MAPE is used as the fitness metric in GA-V1, while R^2 is used in GA-V2, reflecting the methodological evolution between both versions.

3 | Results

3.1 | Evolutionary Behavior Leading to the Redesign From GA-V1 to GA-V2

GA-V1 was initially evaluated on the irrigated dataset to assess its capacity to explore the high-dimensional space generated by phenotypic, UAV-derived, and environmental variables. The best result obtained with GA-V1 corresponded to the Extreme Gradient Boost model, achieving $R^2 = 0.2796$, RMSE = 721.89, and A-MAPE = 0.7917.

Although the RMSE remained within the same magnitude as later experiments, the low R^2 indicated limited ability to explain yield variability. This behavior suggested premature convergence and insufficient exploration of the joint feature–hyperparameter space when model selection, feature selection, and hyperparameter tuning were performed within a single evolutionary population.

GA-V2 was subsequently applied to the same dataset using a swarm-based architecture in which independent populations specialize by the model type. Under this configuration, the best performance corresponded to the LASSO model, with $R^2 = 0.6199$ and A-MAPE = 0.8356, while maintaining a similar RMSE. The improvement in R^2 indicates enhanced capacity to identify explanatory relationships rather than only minimizing absolute error.

Given the limited performance of GA-V1, bootstrap validation was performed exclusively for GA-V2. Table 5 summarizes the experimental results that motivated the transition between evolutionary designs in the irrigated dataset.

3.2 | Predictive Performance Across Experimental Scenarios

After establishing the superiority of GA-V2 in the evolutionary search process, its predictive performance was evaluated across the two experimental scenarios defined in this study: non-irrigated (environment-driven variability) and irrigated (genotype-driven variability).

Table 6 summarizes the best model performance obtained by GA-V2 for each scenario using the complete feature matrix available in each dataset.

The results reveal a clear difference in predictive behavior between scenarios. In the nonirrigated dataset, where climatic stress dominates variability, the model achieved moderate predictive capacity. In contrast, the irrigated dataset exhibited substantially higher predictive accuracy, confirming that yield variability in this scenario is primarily genotype-driven and more learnable by ML models.

This contrast highlights the importance of evaluating the optimization framework under distinct sources of variability rather than reporting performance over a single homogeneous dataset.

TABLE 5 | Experimental evidence that motivated the transition from GA-V1 to GA-V2 in the nonirrigated dataset.

GA version	Best model	R^2	RMSE	A-MAPE	Outcome
GA-V1	Extreme Gradient Boost Regression	0.2796	721.89	0.7917	Discarded due to poor variance explanation
GA-V2	LASSO	0.6199	721.6764	0.8356	Adopted for further validation (bootstrap)

TABLE 6 | Best predictive performance achieved by GA-V2 across experimental scenarios.

Scenario	Best model	R^2	RMSE	A-MAPE
Nonirrigated (Phenotypic + Climatic)	LASSO	0.6199	721.6764	0.8356
Irrigated (all variables combined)	LASSO	0.8005	42.6055	0.9713

3.3 | Analysis of Environmental Variability and Feature Relevance

The variables selected by GA-V2 differ markedly between non-irrigated and irrigated environments.

In the nonirrigated dataset, the variables with selection frequency $\geq 90\%$ are dominated by NDVI-derived descriptors and biomass-related indicators, accompanied by a limited subset of climatic variables. These variables appear consistently across the best evolutionary solutions.

In contrast, in the irrigated dataset, the variables with selection frequency $\geq 90\%$ are primarily phenotypic measurements, while NDVI and climatic descriptors show substantially lower recurrence (see Tables 7 and 8).

For the irrigated dataset, the stability of this selection pattern was evaluated using bootstrap resampling. The results show low dispersion in performance metrics across resamples, indicating that the same subset of phenotypic variables is repeatedly favored by the evolutionary process.

Table 9 presents the bootstrap validation results for the irrigated dataset using GA-V2.

3.4 | Statistical Robustness and Computational Performance

The statistical robustness of the proposed framework is evidenced by the bootstrap validation results previously presented

for the irrigated dataset. The low dispersion observed in R^2 , RMSE, and A-MAPE across resamples demonstrates that GA-V2 converges toward stable feature subsets and hyperparameter configurations rather than fitting particular samples of the data.

From a computational perspective, the evolutionary process is governed by the training cost of the embedded ML models, particularly XGBoost and RFs, which consume the majority of CPU resources during fitting. Although GA-V2 introduces a swarm-based architecture, the practical level of parallelism is constrained by the internal behavior of these algorithms. Consequently, the runtime of the evolutionary process is dominated by the most computationally demanding model evaluated in each generation.

Table 10 summarizes the computational complexity and observed runtime of GA-V1 and GA-V2 in terms of population size, generations, and average execution time per evolutionary cycle.

To contextualize the predictive performance of the proposed approach, a comparison with recent wheat yield prediction studies based on remote sensing, climatic modeling, and hybrid AI methods is presented. Despite relying on classical regression models, the evolutionary optimization strategy achieves competitive performance relative to more complex deep learning or ensemble approaches reported in agricultural systems, remote sensing, and smart agricultural technology.

Table 11 presents this comparative analysis with representative studies from the literature.

TABLE 7 | Selection frequency and characterization of high-impact features (frequency $\geq 90\%$) for the nonirrigated dataset.

Variable category	Metric	Period/date	Frequency
Phenotypic			
Lodging	Value	Final	10/10
Spike length	Value	Final	9/10
Climatic			
Specific humidity at 2 m	Mean	2018, November	10/10
Temperature at 2 m	Mean	2018, November	10/10
Specific humidity at 2 m	Minimum	2018, October	9/10
Wind speed at 10 m	Maximum	2018, October	9/10
Wind speed at 10 m	Mean	2018, October	9/10
Ultraviolet-B	Range	2018, March	9/10
Ultraviolet-A	Mean	2018, December	9/10
Dew point temperature 2 m	Minimum	2018, September	9/10
Corrected precipitation	Minimum	2018, March	9/10
Wind speed at 50 m	Range	2018, May	9/10
Specific humidity at 2 m	Maximum	2015, May	9/10
Mean radiation	Range	2018, September	9/10
Mean radiation	Mean	2018, January	9/10
Wind speed at 50 m	Mean	2018, March	9/10

TABLE 8 | Selection frequency and characterization of high-impact features (frequency $\geq 70\%$) for the irrigated dataset.

Variable category	Metric	Period/date	Frequency
Phenotypic			
Harvest Index	Value	Final	10/10
Biomass	Value	Final	10/10
Thousand grain weight	Value	Final	10/10
Climatic			
High wind speed	Mean	2016, February	8/10
Biologically effective UV dose	Mean	2016, February	8/10
Wind speed	Mean	2015, December	7/10
Heat degree days	Mean	2016, February	7/10
Dew point temperature	Mean	2015, December	7/10
Minimum temperature	Mean	2016, February	7/10
Rain rate	Mean	2016, February	7/10
Indoor Heat Index	Mean	2016, February	7/10
Solar energy	Mean	2016, February	7/10

TABLE 9 | Statistical validation of the GA-optimized model performance through bootstrapping.

Metric	Mean (bootstrap)	Standard error	95% confidence interval
R^2	0.7978	0.000080	[0.7977, 0.7980]
Accuracy-MAPE	0.94488	0.000002	[0.9448, 0.9448]

TABLE 10 | Computational complexity and performance metrics.

Algorithm	Dataset	Climate source	Model pool	Total model evaluations	Average runtime
GA-V1	Nonirrigated	NASA POWER	SVR, RF, XGBoost	4000	144 h
GA-V2	Nonirrigated	NASA POWER	SVR, RF, XGBoost, LASSO	32,000	39 h
GA-V2	Irrigated	Local	SVR, RF, XGBoost, LASSO	32,000	18 h

4 | Discussion

The evolutionary process exhibited differentiated selection patterns depending on the dominant source of variability in each dataset (Tables 7 and 8). In the nonirrigated scenario, climatic descriptors derived from the NASA POWER platform appeared with high selection frequency, while phenotypic and UAV variables were less recurrent. This indicates that, under water-limited conditions, yield variability is primarily associated with environmental stress, consistent with previous reports describing strong climate–yield relationships in cereals [44, 45].

In contrast, in the irrigated dataset, phenotypic and UAV-derived variables such as biomass and vegetation indices showed the highest recurrence, whereas climatic descriptors had limited influence. Under controlled water availability, yield differences are more closely related to physiological expression and canopy structure, as described in crop physiology studies [46–48]. The consistency of these patterns across bootstrap resamples indicates that the observed selection behavior is stable rather than stochastic Table 9 [43, 49].

The performance improvement observed with GA-V2 is associated with the structure of the evolutionary search rather than

with the intrinsic complexity of the regression models (Tables 5 and 6). In the baseline configuration [14] and in GA-V1, feature subsets, model type, and hyperparameters compete within a single population, creating a heterogeneous search space in which candidate solutions frequently transition between fundamentally different model families. Such transitions may disrupt promising search trajectories and reduce the effectiveness of selection pressure. In contrast, GA-V2 organizes the optimization process into model-specialized populations, allowing evolutionary operators to refine feature subsets and hyperparameters within a consistent model structure. By eliminating direct competition among different learning algorithms, each population can explore its own search landscape more effectively, facilitating finer hyperparameter adaptation and more stable convergence. This mechanism is consistent with the substantial performance improvement observed in Table 5 and distinguishes the proposed framework from many existing ML and hybrid optimization approaches reported in the agricultural literature.

The observed behavior is also consistent with principles commonly reported in structured multipopulation evolutionary optimization. In GA-V1, feature selection, model choice, and hyperparameter tuning are jointly optimized within a single

TABLE 11 | Performance comparison between the proposed GA–ML framework and relevant previous studies.

Study	Methodology	Geographic scope	Key variables	Max R2	Observation
This study	GA + ML (XGBoost/RF)	Multiregional	NASA POWER + Phenotypic	0.8363	Full automation and global portability.
Moreno Sánchez et al. [14]	Manual Expert Selection	Local (Mexico)	Historical records	0.8700	High precision; required one month of manual curation.
Khechba et al. [12]	RF, XGBoost	Local	NDVI, moisture stress, chlorophyll	0.8100	High local precision using biophysical indices.
Raza et al. [11]	DTs, RF, GB, BTs	Local	NDVI, EVI, ARVI	0.7500	Based on tree ensembles and vegetation indices.
Haseeb et al. [10]	RF, SVM, LASSO	Local	GNDVI, NDVI, EVI	0.7200	Multiple ML models with standard spectral indices.
Akcapinar and Çakmak [42]	Satellite Indices + ML	Local (Turkey)	Drought indices	0.72–0.85	Specifically tailored for water-stressed areas.
Lokeshwari et al. [43]	ABC-CNN	Local (India)	Experimental plots	0.88–0.92	High precision in single-site controlled environments.

Note: The bold value indicates better result of this experiment.

population, creating dependencies among genetic components whose contribution to fitness varies according to the surrounding configuration. Such interactions may introduce epistatic interference, making it more difficult for evolutionary operators to preserve beneficial combinations of features and hyperparameters across generations. By organizing the search into model-specialized populations, GA-V2 reduces this source of interference and allows feature subsets and hyperparameter configurations to evolve within a more homogeneous context. This structural separation favors the preservation of adaptive genetic building blocks and contributes to the more stable convergence behavior observed during optimization.

Recent wheat yield prediction studies commonly rely on pre-defined vegetation indices, climatic variables, or UAV-derived descriptors selected based on domain knowledge prior to model training [30, 39, 49, 50]. Evolutionary–ML approaches closer to this work often apply GA as a preprocessing feature filter, while model configuration and hyperparameters are determined separately [27, 50], as summarized in Table 11. While these approaches reduce computational requirements and simplify model interpretation, they may restrict the search space and overlook interactions between feature selection and model configuration that influence predictive performance.

In the present framework, feature subset, model selection, and hyperparameter configuration are encoded within the same chromosome and evolved simultaneously (Figure 1). Unlike approaches in which feature selection and model configuration are optimized independently, the proposed framework allows these components to coevolve within a unified optimization process, enabling the discovery of interactions that may be overlooked under sequential optimization strategies.

Beyond predictive performance, ensuring methodological transparency and reproducibility is essential for the adoption of evolutionary approaches in agricultural data science. To address reproducibility concerns commonly associated with evolutionary approaches, chromosome structure, operators, parameter settings, and stopping criteria were explicitly documented following recommendations for analytic transparency in agricultural ML research [15, 16]. Bootstrap resampling was used to estimate performance variability and to assess the stability of GA-selected configurations beyond single-split validation Table 9 [41]. Although this strategy helps assess configuration stability across multiple resamples and may reduce the risk of overfitting, the relatively limited number of environments evaluated in this study may still constrain the robustness of the selected configurations.

The differentiated variable selection observed between irrigated and nonirrigated conditions has implications for data acquisition strategies. Under water-limited conditions, climatic and stress-related descriptors contribute more strongly to yield explanation, whereas under irrigation, phenotypic and UAV-derived measurements become more informative. This information can guide prioritization of measurements in breeding and field evaluation programs.

The evolutionary process remains computationally demanding, particularly for tree-based models that dominate runtime (Table 10). Consequently, the proposed framework may be less suitable for applications with strict computational constraints or

limited hardware resources. Compared with conventional ML workflows based on manual tuning or restricted search strategies, the broader optimization space explored by the proposed framework increases computational cost but provides greater flexibility for identifying robust model configurations and potentially uncovering interactions that would remain unexplored under more constrained optimization schemes. In addition, validation was conducted on two contrasting scenarios; extending evaluation to additional environments, crops, and management conditions would strengthen generalization claims. Future work may incorporate distributed computing strategies and integration with crop physiological models to enhance interpretability.

5 | Conclusions

This study presented an evolutionary optimization framework for wheat yield prediction designed to operate over high-dimensional agroenvironmental feature spaces with limited training instances. The proposed approach integrates feature selection, model choice, and hyperparameter tuning within a single GA process, avoiding predefined feature grouping or manual model configuration.

Evaluation under irrigated and nonirrigated conditions showed that the evolutionary process adapts its selection behavior according to the dominant source of variability. In the non-irrigated dataset, climatic descriptors were consistently selected, whereas in the irrigated dataset, phenotypic and UAV-derived variables showed greater relevance. These patterns indicate that the framework can identify context-dependent predictors without prior assumptions.

The comparison between GA-V1 and GA-V2 demonstrated that the swarm-based evolutionary design improves convergence and predictive performance by organizing the search process around model-specialized populations. Bootstrap validation confirmed that the selected configurations are statistically stable despite the small sample size.

The results indicate that predictive capacity can be enhanced through adaptive exploration of heterogeneous descriptors guided by evolutionary optimization, without requiring more complex predictive models. This approach provides a reproducible strategy for identifying relevant variables and model configurations across contrasting agronomic scenarios.

Future work will extend validation to additional environments and independent datasets and will explore strategies to improve computational efficiency for large-scale breeding applications.

Acknowledgments

The authors acknowledge the Universidad Autónoma del Estado de México (UAEMex) for institutional support and the Centro de Investigación de Mejoramiento del Maíz y Trigo (CIMMYT) for academic and collaborative support during the development of this study.

Funding

This work was supported in part by a doctoral scholarship from the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), Mexico.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets used in this study are publicly available and were originally reported in previous studies. The irrigated dataset consists of wheat breeding trials conducted at the Norman E. Borlaug Experimental Station (CENEB) and has been previously used in related studies [14, 17]. The nonirrigated dataset corresponds to the International Wheat Yield Partnership (IWYP) Wheat Exchange Set, publicly released for multi-environment and genomic studies [18]. All datasets are described in Section 2.1 of this manuscript.

References

1. H. E. Villaseñor-Mir, R. Hortelano-Santa Rosa, E. Martínez-Cruz, et al., "GRATA S2022: Nueva Variedad de Trigo Suave Galletero Para Áreas De Riego En México," *Revista Fitotecnia Mexicana, Sociedad Mexicana de Fitogenetica A.C.* 47 (2024): 93, <https://doi.org/10.35196/rfm.2024.1.93>.
2. E. A. Ogutu, S. L. Madahana, S. Bhavani, and G. Macharia, "Genotype × Environment Interaction: Trade-Offs Between the Agonomic Performance and Stability of Durum (*Triticum turgidum*) Wheat to Stem-Rust Resistance in Kenya," *Frontiers in Plant Science* 15 (2024): 1427483, <https://doi.org/10.3389/fpls.2024.1427483>.
3. M. Kaur and H. Ram, "Grain Yield, Heat and Water Use Efficiency of Wheat Cultivars in Relation to Different Sowing Environments," *Bangladesh Journal of Botany* 52, no. 3 (2023): 701–707, <https://doi.org/10.3329/bjb.v52i3.68878>.
4. L. Schmidt, M. Odening, J. Schlanstein, and M. Ritter, "Exploring the Weather-Yield Nexus With Artificial Neural Networks," *Agricultural Systems* 196 (2022): 103345, <https://doi.org/10.1016/j.agry.2021.103345>.
5. V. S. Manivasagam and O. Rozenstein, "Practices for Upscaling Crop Simulation Models From Field Scale to Large Regions," *Computers and Electronics in Agriculture* 175 (2020): 105554, <https://doi.org/10.1016/j.compag.2020.105554>.
6. C. Trentin, C. Bredemeier, A. L. Vian, M. A. Drum, and F. L. Santos, "Biomass Production and Wheat Grain Yield and Its Relationship With NDVI as a Function of Nitrogen Availability," *Revista Brasileira de Ciências Agrárias-Brazilian Journal of Agricultural Sciences, Revista Brasileira de Ciências Agrarias* 16, no. 4 (2021): 1–7, <https://doi.org/10.5039/agraria.v16i4a34>.
7. L. S. Cedric, W. Y. H. Adoni, R. Aworka, et al., "Crops Yield Prediction Based on Machine Learning Models: Case of West African Countries," *Smart Agricultural Technology* 2 (2022): 100049, <https://doi.org/10.1016/j.atech.2022.100049>.
8. D. Paudel, H. Boogaard, A. de Wit, et al., "Machine Learning for Large-Scale Crop Yield Forecasting," *Agricultural Systems* 187 (2021): 103016, <https://doi.org/10.1016/j.agry.2020.103016>.
9. M. Ali, R. C. Deo, Y. Xiang, et al., "Coupled Online Sequential Extreme Learning Machine Model With Ant Colony Optimization Algorithm for Wheat Yield Prediction," *Scientific Reports* 12, no. 1 (2022): 5488, <https://doi.org/10.1038/s41598-022-09482-5>.
10. M. Haseeb, Z. Tahir, S. A. Mahmood, and A. Tariq, "Winter Wheat Yield Prediction Using Linear and Nonlinear Machine Learning Algorithms Based on Climatological and Remote Sensing Data," *Information Processing in Agriculture* 12, no. 4 (2025): 431–444, <https://doi.org/10.1016/j.inpa.2025.02.004>.
11. A. Raza, M. A. Shahid, M. Zaman, et al., "Improving Wheat Yield Prediction With Multi-Source Remote Sensing Data and Machine Learning in Arid Regions," *Remote Sensing* 17, no. 5 (2025): 774, <https://doi.org/10.3390/rs17050774>.
12. K. Khechba, M. Belgiu, A. Laamrani, A. Stein, A. Amazirh, and A. Chehbouni, "The Impact of Spatiotemporal Variability of

- Environmental Conditions on Wheat Yield Forecasting Using Remote Sensing Data and Machine Learning,” *International Journal of Applied Earth Observation and Geoinformation* 136 (2025): 104367, <https://doi.org/10.1016/j.jag.2025.104367>.
13. J. Zhu, Y. Li, C. Wang, P. Liu, and Y. Lan, “Method for Monitoring Wheat Growth Status and Estimating Yield Based on UAV Multispectral Remote Sensing,” *Agronomy* 14, no. 5 (2024): 991, <https://doi.org/10.3390/agronomy14050991>.
 14. J. C. Moreno Sánchez, H. G. Acosta Mesa, A. Trueba Espinosa, S. Ruiz Castilla, and F. García Lamont, “Improving Wheat Yield Prediction Through Variable Selection Using Support Vector Regression, Random Forest, and Extreme Gradient Boosting,” *Smart Agricultural Technology* 10 (2025): 100791, <https://doi.org/10.1016/j.atech.2025.100791>.
 15. G. W. Sileshi, “Analytic Transparency is Key for Reproducibility of Agricultural Research,” *CABI Agriculture and Bioscience* 4, no. 1 (2023): 2, <https://doi.org/10.1186/s43170-023-00144-8>.
 16. D. Tamayo-Vera, X. Wang, and M. Mesbah, “A Review of Machine Learning Techniques in Agroclimatic Studies,” *Agriculture* 14, no. 3 (2024): 481, <https://doi.org/10.3390/agriculture14030481>.
 17. G. W. P. Gwp, I. Collaborators, K. Ammar, and T. Payne, in *52nd International Durum Screening Nursery* (CIMMYT Research Data & Software Repository Network, 2021), <https://hdl.handle.net/11529/10548599>.
 18. A. Morgounov, F. Ozdemir, M. Keser, et al., in *Phenotypic Data for the Genome-Wide Association Study (GWAS) of the International Treaty Project Wheat Exchange Set* (CIMMYT Research Data & Software Repository Network, 2019), <https://hdl.handle.net/11529/10548355>.
 19. A. H. Sparks, “Nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R,” *Journal of Open Source Software, The Open Journal* 3, no. 30 (2018): 1035, <https://doi.org/10.21105/joss.01035>.
 20. S. Gawdiya, D. Kumar, B. Ahmed, et al., “Field Scale Wheat Yield Prediction Using Ensemble Machine Learning Techniques,” *Smart Agricultural Technology* 9 (2024): 100543, <https://doi.org/10.1016/j.atech.2024.100543>.
 21. S. Cheng, C. Feng, L. U. Wingen, et al., “Harnessing Landrace Diversity Empowers Wheat Breeding,” *Nature* 632, no. 8026 (2024): 823–831, <https://doi.org/10.1038/s41586-024-07682-9>.
 22. E. Cheng, B. Zhang, D. Peng, et al., “Wheat Yield Estimation Using Remote Sensing Data Based on Machine Learning Approaches,” *Frontiers in Plant Science* 13 (2022): 1090970, <https://doi.org/10.3389/fpls.2022.1090970>.
 23. M. J. Castelo Cabay, V. C. Armendariz Rodríguez, E. F. Merino Villa, et al., “Exploring the Efficacy of Artificial Neural Networks in Climate Prediction: Societal and Environmental Implications,” in *Tungurahua, Ecuador* (Puebla, Mexico: Institute of Electrical and Electronics Engineers, 2024).
 24. M. Phélippe-Guinvarc’h, in *French Wine: Combination of Multiple Open Data Sources to Mapping the Expected Harvest Value* (2024), <https://arxiv.org/abs/2406.19732>.
 25. V. Onegina, L. Kucher, A. Kucher, V. Krupin, M. Kłodziński, and V. Logos, “Unlocking Innovation Capacity: Strategies for Micro-Small, and Medium Enterprises in Ukrainian Agriculture,” *Agriculture* 15, no. 1 (2025): 65, <https://doi.org/10.3390/agriculture15010065>.
 26. S. Shetty and M. T. R., “MResGat: Multi-Head Residual Dilated Convolution Assisted Gated Unit Framework for Crop Yield Prediction,” *Scalable Computing: Practice and Experience* 25, no. 6 (2024): <https://doi.org/10.12694/scpe.v25i6.3317>.
 27. J. Ding, J. Du, H. Wang, and S. Xiao, “A Novel Two-Stage Feature Selection Method Based on Random Forest and Improved Genetic Algorithm for Enhancing Classification in Machine Learning,” *Scientific Reports* 15, no. 1 (2025): 16828, <https://doi.org/10.1038/s41598-025-01761-1>.
 28. S. Raschka, in *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning* (2020), <https://arxiv.org/abs/1811.12808>.
 29. S. N. Sivanandam and S. N. Deepa, in *Genetic Algorithms* (Springer, 2008), https://doi.org/10.1007/978-3-540-73190-0_2.
 30. M. Crepinšek, S.-H. Liu, and M. Mernik, “Exploration and Exploitation in Evolutionary Algorithms: A Survey,” *ACM Computing Surveys, Association for Computing Machinery* 45, no. 3 (2013): 1–33, <https://doi.org/10.1145/2480741.2480752>.
 31. Z. Y. Taha, A. A. Abdullah, and T. A. Rashid, “Optimizing Feature Selection With Genetic Algorithms: A Review of Methods and Applications,” *Knowledge and Information Systems* 67, no. 11 (2025): 9739–9778, <https://doi.org/10.1007/s10115-025-02515-1>.
 32. M. S. Chiu and J. Wang, “Evaluation of Machine Learning Regression Techniques for Estimating Winter Wheat Biomass Using Biophysical, Biochemical, and UAV Multispectral Data,” *Drones* 8, no. 7 (2024): 287, <https://doi.org/10.3390/drones8070287>.
 33. S. Yang, L. Li, S. Fei, et al., “Wheat Yield Prediction Using Machine Learning Method Based on UAV Remote Sensing Data,” *Drones* 8, no. 7 (2024): 284, <https://doi.org/10.3390/drones8070284>.
 34. G. Lischoid, H. Webber, M. Sommer, C. Nendel, and F. Ewert, “Machine Learning in Crop Yield Modelling: A Powerful Tool, but No Surrogate for Science,” *Agricultural and Forest Meteorology* 312 (2022): 108698, <https://doi.org/10.1016/j.agrformet.2021.108698>.
 35. A. Vannoppen, A. Gobin, L. Kotova, et al., “Wheat Yield Estimation From NDVI and Regional Climate Models in Latvia,” *Remote Sensing* 12, no. 14 (2020): 2206, <https://doi.org/10.3390/rs12142206>.
 36. Y. Chhabra, S. Rawat, and S. Gautam, “Modelling and Performance Evaluation of Machine Learning Techniques in Forecasting Cereal Yield for Shri Ganganagar Region, Rajasthan, India,” *Archives of Current Research International* 24, no. 10 (2024): 110–121, <https://doi.org/10.9734/acri/2024/v24i10913>.
 37. A. S. Kharal, S. A. Mahar, M. I. Mushtaque, A. Magsi, and J. A. Mahar, “A Model for Wheat Yield Prediction to Reduce the Effect of Climate Change Using Support Vector Regression,” *VFAST Transactions on Software Engineering* 12, no. 2 (2024): 192–212, <https://doi.org/10.21015/vtse.v12i2.1855>.
 38. L. Breiman, “Random Forests,” *Machine Learning* 45, no. 1 (2001): 5–32, <https://doi.org/10.1023/A:1010933404324>.
 39. T. Chen and C. Guestrin, in *XGBoost: A Scalable Tree Boosting System* (Association for Computing Machinery, 2016), <https://doi.org/10.1145/2939672.2939785>.
 40. R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society: Series B* 58, no. 1 (2018): 267–288, <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
 41. R. T. Nakatsu, “Validation of Machine Learning Ridge Regression Models Using Monte Carlo, Bootstrap, and Variations in Cross-Validation,” *Journal of Intelligent Systems* 32, no. 1 (2023): 20220224, <https://doi.org/10.1515/jisys-2022-0224>.
 42. M. C. Akcapınar and B. Çakmak, “Yield Prediction Models for Some Wheat Varieties With Satellite-Based Drought Indices and Machine Learning Algorithms,” *Irrigation and Drainage* 74, no. 1 (2025): 237–250, <https://doi.org/10.1002/ird.2989>.
 43. M. Lokeshwari, G. K. Jha, A. Lama, et al., “Wheat Yield Prediction Through Artificial Bee Colony-Enhanced Convolutional Neural Network,” *Indian Journal of Genetics and Plant Breeding* 85 (2025): 78–86, <https://doi.org/10.31742/ISGPB.85.1.9>.
 44. F. Yu, H. Noor, M. F. Seleiman, and F. Noor, “Effects of Different Winter Wheat (*Triticum aestivum* L.) Varieties Addressing the

Agriculture Climate Interactions in Temperature Regions of Yield,” *Atmosphere* 16, no. 2 (2025): 189, <https://doi.org/10.3390/atmos16020189>.

45. A. M. S. Kheir, A. Govind, V. Nangia, et al., “Developing Automated Machine Learning Approach for Fast and Robust Crop Yield Prediction Using a Fusion of Remote Sensing, Soil, and Weather Dataset,” *Environmental Research Communications* 6, no. 4 (2024): 041005, <https://doi.org/10.1088/2515-7620/ad2d02>.

46. H. S. Saini and M. E. Westgate, “Reproductive Development in Grain Crops During Drought,” *Advances in Agronomy* 68 (1999): <https://www.sciencedirect.com/science/article/pii/S0065211308608433>.

47. P. Berry, M. Sterling, J. Spink, et al., “Understanding and Reducing Lodging in Cereals,” *Advances in Agronomy* 84 (2004): <https://www.sciencedirect.com/science/article/pii/S0065211304840057>.

48. J. F. Ledent, “The Physiology of Crop Yield,” *Annals of Botany* 2nd ed., 100, no. 4 (2007): 889–890, <https://doi.org/10.1093/aob/mcm187>.

49. M. Zheng, H. Hu, Y. Niu, Q. Shen, F. Jia, and X. Geng, “Modeling of Winter Wheat Yield Prediction Based on Solar-Induced Chlorophyll Fluorescence by Machine Learning Methods,” *European Journal of Remote Sensing* 58, no. 1 (2025): 2455940, <https://doi.org/10.1080/22797254.2025.2455940>.

50. T. Mahmud, N. Datta, R. Chakma, et al., “An Approach for Crop Prediction in Agriculture: Integrating Genetic Algorithms and Machine Learning,” *IEEE Access* 12 (2024): 173583–173598, <https://doi.org/10.1109/ACCESS.2024.3478739>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Supporting Information.** Supporting Table S1 presents the ten best hyperparameter configurations identified by GA-V2 for the LASSO and XGBoost models. The table is provided as a separate supporting file accompanying this manuscript.

Appendix A

The source code used to implement the evolutionary optimization framework, machine learning models, and preprocessing pipeline is publicly available in the following repositories:

- Genetic algorithm and machine learning optimization framework: https://github.com/carlosmorenophd/machine_phen
- Dataset integration and feature transformation pipeline: https://github.com/carlosmorenophd/phen_transform_dataset

These repositories include all scripts required to reproduce the evolutionary search process, model training, feature engineering, and preprocessing steps described in this study.