Review

Not peer-reviewed version

Genotype × Environment Interaction Analysis in Plant Breeding: Integrating Contemporary Machine Learning Approaches

Yared Semahegn *

Posted Date: 19 March 2026

doi: 10.20944/preprints202603.1584.v1

Keywords: genotype × environment interaction; plant breeding; multi-environment trials; machine learning; genomic prediction

doi:10.20944/preprints202603.1584.v1

Review

Genotype × Environment Interaction Analysis in Plant Breeding: Integrating Contemporary Machine Learning Approaches

Yared Semahegn

Holetta Agricultural Research Center, Holetta, Ethiopia; yaredsemahegnb@gmail.com

Abstract

Genotype × environment (G × E) interaction is a central challenge in plant breeding, as differential genotype performance across environments complicates selection decisions and limits genetic gain. Traditional statistical models such as analysis of variance (ANOVA), stability parameters, and mixed models have long been used to dissect G × E interactions. However, the increasing availability of large-scale phenotypic, environmental, and genomic datasets necessitates more flexible and powerful analytical approaches. Recent advances in machine learning (ML) provide novel opportunities to model complex, nonlinear G × E patterns and improve prediction accuracy for genotype performance and stability. This review summarizes classical and modern G × E analysis methods, highlights emerging machine learning techniques, and discusses their applications, and limitations in plant breeding programs.

Keywords: genotype × environment interaction; plant breeding; multi-environment trials; machine learning; genomic prediction

---

1. Introduction

Plant breeding aims to develop cultivars with high yield, stability, and adaptability across diverse agro-ecological conditions (Allard and Bradshaw, 1964; Falconer and Mackay, 1996). However, the expression of most economically important traits is controlled by quantitative genetic mechanisms and is strongly influenced by environmental conditions, leading to genotype × environment interactions (GEI) (Cooper and Hammer, 1996; Crossa, 1990). GEI arises when genotypes respond differently to environmental variation, resulting in changes in their relative performance across locations, seasons, or management practices (Kang, 2002; Yan and Kang, 2003). Such differential responses complicate the identification of consistently superior genotypes (Gauch, 1992; Yan et al., 2007). Consequently, GEI can reduce the efficiency of selection and slow genetic progress if it is not properly accounted for in breeding programs. In the context of increasing climate variability and rapidly changing production environments, understanding and effectively exploiting GEI has become even more critical for the successful development of improved cultivars. Multi-environment trials (METs), in which genotypes are evaluated across a range of environments, are therefore widely used to quantify GEI and to identify genotypes that combine high yield potential with broad or specific adaptation (Crossa et al., 2002; Yan and Tinker, 2006).

Despite their long-standing use and important contribution to the interpretation of GEI, classical approaches for analyzing genotype × environment interaction have notable limitations in the era of large and complex datasets generated by modern breeding programs. Traditional statistical methods, such as combined analysis of variance and conventional stability statistics, were primarily developed for relatively small and balanced datasets involving a limited number of genotypes and environments (Gauch, 1992; Piepho, 1998). However, contemporary METs frequently generate large volumes of data across multiple locations, years, and management conditions (Smith et al., 2005; van Eeuwijk et al., 2016), often accompanied by high-dimensional environmental and genomic information. Under these circumstances, classical GEI methods may have limited capacity to capture complex interaction patterns, nonlinear relationships, and heterogeneous variance structures inherent in such datasets (Crossa et al., 2002; van Eeuwijk et al., 2016). Moreover, many of these approaches rely on strict assumptions, including balanced designs, independence of errors, and homogeneity of variance, that are rarely satisfied in large-scale breeding experiments (Piepho et al., 2008). As a result, their application may lead to reduced analytical efficiency, biased parameter estimates, or loss of valuable information, underscoring the need for more flexible and robust analytical frameworks capable of handling high-dimensional and unbalanced data.

Advances in data science have introduced new analytical opportunities for addressing these challenges (Crossa et al., 2017; Montesinos-López et al., 2018). In particular, the emergence of machine learning (ML) techniques has opened promising avenues for improving predictive accuracy and efficiency in modern plant breeding. ML approaches are increasingly used to analyze large and complex datasets generated from high-throughput phenotyping, genotyping, and environmental monitoring systems (Heslot et al., 2012; Ma et al., 2018). Unlike traditional statistical models, ML algorithms are capable of capturing complex and nonlinear relationships among genotypes, environments, and management factors (Montesinos-López et al., 2018; Washburn et al., 2021). This capability makes ML particularly valuable for predictive breeding, where the objective is to forecast the performance of breeding lines in untested environments or future growing seasons. By integrating genomic, phenotypic, and environmental information, ML models can enhance genomic prediction, improve the characterization of genotype × environment interactions, and support more informed selection decisions (Crossa et al., 2017). Consequently, the integration of machine learning approaches into the analysis of multi-environment trials offers a promising pathway to overcome the limitations of conventional GEI analysis and to accelerate genetic gain in plant breeding programs. This review aims to: i) critically examine conventional statistical approaches used for GEI analysis in plant breeding and highlight their strengths and limitations; ii) synthesize recent advances in machine learning and artificial intelligence applied to GEI analysis and multi-environment trials; iii) identify key challenges, knowledge gaps, and future research directions for the adoption of machine learning in GEI-driven plant breeding; and, iv) provide practical insights for breeders and researchers on selecting and implementing machine learning approaches in breeding programs.

2. Biological and Statistical Foundations of G × E Interaction

GEI is a fundamental concept in quantitative genetics and plant breeding, referring to the differential response of genotypes to varying environmental conditions (Allard & Bradshaw, 1964; Falconer & Mackay, 1996; Cooper & Hammer, 1996). In the field of Plant Breeding, crop performance such as grain yield, oil content, or disease resistance is typically governed by complex quantitative traits whose expression is influenced by both genetic makeup and environmental factors (Falconer & Mackay, 1996; Bernardo, 2010). From a biological perspective, GEI arises because genotypes possess different physiological, morphological, and biochemical mechanisms that determine how they respond to environmental variables such as temperature, rainfall, soil fertility, and management practices (Cooper & Hammer, 1996; Ceccarelli, 1989). Consequently, a genotype that performs best in one environment may not necessarily maintain superior performance in another environment. This differential response reflects the plasticity of genotypes and their capacity to exploit specific environmental conditions.

From a statistical perspective, GEI is described as the deviation from the expected additive effects of genotype and environment in determining phenotypic performance (Falconer & Mackay, 1996; Crossa, 1990). In the context of GEI, this interaction is typically detected when the difference in performance between two genotypes changes across environments (Gauch, 1992; Yan & Kang, 2003). Statistically, GEI can be partitioned into different components that help explain the nature of genotype responses across environments. One commonly recognized distinction is between non-crossover (ordinal) interaction and crossover (disordinal) interaction (Baker, 1988; Kang, 2002). Non-crossover interaction occurs when genotypes maintain the same ranking across environments, but the magnitude of differences between them varies. In contrast, crossover interaction occurs when genotype rankings change across environments, meaning that a genotype superior in one environment may become inferior in another environment (Baker, 1988; Gauch, 1992). The diagrammatic view of crossover and non-crossover G×E interaction is presented in Figure 1. Crossover interactions are particularly important in plant breeding because they complicate selection decisions and may require environment-specific cultivar recommendations.

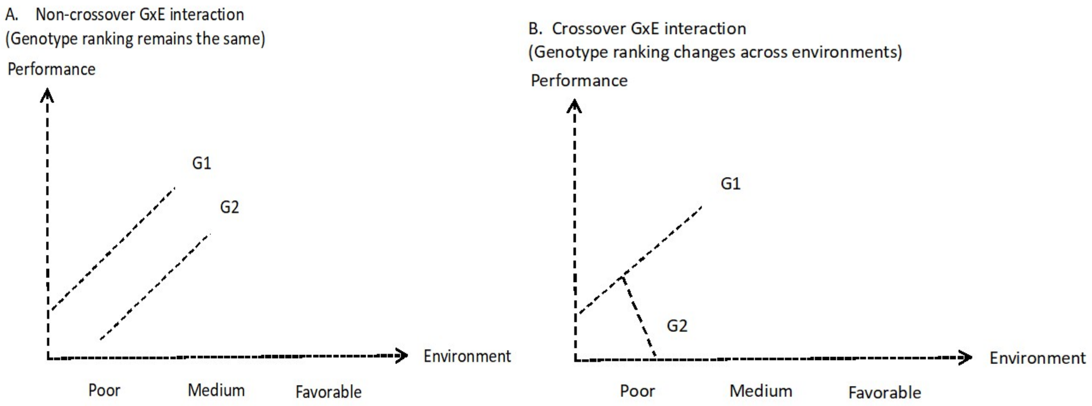

Figure 1. Diagram of crossover and non-crossover G×E interaction Another important aspect of GEI is the distinction between predictable and unpredictable environmental variation (Cooper & Hammer, 1996; Annicchiarico, 2002). Predictable environmental factors include soil type, altitude, or management practices that remain relatively stable across seasons, whereas unpredictable factors include seasonal rainfall variation, temperature fluctuations, and pest or disease outbreaks (Cooper & Hammer, 1996). Genotypes that respond differently to these environmental conditions exhibit varying degrees of phenotypic plasticity, which contributes to the magnitude of GEI observed in multi-environment trials.

Understanding the biological and statistical nature of GEI has important implications for adaptation, stability, and cultivar recommendation. Adaptation refers to the ability of a genotype to perform well under specific environmental conditions (Allard & Bradshaw, 1964; Ceccarelli, 1989). Plant breeders often distinguish between broad adaptation, where genotypes perform consistently across a wide range of environments, and specific adaptation, where genotypes perform exceptionally well in particular environments but not necessarily in others (Annicchiarico, 2002; Yan & Kang, 2003). The presence of strong crossover G×E interactions often favors breeding for specific adaptation, whereas weak interactions may allow the identification of broadly adapted cultivars.

Closely related to adaptation is the concept of phenotypic stability, which refers to the consistency of genotype performance across environments (Eberhart & Russell, 1966; Becker & Léon, 1988). Stability analysis helps breeders identify genotypes that maintain relatively stable yields despite environmental fluctuations. Several statistical approaches have been developed to quantify stability, including regression-based methods, variance-based statistics, and multivariate techniques (Lin et al., 1986; Gauch, 1992; Yan & Tinker, 2006). These approaches aim to separate the main genotype effect from the interaction component, enabling breeders to identify genotypes that combine high mean performance with desirable stability. The implications of GEI extend directly to cultivar evaluation and recommendation. In breeding programs, multi-environment trials are conducted to evaluate candidate cultivars across diverse agro-ecological zones (Crossa, 1990; Smith et al., 2005). Significant G×E interaction requires careful analysis to determine whether a genotype should be recommended broadly across regions or targeted to specific environments (Yan & Kang, 2003; van Eeuwijk et al., 2016). Accurate characterization of G×E interaction therefore supports more reliable cultivar deployment, enhances genetic gain, and improves the efficiency of breeding programs.

doi:10.20944/preprints202603.1584.v1

3. Classical Approaches to G × E Interaction Analysis

3.1. Analysis of Variance (ANOVA)

Analysis of Variance (ANOVA) is one of the earliest and most widely used classical statistical approaches for assessing GEI in METs (Yates & Cochran, 1938; Crossa, 1990). In plant breeding experiments, ANOVA partitions the total phenotypic variation observed in a dataset into components attributable to genotype, environment, and the GEI (Falconer & Mackay, 1996; Bernardo, 2010), along with experimental error. This decomposition enables breeders to determine the relative contributions of genetic and environmental factors to trait expression and to test whether genotypes perform consistently across different environments. In the standard two-way ANOVA model for METs, genotype and environment effects are considered along with their interaction term. The significance of the GEI component indicates whether genotypes respond differently to environmental changes (Gauch, 1992; Yan & Kang, 2003). A significant interaction suggests that the ranking of genotypes varies across environments, complicating the identification of broadly adapted cultivars. Conversely, a non-significant interaction implies relatively stable genotype performance across environments, allowing breeders to rely more confidently on overall mean performance for selection. Although ANOVA provides a useful initial assessment of GEI, it has several limitations. The method mainly detects the presence and magnitude of interaction but provides limited insight into the specific patterns or structure of the interaction (Gauch, 1992; Crossa et al., 2002). It does not identify which genotypes are specifically adapted to particular environments or reveal crossover interactions where genotype rankings change across environments. Furthermore, classical ANOVA assumes balanced data and homogeneous error variance, assumptions that are often violated in large multi-environment datasets (Piepho et al., 2008). As a result, ANOVA is frequently complemented by more advanced analytical methods such as stability statistics, regression approaches, and multivariate techniques to better understand genotype performance and environmental responses.

3.2. Stability and Adaptability Models

Various regression-based and variance-based statistics have been developed to evaluate stability and adaptability of genotypes across environments (Lin et al., 1986; Becker & Léon, 1988; Kang, 2002). These models aim to describe how genotypes respond to environmental variation and to identify cultivars that combine high productivity with consistent performance. Among the most widely used approaches are regression-based stability models and variance-based stability statistics. One of the earliest regression approaches is the Finlay–Wilkinson regression model (Finlay & Wilkinson, 1963), which evaluates genotype performance by regressing the mean yield of each genotype on an environmental index, typically calculated as the mean performance of all genotypes in each environment. In this model, the regression coefficient (bi) represents the responsiveness of a genotype to environmental changes. Genotypes with bi ≈ 1 are considered to have average responsiveness and broad adaptation, those with bi > 1 are more responsive and better suited to favorable environments, and those with bi < 1 are less sensitive to environmental improvement and may perform relatively better under unfavorable conditions. Although simple and intuitive, the Finlay–Wilkinson model primarily measures responsiveness rather than stability.

The Eberhart–Russell model (Eberhart & Russell, 1966) extends the Finlay–Wilkinson approach by incorporating both the regression coefficient (bi) and the deviation from regression (S^2di) as stability parameters. In this framework, an ideal genotype is characterized by a high mean yield, a regression coefficient close to unity (bi ≈ 1), and a small deviation from regression (S^2di ≈ 0). The regression coefficient measures the response of a genotype to environmental improvement, while the deviation from regression captures the unpredictability of genotype performance across environments. This model therefore provides a more comprehensive assessment of both adaptability and stability.

In addition to regression-based models, several variance-based stability statistics have been proposed. Wricke's ecovalence (Wricke, 1962) quantifies the contribution of each genotype to the total

G×E interaction sum of squares. A genotype with a small ecovalence value contributes little to the interaction and is therefore considered more stable across environments. Conversely, large values indicate that the genotype shows strong interaction with environments and thus unstable performance. Similarly, Shukla's stability variance (Shukla, 1972) provides an unbiased estimate of the variance attributed to each genotype's interaction with environments after accounting for experimental error. Genotypes with lower stability variance are considered more stable because their performance varies less across environmental conditions. Compared with Wricke's ecovalence, Shukla's method provides a statistical test for the stability parameter, allowing breeders to determine whether differences in stability among genotypes are significant. These classical stability and adaptability models have been widely applied in plant breeding programs due to their conceptual simplicity and ease of interpretation. However, most of them rely on genotype means and therefore may overlook complex patterns of interaction present in multi-environment datasets. Consequently, they are often complemented by multivariate approaches such as Additive Main Effects and Multiplicative Interaction (AMMI) and Genotype plus Genotype × Environment interaction (GGE) biplot analyses, as well as modern mixed-model and machine learning methods, which can better capture the structure of genotype × environment interactions in large and unbalanced datasets.

3.3. Multivariate Methods

To better understand the structure and pattern of GEI, multivariate statistical approaches have been widely adopted in the analysis of MET data. Among these, the AMMI model and the GGE biplot method are the most commonly used techniques. These approaches combine analysis of variance with multivariate analysis to provide a more detailed interpretation of genotype performance across environments. The AMMI model integrates ANOVA for the additive main effects of genotype and environment with principal component analysis (PCA) applied to the multiplicative GEI component (Gauch, 1992; Zobel et al., 1988). In the AMMI framework, the total variation is first partitioned into genotype effects, environment effects, and the GEI using ANOVA. The interaction portion is then decomposed into interaction principal component axes (IPCAs) through PCA (Crossa, 1990; Gauch, 1992). This approach allows researchers to identify patterns of interaction and to determine how individual genotypes and environments contribute to the overall interaction structure. AMMI is particularly useful for detecting specific adaptation, improving the accuracy of genotype performance estimates, and identifying stable genotypes through stability measures such as the AMMI Stability Value (ASV). Additionally, AMMI biplots provide a graphical representation of genotype and environment relationships, facilitating interpretation of complex interaction patterns.

The GGE biplot method focuses on the combined effects of genotype (G) and GEI, which together represent the variation relevant for genotype evaluation and selection. Unlike AMMI, which separates genotype and environment main effects, GGE biplot removes the environment main effect and analyzes only G + GEI. The method uses singular value decomposition (SVD) to generate principal components that are displayed in a biplot (Yan et al., 2000; Yan & Kang, 2003). This graphical tool allows breeders to visualize several important aspects of MET data, including the "which-won-where" pattern that identifies the best-performing genotype in specific environments, the discriminating ability and representativeness of test environments, and the stability and mean performance of genotypes (Yan & Tinker, 2006). The GGE biplot is therefore particularly useful for identifying mega-environments and recommending cultivars with either broad or specific adaptation. Both AMMI and GGE biplot approaches provide powerful visualization tools and deeper insights into G×E interaction patterns compared with classical ANOVA-based methods. However, these techniques generally rely on balanced datasets and are based on fixed-effect models, which may limit their applicability when dealing with unbalanced trials or complex experimental designs. As a result, they are often complemented with mixed-model approaches and modern predictive methods that can better accommodate heterogeneous variance structures and large-scale breeding data.

3.4. Mixed Models and BLUP

Linear mixed models have become an important extension of classical approaches for analyzing GEI in multi-environment trials (Smith et al., 2005; Piepho et al., 2008). Unlike traditional fixed-effect models, mixed models allow certain factors to be treated as random effects, enabling a more flexible and realistic representation of biological and experimental variation (Patterson & Thompson, 1971; Robinson, 1991). In plant breeding experiments, genotypes, environments, and their interactions can be modeled as either fixed or random depending on the objectives of the study. Mixed models are typically estimated using the restricted maximum likelihood (REML) method, which provides unbiased estimates of variance components associated with genotype, environment, and G×E interaction. A key advantage of mixed models is their ability to generate Best Linear Unbiased Predictions (BLUPs) for genotypic performance. BLUPs combine information from multiple environments while accounting for environmental variability and experimental error, producing more reliable estimates of genotype performance than simple arithmetic means. By borrowing strength across environments and adjusting for imbalance in the data, BLUPs improve the accuracy of genotype ranking and selection decisions (Henderson, 1975; Bernardo, 2010). This is particularly important in breeding programs where trials are often conducted across many locations and years with missing observations.

Mixed models also allow breeders to explicitly model heterogeneous variances and correlation structures among environments. For example, spatial variation within trials, heterogeneity of error variances across environments, and genetic correlations between environments can be incorporated into the model. These features make mixed models especially useful for analyzing large and complex MET datasets, where classical approaches such as ANOVA may fail to adequately capture the underlying data structure. Furthermore, the mixed-model framework provides the foundation for advanced methodologies such as factor analytic models, which are widely used to describe the structure of G×E interactions in breeding programs (Smith et al., 2001; van Eeuwijk et al., 2016). These models can identify patterns of genetic correlation among environments and improve prediction of genotype performance in untested environments. As a result, mixed models and BLUP-based approaches have become central tools in modern plant breeding for evaluating genotype stability, predicting performance, and supporting cultivar recommendation across diverse environments.

4. Rationale for Machine Learning in G×E Analysis

The increasing availability of large-scale phenotypic, genomic, and environmental datasets has created new opportunities for improving the analysis and prediction of GEI in plant breeding (Crossa et al., 2017; Hickey et al., 2019). Traditional statistical approaches such as analysis of variance, regression-based stability models, and multivariate methods have played a fundamental role in understanding GEI. However, these approaches often rely on assumptions of linearity, simple interaction structures, and relatively small numbers of predictors. As breeding programs increasingly generate high-dimensional data from genomic markers, environmental covariates, and high-throughput phenotyping platforms, these assumptions may limit the ability of classical methods to fully capture the complexity of genotype responses across diverse environments. Some features of classical and machine learning models are presented in Table 1. ML methods provide an alternative analytical framework capable of addressing these challenges by modeling complex, nonlinear relationships and high-order interactions without requiring strong parametric assumptions (Heslot et al., 2012; Ma et al., 2018).

Table 1. Features of classical and machine learning models.

Feature	Classical Models	Machine Learning	Deep Learning
Data size handling	Small–moderate	Large	Very large
Relationship modeling	Mostly linear	Nonlinear	Highly nonlinear
G×E interaction modeling	Explicit statistical models	Implicit via algorithm	Learned automatically
Interpretability	High	Moderate	Low
Prediction accuracy	Moderate	High	Very high
Computational requirement	Low	Medium	High

One key motivation for applying machine learning to G×E analysis lies in its ability to model nonlinear and complex interaction structures. In real agricultural systems, genotype responses to environmental conditions are rarely purely linear. Environmental variables such as temperature, rainfall distribution, soil properties, and management practices interact in complex ways that influence phenotypic expression. Classical regression-based approaches typically assume additive or low-order interaction effects, which may oversimplify biological reality. In contrast, machine learning algorithms such as random forests, gradient boosting machines, support vector machines, and deep neural networks can automatically capture nonlinear patterns and higher-order interactions among predictors (Washburn et al., 2021; Reinoso-Peláez et al., 2022). This flexibility allows ML models to better represent the intricate relationships among genotype, environment, and phenotype, thereby potentially improving the accuracy of performance predictions across heterogeneous environments.

Another major rationale for adopting machine learning approaches is the high dimensionality of genomic and environmental covariates available in modern breeding programs. Advances in genotyping technologies now allow breeders to access thousands to millions of molecular markers, while environmental characterization can include numerous climatic, soil, and management variables (Meuwissen et al., 2001; Crossa et al., 2017). Integrating these diverse datasets within a single analytical framework presents a substantial statistical challenge, particularly when the number of predictors greatly exceeds the number of observations. Many machine learning algorithms are specifically designed to handle such high-dimensional data through techniques such as regularization, feature selection, and ensemble learning (Hastie et al., 2009; Montesinos-López et al., 2018). These capabilities enable ML models to extract informative patterns from complex datasets while mitigating problems such as multicollinearity and overfitting that commonly arise in traditional regression-based analyses.

Machine learning also aligns well with the shift from inferential to predictive breeding frameworks. Conventional statistical methods in plant breeding have primarily focused on inference-testing hypotheses about genotype performance, partitioning variance components, and identifying statistically significant interactions. While these objectives remain important for understanding biological processes, modern breeding increasingly emphasizes prediction that accurately forecasting the performance of new genotypes in untested environments (Meuwissen et al., 2001; Crossa et al., 2017). Predictive modeling is particularly critical in genomic selection and climate-resilient breeding strategies, where the goal is to accelerate genetic gain by selecting superior genotypes based on predicted performance rather than solely on observed data. Machine learning methods are inherently prediction-oriented and are often optimized using cross-validation and other predictive performance metrics (Heslot et al., 2012; Montesinos-López et al., 2022), making them well suited for breeding programs that prioritize decision-making based on predictive accuracy.

Furthermore, the integration of machine learning with G×E analysis supports the development of environment-specific breeding strategies. By leveraging detailed environmental covariates, ML models can identify patterns of genotype adaptation across environmental gradients and help define environmental clusters or “mega-environments” (van Eeuwijk et al., 2016; Jarquín et al., 2014**)**. This information can guide targeted cultivar deployment, improve recommendations for specific agro-ecological zones, and enhance the identification of genotypes with broad or specific adaptation (Crossa et al., 2017; Washburn et al., 2021). As climate variability continues to increase, the ability of machine learning models to incorporate dynamic environmental information becomes particularly valuable for anticipating genotype performance under future conditions. Despite these advantages, machine learning approaches should be viewed as complementary rather than replacements for traditional statistical models (van Eeuwijk et al., 2016). Classical methods provide interpretable frameworks for understanding variance components, stability parameters, and biological mechanisms underlying GEI. Machine learning models, on the other hand, excel at capturing complex patterns and improving predictive performance but may offer limited interpretability. Consequently, integrating ML approaches with established statistical methodologies can provide a more comprehensive framework for both understanding and predicting genotype performance across environments, ultimately supporting more efficient and data-driven plant breeding decisions (Crossa et al., 2017; Montesinos-López et al., 2022).

5. Machine Learning Techniques for G × E Analysis

Recent advances in data collection technologies including genomic sequencing, environmental monitoring, and high-throughput phenotyping have generated large and complex datasets in plant breeding programs (Araus & Cairns, 2014; Crossa et al., 2017; Montesinos-López et al., 2022). These datasets often contain nonlinear relationships, high-dimensional predictors, and complex GEI structures that are difficult to capture using traditional statistical approaches (van Eeuwijk et al., 2016). ML provides a flexible computational framework capable of modeling such complexities while emphasizing predictive accuracy (Hastie et al., 2009; Heslot et al., 2012). These approaches used in G×E analysis can generally be categorized into supervised learning, unsupervised learning, and deep learning approaches, each serving distinct purposes in understanding and predicting genotype performance across diverse environments.

5.1. Supervised Learning Approaches

Supervised learning methods are among the most widely applied machine learning techniques in plant breeding for G×E analysis (Heslot et al., 2012; Crossa et al., 2017). These methods learn predictive relationships between input variables (e.g., genomic markers, environmental covariates, and management factors) and output variables such as yield, stress tolerance, or other agronomic traits. Once trained, the models can be used to predict the performance of genotypes in untested environments or future seasons. Random Forest (RF) is an ensemble learning method that constructs a large number of decision trees and aggregates their predictions to improve accuracy and robustness (Breiman, 2001). RF models are particularly useful for G×E studies because they can capture nonlinear relationships and complex interactions among predictors without requiring explicit model specification (Heslot et al., 2012; Washburn et al., 2021). Additionally, random forests provide measures of variable importance, which can help identify key genomic regions or environmental factors influencing genotype performance. Support Vector Machines (SVM) are another powerful supervised learning method used for both regression and classification tasks. SVM models map input data into a high-dimensional feature space using kernel functions, enabling them to capture nonlinear relationships between predictors and phenotypic traits (Cortes & Vapnik, 1995). In the context of G×E analysis, SVMs have been used to model genotype responses across environmental gradients and to improve prediction accuracy in genomic selection frameworks (Heslot et al., 2012). Gradient Boosting Machines (GBM), including popular implementations such as extreme gradient boosting (XGBoost), build predictive models sequentially by combining multiple weak learners, typically decision trees (Friedman, 2001). Each new model focuses on correcting the errors of the previous one, resulting in highly accurate predictive models. Gradient boosting methods have shown strong performance in predicting complex traits influenced by multiple interacting genomic and environmental factors (Chen & Guestrin, 2016; Montesinos-López et al., 2018). Artificial Neural Networks (ANNs) represent another supervised learning approach capable of modeling nonlinear relationships between predictors and phenotypic outcomes. ANNs consist of interconnected layers of computational nodes that transform input data through nonlinear activation functions (Haykin, 1994). In plant breeding applications, neural networks have been used to integrate genomic and environmental information to predict genotype performance under varying environmental conditions (Ma et al., 2018).

5.2. Unsupervised Learning Approaches

Unsupervised learning methods analyze data without predefined response variables, focusing instead on identifying hidden patterns, structures, or groupings within the data (Hastie et al., 2009). These methods are especially useful in G×E analysis for exploring environmental similarities, identifying genotype groups, and simplifying complex datasets. Clustering techniques, such as k-means clustering and hierarchical clustering, are commonly used to group environments or genotypes based on similarity in environmental conditions or performance patterns (Hartigan & Wong, 1979; Everitt et al., 2011). In multi-environment trials, clustering methods can help identify mega-environments, which are groups of locations with similar genotype responses (Yan & Kang, 2003; van Eeuwijk et al., 2016). Such classifications assist breeders in developing region-specific cultivar recommendations and improving selection strategies.

Dimensionality reduction techniques are also widely used to simplify high-dimensional datasets while preserving the most important sources of variation. Methods such as principal component analysis (PCA) (Jolliffe, 2002), t-distributed stochastic neighbor embedding (t-SNE) (van der Maaten & Hinton, 2008), and other manifold learning techniques reduce complex datasets into lower-dimensional representations that are easier to visualize and analyze. These approaches are particularly useful for exploring relationships among environments, genotypes, and environmental covariates. Another important application of unsupervised learning is environment stratification, which involves grouping environments based on climatic, soil, and management variables (Cooper & Hammer, 1996; van Eeuwijk et al., 2016). By identifying environmental clusters with similar conditions, breeders can design more efficient testing networks and improve the targeting of genotypes to specific production zones. Environment stratification also helps reduce noise in G×E analysis by focusing comparisons within relatively homogeneous environmental groups.

5.3. Deep Learning Models for G×E Prediction

Deep learning is a specialized branch of machine learning that uses multi-layer neural networks to automatically learn complex hierarchical representations of data (LeCun et al., 2015). These models are particularly well suited for handling large, heterogeneous datasets such as genomic sequences, temporal climate records, spatial imagery, and high-throughput phenotyping data (Montesinos-López et al., 2022; Ma et al., 2018). In G×E prediction, deep learning models can incorporate temporal environmental data, such as time-series weather variables throughout the growing season (Khaki & Wang, 2019). Recurrent neural networks (RNNs) and related architectures like long short-term memory (LSTM) networks are designed to capture sequential dependencies in time-series data (Hochreiter & Schmidhuber, 1997), enabling them to model how environmental conditions at different growth stages influence crop performance. Deep learning methods are also effective for analyzing spatial data, including satellite imagery, drone-based observations, and field sensor measurements. Convolutional neural networks (CNNs) are widely used for extracting spatial features from images (Krizhevsky et al., 2012) and have been applied to tasks such as canopy characterization, disease detection, and biomass estimation in crop breeding trials (Kamilaris & Prenafeta-Boldú, 2018).

Furthermore, deep learning facilitates the integration of high-throughput phenotyping (HTP) data, which may include hyperspectral imaging, thermal imaging, and LiDAR measurements (Araus & Cairns, 2014; Singh et al., 2016). These technologies generate massive datasets that capture detailed plant characteristics across time and space. Deep neural networks can process these complex data sources and link them with genomic and environmental variables to improve the prediction of genotype performance under diverse environmental conditions. Conceptual diagram of deep learning framework for G×E genomic prediction is presented in Figure 2. The integration of deep learning with genomic and environmental datasets represents a promising frontier for next-generation predictive breeding, enabling breeders to better understand complex G×E interactions and accelerate the development of climate-resilient crop varieties.

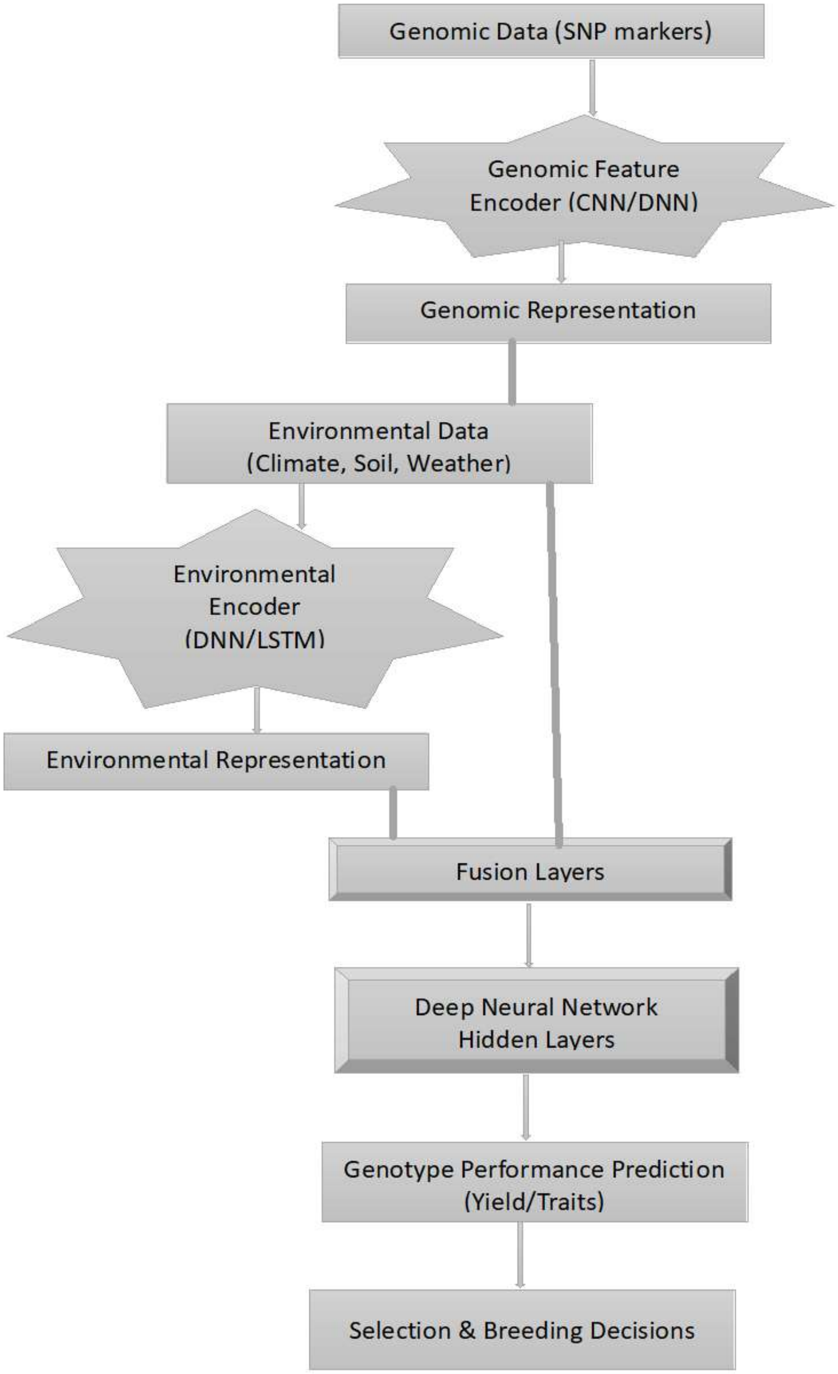

Figure 2. Conceptual diagram: Deep learning framework for G×E genomic prediction.

6. Integration of Genomic, Environmental, and Phenotypic Information

Advances in plant breeding have generated large and diverse datasets that include genomic information, environmental measurements, and detailed phenotypic observations (Crossa et al., 2017; Hickey et al., 2019). Understanding GEI increasingly requires the integration of these multiple sources of information rather than analyzing them separately. Integrative analytical approaches enable breeders to better capture the biological complexity underlying crop performance, improve predictive accuracy, and accelerate genetic gain (van Eeuwijk et al., 2016; Montesinos-López et al., 2022). The combination of genomic prediction models, envirotyping strategies, and machine learning–based multi-modal data integration has therefore emerged as a promising framework for studying and predicting genotype performance across diverse environmental conditions. The conceptual diagram of ML framework for G×E analysis is presented in Figure 3.

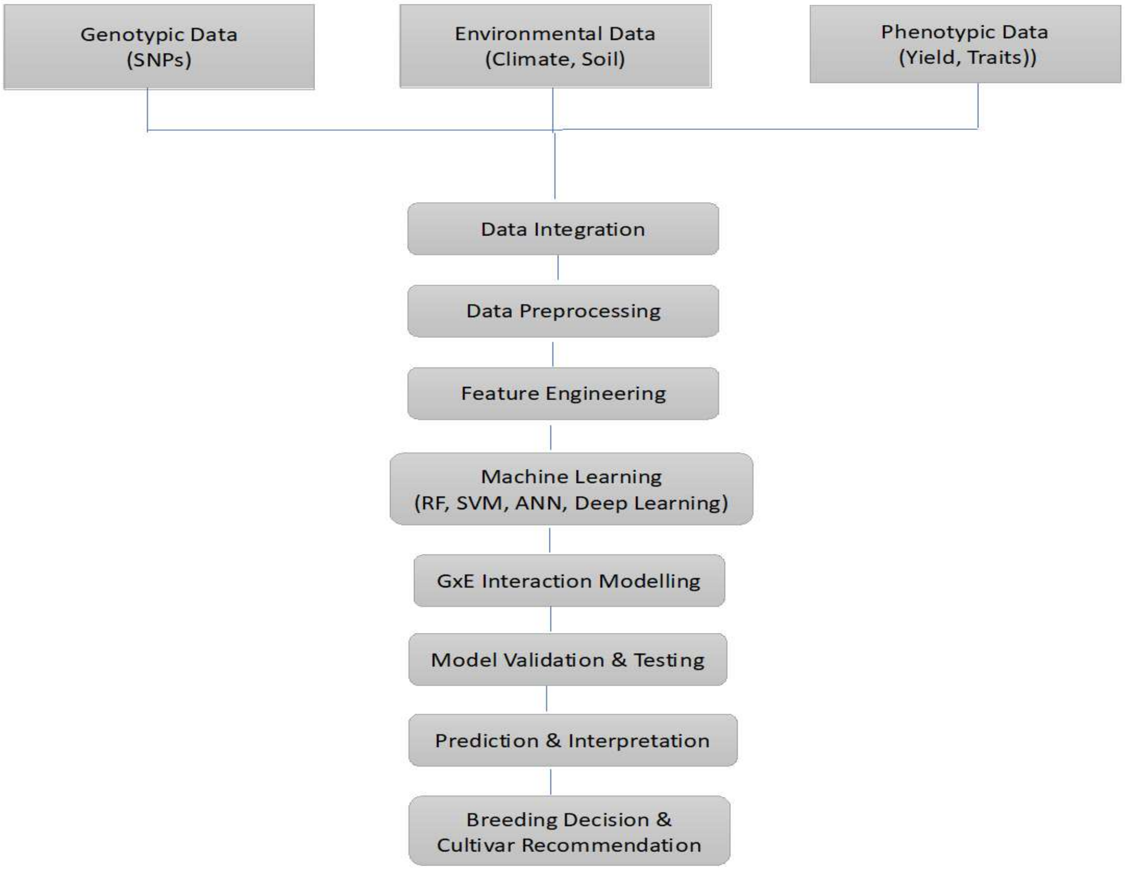

Figure 3. Conceptual diagram of ML framework for G×E analysis.

Genomic prediction has become a central tool in modern breeding programs, allowing the estimation of breeding values using genome-wide molecular markers (Meuwissen et al., 2001). Traditional genomic prediction models often assume that marker effects remain constant across environments; however, in multi-environment trials, genotype performance is strongly influenced by environmental variability. Consequently, incorporating G×E effects into genomic prediction models has become increasingly important (Jarquín et al., 2014; Crossa et al., 2017). Approaches such as reaction norm models, marker-by-environment interaction models, and multi-environment genomic prediction frameworks allow marker effects to vary across environments, thereby capturing differences in genotype responses under contrasting environmental conditions (Jarquín et al., 2014). These models improve prediction accuracy and enable breeders to identify genotypes with either broad adaptation across environments or specific adaptation to particular agro-ecological zones. Such information is particularly valuable for traits like yield and stress tolerance, where environmental factors strongly influence phenotypic expression.

A critical component of integrating environmental information into G×E analysis is envirotyping, which involves the systematic characterization of environmental conditions that influence crop development and productivity (Xu, 2016; Costa-Neto et al., 2021). Envirotyping uses environmental covariates such as temperature, rainfall, solar radiation, humidity, soil properties, and

doi:10.20944/preprints202603.1584.v1

management practices to describe the environments in which crops are evaluated (van Eeuwijk et al., 2016; Cooper et al., 2014). Advances in environmental monitoring systems, remote sensing technologies, and global climate databases have greatly increased the availability of high-resolution environmental data. These climate and environmental covariates can be incorporated into predictive models to represent environmental variability more accurately and to improve the ability to predict genotype responses under different climatic conditions. Additionally, envirotyping facilitates the identification of environmental similarity among testing locations, enabling the classification of environments into mega-environments and improving the targeting of cultivars to specific production regions (Yan & Tinker, 2006; Crossa et al., 2017). In the context of climate change, the use of environmental covariates also allows breeders to evaluate genotype performance under projected future climate scenarios.

The integration of genomic, environmental, and phenotypic data poses substantial analytical challenges because these datasets often differ in scale, dimensionality, and structure (Montesinos-López et al., 2018). Machine learning approaches provide powerful tools for addressing these challenges through multi-modal data integration, where multiple types of data are analyzed simultaneously within a unified modeling framework. Machine learning models can integrate genomic marker information, environmental covariates, and phenotypic measurements to capture complex relationships among genotype, environment, and phenotype (Hastie et al., 2009; Crossa et al., 2017). Algorithms such as ensemble learning models and deep neural networks are particularly well suited for handling high-dimensional and heterogeneous datasets, enabling them to detect nonlinear relationships and higher-order interactions that may not be easily captured by traditional statistical methods (LeCun et al., 2015; Montesinos-López et al., 2022).

By leveraging multi-modal data integration, machine learning models can generate more comprehensive predictive frameworks that reflect the complex biological processes governing crop performance. For example, genomic marker data can be combined with environmental time-series information and high-throughput phenotyping measurements to improve the prediction of genotype performance across environments (Araus & Cairns, 2014; Jarquín et al., 2014). Such integrative models enhance the ability of breeders to identify adaptive traits, optimize selection strategies, and design more efficient multi-environment testing programs (Hickey et al., 2019; Crossa et al., 2017). The integration of genomic, environmental, and phenotypic information represents a critical step toward next-generation predictive breeding systems capable of addressing the challenges of climate variability, resource limitations, and the increasing demand for sustainable agricultural productivity.

7. Model Evaluation, Validation, and Interpretability in G × E Analysis

As machine learning and advanced statistical models become increasingly applied to genotype × environment (G×E) studies, rigorous procedures for model evaluation, validation, and interpretability are essential to ensure reliability and practical usefulness in plant breeding. Unlike traditional inferential models that primarily focus on hypothesis testing and parameter estimation, predictive models must be carefully validated to assess their ability to generalize to new genotypes or environments. In the context of METs, evaluation frameworks must consider the hierarchical and structured nature of the data, where observations are nested within genotypes, locations, and years. Consequently, specialized cross-validation strategies, appropriate predictive performance metrics, and interpretable modeling approaches play a critical role in assessing and improving predictive models for G×E analysis. Cross-validation is widely used to evaluate predictive models by partitioning the dataset into training and validation subsets (Schrauf et al., 2021). However, standard random cross-validation may not be appropriate for multi-environment trials because it can lead to overly optimistic estimates of prediction accuracy when observations from the same genotype or environment appear in both training and testing sets. To address this issue, several cross-validation strategies have been developed specifically for G×E prediction studies. These strategies simulate realistic breeding scenarios, such as predicting the performance of new genotypes in known environments, predicting existing genotypes in new environments, or predicting new genotype–

environment combinations that have not yet been observed (Roorkiwal et al., 2018). For example, leave-one-environment-out validation assesses a model's ability to predict genotype performance in environments that were not included during model training, while leave-one-genotype-out validation evaluates the prediction of untested genotypes. These cross-validation schemes provide more realistic estimates of model performance and better reflect practical breeding decisions (Montesinos-López et al., 2026).

In addition to validation strategies, the evaluation of predictive models relies on quantitative performance metrics that measure the agreement between predicted and observed phenotypic values. Commonly used metrics include the correlation between predicted and observed values, root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination. These metrics help quantify both the accuracy and reliability of predictions across environments (Pan et al., 2024). In genomic prediction and machine learning–based G×E studies, prediction accuracy is often assessed as the correlation between predicted breeding values and observed phenotypic performance. Such metrics enable researchers to compare different modeling approaches and identify methods that provide the most reliable predictions across diverse environmental conditions (Ould Estaghvirou et al., 2013).

While many machine learning models offer high predictive accuracy, they are often criticized for their limited interpretability. Complex models such as ensemble algorithms and deep neural networks are sometimes described as "*black boxes*," making it difficult to understand how specific variables influence predictions. To address this challenge, the emerging field of explainable artificial intelligence (XAI) has introduced techniques that improve the interpretability of machine learning models (Lunberg and Lee 2017; Lipton 2018; Rudin 2019; Yu et al., 2025). These methods aim to reveal how input variables such as genomic markers, environmental covariates, or phenotypic traits contribute to model predictions. Feature importance measures can identify the most influential genomic or environmental variables affecting crop performance. Partial dependence plots illustrate how predicted outcomes change across the range of a particular variable while holding others constant. More advanced techniques, such as Shapley additive explanations (SHAP), provide a consistent framework for quantifying the contribution of each predictor to individual model predictions. These interpretability tools allow researchers to better understand the biological and environmental factors underlying genotype performance and to gain insights into complex GEI. Improving model interpretability is particularly important in plant breeding because breeders require not only accurate predictions but also actionable biological insights. Explainable models can help identify key environmental drivers of genotype adaptation, reveal genomic regions associated with stress tolerance, and support decision-making in cultivar selection and deployment (Novielli et al., 2025; Yu et al., 2025). By combining robust validation strategies, appropriate performance metrics, and explainable artificial intelligence techniques, researchers can ensure that machine learning models used in G×E analysis are both predictively reliable and biologically informative, ultimately enhancing their utility in modern crop improvement programs.

8. Applications of Machine Learning–Based G × E Analysis in Crop Improvement

The integration of ML approaches into G×E analysis has opened new opportunities for improving crop breeding and accelerating genetic gain (Montesinos-López et al., 2018; Washburn et al., 2021; Crossa et al., 2017). By leveraging large datasets that combine genomic, environmental, and phenotypic information, ML-based G×E models can better capture complex patterns of genotype performance across diverse environments. Some comparison of machine learning algorithms used in G×E analysis is presented in Table 2. These algorithms are increasingly being applied in crop improvement programs to enhance the prediction of genotype performance, identify adaptive traits, and support the development of cultivars with improved yield stability and resilience to climatic variability.

Table 2. Comparison of machine learning algorithms used in G × E analysis.

ML Algorithm	Key Characteristics	Advantages for G × E Analysis	Limitations	References
Random Forest (RF)	Ensemble tree-based algorithm	Handles nonlinear relationships; robust to noise; suitable for high-dimensional genomic data	Less interpretable; may overfit small datasets	González-Camacho et al., 2018; Montesinos-López et al., 2018
Support Vector Machine (SVM)	Kernel-based regression and classification	Effective in high-dimensional genomic datasets; captures nonlinear relationships	Requires careful parameter tuning; computationally intensive	Heslot et al., 2012
Gradient Boosting Machines (GBM)	Sequential ensemble learning method	High prediction accuracy; captures complex interactions	Risk of overfitting; high computational demand	Montesinos-López et al., 2019
Artificial Neural Networks (ANN)	Multilayer neural network	Captures complex nonlinear G×E interactions	Requires large datasets; limited interpretability	Heslot et al., 2012
Deep Neural Networks (DNN)	Multi-layer deep learning architecture	Models complex genomic–environmental relationships; high predictive power	Requires large training datasets and computing resources	Ma et al., 2018; Washburn et al., 2021
Convolutional Neural Networks (CNN)	Deep learning architecture detecting local patterns	Effective for genomic marker structure and spatial genomic patterns	Large computational requirement	Ma et al., 2018
Recurrent Neural Networks (RNN/LSTM)	Neural network for sequential data	Captures temporal environmental variation (e.g., weather patterns)	Complex training; large datasets required	Washburn et al., 2021
k-Nearest Neighbors (k-NN)	Distance-based learning method	Simple implementation; no assumptions about data distribution	Sensitive to noise and large datasets	Heslot et al., 2012

Evidence from several major crop species demonstrates the growing importance of machine learning in addressing the challenges associated with environmental heterogeneity and climate change. In crops such as maize, wheat, and rice, machine learning models have been used to improve the prediction of genotype performance across multi-environment trials by integrating genomic markers with environmental covariates (Montesinos-López et al., 2018; Li et al., 2018; Washburn et al., 2021). Table 3 shows various machine learning–based G×E analysis in different crops. Studies have shown that ensemble learning algorithms and neural network–based models can outperform traditional statistical approaches in capturing nonlinear relationships between genotypes and environmental factors (Montesinos-López et al., 2019; Heslot et al., 2012). For example, machine learning methods have been successfully applied to predict grain yield across variable climatic conditions by incorporating weather data, soil characteristics, and management practices alongside genomic information. These predictive models enable breeders to evaluate large numbers of candidate genotypes more efficiently and to prioritize those with superior performance across diverse environments (Fernandes et al., 2024).

Table 3. Machine learning–based G×E analysis in different crops.

Crop	Machine Learning Method	Objective/Trait	Key Findings	Reference
Wheat	Random Forest (RF), Gradient Boosting	Grain yield prediction across environments	ML models improved prediction accuracy compared with traditional statistical models in multi-environment trials	Montesinos-López et al., 2018
Maize	Deep Neural Networks (DNN), CNN	Genomic prediction under G×E	Deep learning captured nonlinear genomic × environmental interactions and improved yield prediction	Washburn et al., 2021
Rice	Support Vector Machine (SVM), Random Forest	Yield prediction and stability analysis	ML methods effectively modeled complex environmental effects influencing rice productivity	Li et al., 2018
Wheat	Gradient Boosting Machines (GBM), RF	Genomic selection and yield prediction	ML approaches showed higher predictive ability when integrating environmental covariates	Jarquin et al., 2014
Barley	Artificial Neural Networks (ANN)	Yield stability across environments	ANN captured nonlinear genotype responses and improved prediction of adaptation patterns	Heslot et al., 2012
Potato	Random Forest, SVM	Genotype performance prediction	Machine learning models improved prediction of tuber yield across locations	Tatsumi & Usami, 2024
Soybean	Deep Learning, CNN	Crop yield prediction	Deep learning models handled complex relationships	Khalilzadeh et al., 2024
Cotton	Random Forest, XGboost	Lint yield prediction	ML models improved predictive accuracy	Dhaliwal et al., 2022

In addition to improving predictive accuracy, machine learning–based G×E analyses contribute to a better understanding of environmental drivers influencing crop performance. By identifying key environmental variables that interact with genotype performance, these models help breeders determine the environmental conditions under which particular genotypes perform best (Jarquín et al., 2014; Crossa et al., 2017). This information is valuable for defining mega-environments, optimizing trial locations, and improving cultivar recommendation strategies. Furthermore, ML approaches can assist in identifying genotype–environment combinations that maximize productivity, thereby improving the efficiency of breeding programs and cultivar deployment (Fernandes et al., 2024).

Machine learning models have also been used to support the selection of genotypes with greater yield stability across environments (İPEKEŞEN et al., 2024), which is a major objective in crop breeding. Stability analysis traditionally relies on statistical methods that evaluate genotype performance across multiple environments; however, ML models can capture more complex interaction patterns and detect subtle relationships between environmental variables and genotype responses (González-Camacho et al., 2018). By analyzing large multi-environment datasets, machine learning methods can identify genotypes that consistently perform well under a wide range of environmental conditions or those that are specifically adapted to particular stress environments such as drought or heat.

Another important application of machine learning–based G×E analysis is the development of climate-resilient crop varieties. As climate change increases the frequency and intensity of environmental stresses, breeders must develop cultivars capable of maintaining productivity under unpredictable conditions. Machine learning models can incorporate long-term climatic data, environmental covariates, and genomic information to predict how genotypes are likely to perform under future climate scenarios (Ma et al., 2018; Crossa et al., 2017). This capability enables breeders to anticipate environmental challenges and select genotypes with traits that enhance tolerance to heat, drought, or other stress factors. Furthermore, the integration of high-throughput phenotyping and remote sensing data with machine learning models has expanded the scope of G×E analysis in crop improvement. Data collected from drones, satellites, and field sensors provide detailed information about crop growth, canopy structure, and stress responses over time. When combined with genomic and environmental data, these datasets allow machine learning models to capture dynamic genotype responses throughout the growing season, leading to more accurate predictions of crop performance and improved understanding of adaptation mechanisms (Araus & Cairns, 2014; Ma et al., 2018).

9. Challenges and Limitations

Despite the growing interest in applying ML approaches to G×E analysis, several challenges and limitations remain that affect their effectiveness and practical implementation in crop improvement programs. One of the most significant constraints in ML-based G×E analysis is the availability and quality of data. Machine learning algorithms typically require large, high-quality datasets to effectively learn complex patterns and interactions among genomic, environmental, and phenotypic variables (Crossa et al., 2017; Crossa et al., 2014). However, in many breeding programs, especially in developing regions, multi-environment trial datasets may be limited in size or contain missing and inconsistent observations. Environmental data may also be incomplete or measured at low temporal and spatial resolution, reducing the ability to accurately characterize environmental variability. Additionally, phenotypic measurements may be affected by experimental errors, inconsistent trial management, or differences in data collection protocols across locations and seasons. Such limitations can reduce the predictive performance of machine learning models and may lead to biased or unreliable predictions. Ensuring standardized data collection protocols, improving environmental monitoring systems, and integrating high-throughput phenotyping technologies are therefore critical steps toward improving data quality for ML-based G×E analysis.

Another important challenge concerns model transparency and reproducibility. Many machine learning algorithms, particularly complex ensemble models and deep neural networks, operate as "*black-box*" systems in which the relationships between input variables and predictions are not easily interpretable (Lipton, 2018; Rudin, 2019; Pineau et al., 2021). This lack of transparency can make it difficult for breeders and researchers to understand the biological mechanisms underlying model predictions or to identify which genomic and environmental factors drive genotype performance. Moreover, reproducibility can be compromised if model training procedures, parameter settings, or preprocessing steps are not clearly documented. Variability in computational environments, software implementations, and data preprocessing pipelines can lead to differences in model results, making it challenging to replicate studies or compare results across research groups. To address these issues, researchers increasingly emphasize transparent reporting practices, open data sharing, and the use of explainable artificial intelligence methods that help clarify how machine learning models generate predictions.

A further limitation relates to the practical adoption of machine learning methods in breeding programs. While ML models can offer improved predictive performance, their implementation often requires specialized computational skills, access to high-performance computing resources, and expertise in data science. Many breeding programs, particularly in public research institutions or developing countries, may lack the technical infrastructure or trained personnel needed to apply advanced machine learning methods effectively. In addition, breeders may prefer analytical approaches that provide clear biological interpretation and straightforward decision-making criteria. If machine learning models are perceived as complex, opaque, or difficult to integrate into existing workflows, their adoption may remain limited despite their potential advantages. Bridging the gap between machine learning research and practical breeding applications therefore requires the development of user-friendly analytical tools, standardized workflows, and interdisciplinary collaboration between plant breeders, statisticians, and data scientists. Integrating machine learning models with established statistical frameworks and breeding pipelines can also facilitate their acceptance and practical use. As computational tools become more accessible and training opportunities expand, the effective application of machine learning in G×E analysis is expected to grow, ultimately supporting more efficient and data-driven crop improvement programs.

10. Future Perspectives

The integration of ML into G×E analysis represents an evolving frontier in plant breeding, and several promising research directions are expected to shape its future development. Advances in computational biology, environmental data acquisition, and genomic technologies are likely to expand the role of ML in understanding complex genotype responses across diverse environments. Future research will increasingly focus on hybrid analytical frameworks, improved interpretability of machine learning models, integration of climate change projections, and the development of real-time decision support systems to enhance breeding efficiency. Conventional statistical models, particularly linear mixed models and genomic prediction methods, have strong theoretical foundations and provide interpretable estimates of genetic effects and variance components. However, they may have limited capacity to capture complex nonlinear interactions between genotypes and environments. Hybrid models that integrate mixed-model methodology with machine learning algorithms can potentially leverage the strengths of both approaches (Figure 4). In such frameworks, mixed models can be used to account for experimental design structures, random effects, and genetic relationships, while machine learning algorithms can capture nonlinear patterns and higher-order interactions among genomic and environmental variables. This integration is expected to improve predictive accuracy while maintaining the statistical rigor required for breeding decisions.

Another critical area of future research involves the advancement of explainable machine learning approaches for biological interpretation. While many ML models offer strong predictive performance, their "*black-box*" nature often limits their usefulness for understanding the biological mechanisms underlying G×E interactions. Emerging techniques in explainable artificial intelligence (XAI), such as feature importance analysis, partial dependence plots, and model-agnostic interpretation methods, can help reveal how genomic markers, environmental variables, and management factors contribute to genotype performance. By improving interpretability, these methods enable breeders to identify key drivers of adaptation and stability, thereby bridging the gap between predictive modeling and biological insight. As explainable ML methods continue to evolve, they are expected to enhance confidence in ML-based predictions and support more informed selection decisions.

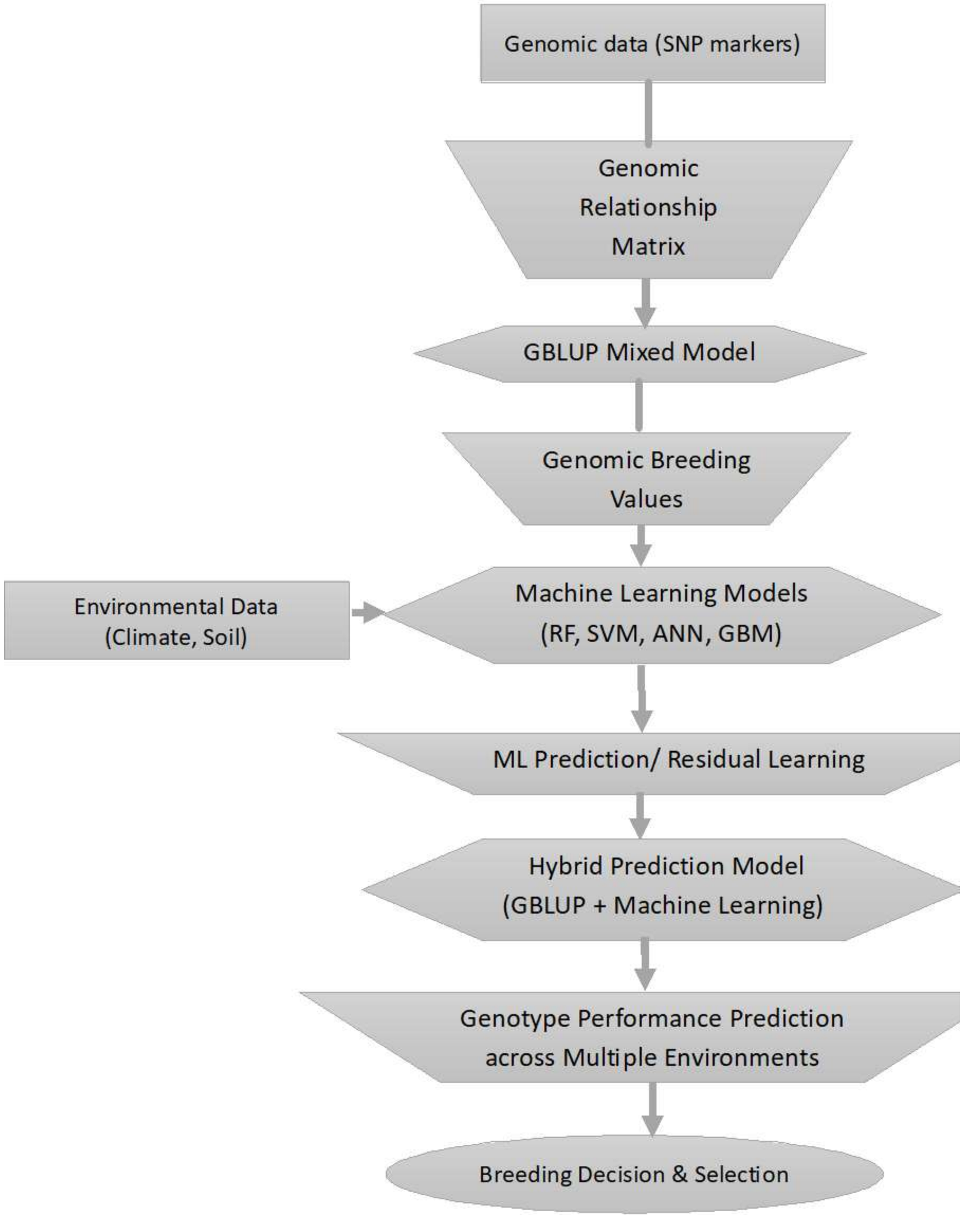

Figure 4. Conceptual diagram: Hybrid ML–GBLUP framework.

The integration of climate change scenarios into G×E prediction models is also likely to become increasingly important. Climate variability and long-term changes in temperature, precipitation patterns, and extreme weather events are expected to significantly influence crop performance and agricultural sustainability. Machine learning models that incorporate climate projections, environmental covariates, and historical weather data can help simulate genotype performance under future climatic conditions. Such predictive frameworks can assist breeders in identifying genotypes with improved resilience to environmental stresses and in designing breeding strategies aimed at developing climate-adapted cultivars. The combination of ML-based G×E models with climate forecasting tools may therefore play a key role in supporting climate-smart agriculture.

Finally, future advances will likely lead to the development of real-time decision support systems for plant breeders. With the increasing availability of high-throughput phenotyping platforms, remote sensing technologies, and environmental monitoring systems, large volumes of data can be generated continuously during the crop growing season. Machine learning algorithms integrated into digital breeding platforms can analyze these data streams to provide real-time insights into genotype performance, environmental conditions, and potential stress factors. Such systems could support dynamic decision-making processes, including genotype selection, trial site allocation, and adaptive management strategies. By combining predictive modeling with user-friendly interfaces and automated data integration, real-time decision support tools have the potential to transform breeding programs into more efficient, data-driven systems.

11. Conclusions

GEI analysis remains a fundamental component of plant breeding because it enables breeders to understand how genotypes respond to environmental variability and to identify cultivars with superior adaptation and stability. Classical statistical approaches have played a critical role in quantifying G×E patterns and guiding cultivar evaluation across multi-environment trials. These methods continue to provide a strong theoretical foundation for understanding genotype performance and remain valuable tools for inference and interpretation in breeding research. However, the increasing availability of genomic, environmental, and high-throughput phenotypic data has introduced new challenges that extend beyond the capabilities of many traditional analytical methods. Machine learning approaches provide powerful alternatives for modeling complex and nonlinear relationships among these diverse data sources. By effectively capturing intricate interaction patterns and integrating high-dimensional datasets, machine learning models can significantly improve the prediction of genotype performance across diverse environments. The integration of machine learning with genomic prediction, environmental covariates, and advanced phenotyping technologies represents a transformative shift toward data-driven breeding strategies. These approaches enable more accurate prediction of genotype performance, support the identification of stable and widely adapted cultivars, and facilitate the development of varieties with enhanced resilience to environmental variability and climate change. As computational tools, data infrastructure, and interdisciplinary collaboration continue to advance, the combined use of statistical and machine learning approaches will play an increasingly important role in accelerating crop improvement and ensuring sustainable agricultural productivity in diverse and changing environments.

References

Allard, R. W., & Bradshaw, A. D. (1964). Implications of genotype–environmental interactions in applied plant breeding. *Crop Science*, 4, 503–508.

Annicchiarico, P. (2002). Genotype × Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations. FAO Plant Production and Protection Paper No. 174, FAO, Rome.

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, 19, 52–61.

Baker, R. J. (1988). Tests for crossover genotype–environmental interactions. *Canadian Journal of Plant Science*, 68, 405–410.

Becker, H. C., & Léon, J. (1988). Stability analysis in plant breeding. *Plant Breeding*, 101, 1–23.

Bernardo, R. (2010). Breeding for Quantitative Traits in Plants. 3rd Edition, Stemma Press, Woodbury.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.

doi:10.20944/preprints202603.1584.v1

Ceccarelli, S. (1989). Wide adaptation: How wide? *Euphytica*, 40, 197–205.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, San Francisco, 13-17 August 2016, 785-794. https://doi.org/10.1145/2939672.2939785

Cooper, M., & Hammer, G. L. (1996). Plant adaptation and crop improvement. CAB International.

Cooper, M., Gho, C., Leafgren, R., Tang, T., & Messina, C. (2014). Breeding drought-tolerant maize hybrids for the US corn-belt: discovery to product. *Journal of Experimental Botany*, 65, 6191–6204, https://doi.org/10.1093/jxb/eru064

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297.

Costa-Neto, G., Fritsche-Neto, R., & Crossa, J. (2021). Nonlinear kernels, dominance, and envirotyping data increase the accuracy of genome-based prediction in multi-environment trials. *Heredity*, 126, 92–106.

Crossa, J. (1990). Statistical analyses of multilocation trials. *Advances in Agronomy*, 44, 55–85.

Crossa, J., Cornelius, P. L., & Yan, W. (2002). Biplots of linear-bilinear models for studying crossover genotype × environment interaction. *Crop Science*, 42, 619–633.

Crossa, J., et al. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends in Plant Science*, 22(11), 961–975.

Crossa, J., Pérez, P., Burgueño, J., et al. (2014). Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity*, 112(1), 48–60.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., et al. (2017). Genomic selection in plant breeding: methods, models, and perspectives. *Trends in Plant Science*, 22(11), 961–975.

Dhaliwal, J. K., Panday, D., Saha, D., Lee, J., Jagadamma, S., Schaeffer, S., & Mengistu, A. (2022). Predicting and interpreting cotton yield and its determinants under long-term conservation management practices using machine learning. *Computers and Electronics in Agriculture*, *199*, 107107.

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

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Wiley.

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). Longman.

Fernandes, I. K., Vieira, C. C., Dias, K. O., & Fernandes, S. B. (2024). Using machine learning to combine genetic and environmental data for maize grain yield predictions across multi-environment trials. *Theoretical and Applied Genetics*, *137*(8), 189.

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

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29, 1189–1232.

Gauch, H. G. (1992). Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier.

González-Camacho, J. M., Crossa, J., Pérez-Rodríguez, P., Ornella, L., & Gianola, D. (2016). Genome-enabled prediction using probabilistic neural network classifiers. *BMC genomics*, *17*(1), 208.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, *28*(1), 100-108.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

Haykin, S. (1994). Neural networks: a comprehensive foundation. Prentice hall PTR.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 423-447.

Heslot, N., Yang, H. P., Sorrells, M. E., & Jannink, J. L. (2012). Genomic selection in plant breeding: a comparison of models. *Crop science*, *52*(1), 146-160.

Hickey, L. T., N. Hafeez, A., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C., Tester, M., ... & Wulff, B. B. (2019). Breeding crops to feed 10 billion. *Nature biotechnology*, *37*(7), 744-754.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

İPEKEŞEN, S., RUFAIOGLU, S. B., & BİÇER, B. T. (2024). Machine Learning in Legume Breeding: Modelling Genotype and Environment Interactions. *Ejons International Journal on Mathematic, Engineering and Natural Sciences*, *8*(4), 493-503.

Jarquín, D., Crossa, J., Lacaze, X., et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics*, 127, 595–607.

Jewiss, J. L., Brown, M. E., & Escobar, V. M. (2020). Satellite remote sensing data for decision support in emerging agricultural economies: How satellite data can transform agricultural decision making [perspectives]. *IEEE Geoscience and Remote Sensing Magazine*, *8*(4), 117-133.

Jolliffe, I. (2025). Principal Component Analysis. In *International Encyclopedia of Statistical Science* (pp. 1945-1948). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, *147*, 70-90.

Kang, M. S. (2002). Quantitative genetics, genomics and plant breeding. CAB International.

Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in plant science*, *10*, 621.

Khalilzadeh, Z., Kashanian, M., Khaki, S., & Wang, L. (2024). A hybrid deep learning-based approach for optimal genotype by environment selection. *Frontiers in Artificial Intelligence*, *7*, 1312115.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.

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*, *115*(26), 6679-6684.

Lin, C. S., Binns, M. R., & Lefkovitch, L. P. (1986). Stability analysis: Where do we stand? *Crop Science*, 26, 894–900.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31-57.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

Ma, W., Qiu, Z., Song, J., Li, J., Cheng, Q., Zhai, J., & Ma, C. (2018). A deep convolutional neural network approach for predicting phenotypes from genotypes. *Planta*, *248*(5), 1307-1318.

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

Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). *Multivariate statistical machine learning methods for genomic prediction* (p. 691). Springer Nature.

Montesinos-López, A., Montesinos-López, O. A., Gianola, D., Crossa, J., & Hernández-Suárez, C. M. (2018). Multi-environment genomic prediction of plant traits using deep learners with dense architecture. *G3: Genes, Genomes, Genetics*, *8*(12), 3813-3828.

Montesinos-López, et al.. (2019). A benchmarking between deep learning, support vector machine and Bayesian threshold best linear unbiased prediction for predicting ordinal traits in plant breeding. *G3: Genes, Genomes, Genetics*, *9*(2), 601-618.

Montesinos-López, O. A., Kismiantini, & Montesinos-López, A. (2026). An alternative approach for predicting genotype performance in novel environments without genomic data. *Discover Agriculture*, *4*(1), 22.

Novielli, P., et al. (2025). Leveraging Explainable Artificial Intelligence for Genotype-to-Phenotype Prediction: A Case Study in Arabidopsis thaliana. *Applied System Innovation*, *8*(6), 164.

Ould Estaghvirou, S. B., Ogutu, J. O., Schulz-Streeck, T., Knaak, C., Ouzunova, M., Gordillo, A., & Piepho, H. P. (2013). Evaluation of approaches for estimating the accuracy of genomic prediction in plant breeding. *BMC genomics*, *14*(1), 860.

Pan, S., Liu, Z., Han, Y., Zhang, D., Zhao, X., Li, J., & Wang, K. (2024). Using the Pearson's correlation coefficient as the sole metric to measure the accuracy of quantitative trait prediction: is it sufficient?. *Frontiers in Plant Science*, *15*, 1480463.

Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58, 545–554.

Piepho, H. P. (1998). Methods for comparing the yield stability of cropping systems. *Journal of Agronomy and Crop Science*, 180, 193–213.

Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, *161*(1), 209-228.

Pineau, J., et al. (2021). Improving reproducibility in machine learning research. *Journal of Machine Learning Research*, 22, 1–20.

Reinoso-Peláez, E. L., Gianola, D., & Gonzalez-Recio, O. (2022). Genome-enabled prediction methods based on machine learning. In *Genomic prediction of complex traits: Methods and protocols* (pp. 189-218). New York, NY: Springer US.

Robinson, G. K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical science*, 15-32.

Roorkiwal, M., et al. (2018). Genomic-enabled prediction models using multi-environment trials to estimate the effect of genotype× environment interaction on prediction accuracy in chickpea. *Scientific reports*, *8*(1), 11701.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1, 206–215.

Schrauf, M. F., de los Campos, G., & Munilla, S. (2021). Comparing genomic prediction models by means of cross validation. *Frontiers in Plant Science*, 12, 734512.

Shukla, G. K. (1972). Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity*, 29, 237–245.

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in plant science*, 21(2), 110-124.

Smith, A. B., Cullis, B. R., & Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science*, 143(6), 449-462.

Smith, A., Cullis, B., & Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*, *57*(4), 1138-1147.

Tatsumi, K., & Usami, T. (2024). Plant-level prediction of potato yield using machine learning and unmanned aerial vehicle (UAV) multispectral imagery. *Discover Applied Sciences*, *6*(12), 649.

van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.

van Eeuwijk, F. A., Bustos-Korts, D., & Malosetti, M. (2016). What should students in plant breeding know about the statistical aspects of genotype × environment interactions? *Crop Science*, 56, 2119–2140.

Washburn, J. D., et al. (2021). Predicting phenotypes from genetic, environment, management, and historical data using CNNs. *Theoretical and Applied Genetics*, *134*(12), 3997-4011.

Wricke, G. (1962). Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. *Zeitschrift für Pflanzenzüchtung*, 47, 92–96.

Xu, Y. (2016). Envirotyping for deciphering environmental impacts on crop plants. *Theoretical and Applied Genetics*, 129, 653–673.

Yan, W., & Kang, M. S. (2003). GGE biplot analysis: A graphical tool for breeders. CRC Press.

Yan, W., & Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canadian journal of plant science*, *86*(3), 623-645.

Yan, W., Hunt, L. A., Sheng, Q., & Szlavnics, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science*, 40, 597–605.

Yan, W., Kang, M. S., Ma, B., Woods, S., & Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop science*, *47*(2), 643-653.

Yates, F., & Cochran, W. G. (1938). The analysis of groups of experiments. *The Journal of Agricultural Science*, *28*(4), 556-580.

Yu, T., et al. (2025). EXGEP: a framework for predicting genotype-by-environment interactions using ensembles of explainable machine-learning models. *Briefings in Bioinformatics*, *26*(4), bbaf414.)

Zobel, R. W., Wright, M. J., & Gauch, H. G. (1988). Statistical analysis of yield trials. Agronomy Journal, 80, 388–393.