Journal of Environmental and Agricultural Studies
ISSN: 2710-1401
DOI: 10.32996/jeas
Journal Homepage: www.al-kindipublisher.com/index.php/jeas

JEAS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Integrating Genomic Selection and Machine Learning: A Data-Driven Approach to Enhance Corn Yield Resilience Under Climate Change

Abu Saleh Muhammad Saimon[1], Mohammad Moniruzzaman[2], Md Shafiqul Islam[2], Md Kamal Ahmed[3], Md Mizanur Rahaman[4], Sazzat Hossain[3], Mia Md Tofayel Gonee Manik[4]✉

[1]*Department of Computer Science, Washington University of Science and Technology, Alexandria VA 22314, USA*
[2]*Department of Computer Science, Maharishi International University, Iowa 52557, USA*
[3]*School of Business, International American University, Los Angeles, CA 90010, USA*
[4]*College of Business, Westcliff University, Irvine, CA 92614, USA*

Corresponding Author: Tofayel Gonee Manik, **E-mail**: m.manik.407@westcliff.edu

| ABSTRACT

Genomic selection is a revolutionary approach in breeding, exploiting genetic markers to forecast breeding values and hence accelerating the pace of traits associated with resilience, like drought tolerance, heat resistance, and pest resistance. This study addresses these challenges through ML algorithms such as random forests, support vector machines, and neural networks thereby enhancing prediction accuracy while handling complicated genomic as well as environmental datasets. Relevant ML algorithms for genomic selection are considered in this discussion, as well as strategies for data processing, feature selection, and environmental factors, including climate conditions and soil parameters. These are brought together to form predictive models that indeed cater to genotype-by-environment interactions vital for crop performance evaluation over different environmental conditions. A proposed framework integrates genomic selection with machine learning, benefiting both disciplines by developing a data-driven methodology for yield prediction in corn. The critical machine learning models to be used include multi-layer perceptron and ensemble models. A case study shows the practical applicability of the GS-ML framework, describing the dataset prepared, model testing and validation procedures, and yield resilience prediction results. The conclusion of the study states that GS and ML combined hold great promise in supporting sustainable agriculture and climate resilience. It requires further research, infrastructure development, and policy support to scale this approach across different crops and diverse climate scenarios. The combined use of genomic and ML approaches is profoundly innovative in predictive breeding and will help develop resilient agricultural systems critical for global food security under a changing climate.

| KEYWORDS

Corn (*Zea mays* L.), Genomic Selection, Machine Learning, Yield Resilience

| ARTICLE INFORMATION

ACCEPTED: 02 June 2023 **PUBLISHED:** 29 July 2023 **DOI:** 10.32996/jeas.2023.4.2.6

1. Introduction

Corn (*Zea mays* L.) is one of the most important food items in the world. Additionally, to being a major source of food, it is grown around the world to make biofuel and animal feed. It's an important crop that keeps many farming economies going around the world and makes sure that millions of people always have food. Rising temperatures, changes in rainfall patterns, and more extreme weather events are all signs of climate change, which is a complicated problem that threatens corn production in a big way (Wang et al., 2018). If the weather changes, it will test corn's ability to resist and adjust. This will have a direct effect on yield.

Climate change makes it harder for corn to grow because it changes weather stress, the way insects move, and the amount of water that is available. For example, drought stress hurts corn during all of its stages of growth, from seedlings to full development. This problem is expected to get worse in the future (Lobell et al., 2014). When the earth is dry from a drought, plant height, leaf area, and photosynthetic efficiency all go down, which lowers yields. A study says that droughts caused by climate change could cut maize crops by 10–20% over the next few decades. This trend is also seen in places like North America, Sub-Saharan Africa, and Southeast Asia where corn is a major crop. When corn is in its reproductive stage, heat stress makes the problem even worse. As temperatures rise above perfect levels, pollen viability and fertilization rates drop sharply, leading to fewer kernels and less grain filling (Hatfield & Prueger, 2015). Also, because of higher temperatures, diseases and pests may be able to move to new places more easily. Rising numbers of diseases and bugs, such as the European corn borer and Fusarium spp., make maize production even less stable (Deutsch et al., 2018). These changes make it clear how important it is to come up with breeding plans that make corn more resistant to these changing natural stresses. Because of the need for climate-resilient farming, new ways of breeding crops are becoming more common. Genomic selection (GS) is one of the most important tools for this. In genomic selection, a type of marker-assisted selection, high-density genetic markers found throughout the whole genome are used to predict how useful an organism will be for breeding (Meuwissen et al., 2001). This work demonstrates the potential of combined GS and ML to enhance the stability and versatility of maize yields. The proposed approach is an integration of GS with ML algorithms designed to tackle high-dimensional, non-linear data structures. Consequently, it can efficiently and reliably predict the reaction of different maize varieties under heat and drought conditions. We present a methodology that integrates machine learning with genomic selection. The merger of machine learning with genetic information improves not only accuracy in predictions but also flexibility, thereby permitting the development of more dependable traits. Therefore, these two technologies need to be integrated for producing climate-resilient maize: it will be an important breeding tool as it quickens the selection process and makes agricultural output more resistant to climate change.

Furthermore, this combination approach has several long-term effects on agricultural output due to climate change. Adaptive breeding is achieved by the combination of genetic selection and machine learning. Breeders can develop crops with greater tolerance under a wide range of stresses by considering the complex interplay between genetic makeup and environmental conditions. This resource efficiency, coupled with waste minimization in production, maximizes food security while simultaneously minimizing the impact on the environment. Data-driven predictive breeding technologies will be critical in transitioning to combat climate change's increasing impacts and in supporting the sustainability of global food systems. The combination of GS and ML allows for data-driven approaches to be integrated with sustainable agriculture strategies used in predictive breeding for increased resistance to climate change in agriculture (Crossa et al., 2017). No less than the preceding paragraph, this one is also in urgent need of increased funding for research and development. As promising as the results from integrating GS with ML are, prediction models require further development because these models need constant updating to reflect changes in climate components over time. Research should continue so that better and more intuitive models can be developed, and the methodologies applied elsewhere for other crops under different regions. This will ensure that breeding programs anywhere in the world have access to state-of-the-art tools. Data-driven agriculture requires government support, multi-disciplinary collaboration, and investments in digital infrastructure. Building on the success of GS with ML integration, focused farming can increase agricultural productivity while also making crops more resilient against unexpected challenges of changing climates through adaptation. If that happens, future agricultural practices could be sustainable rather than injurious to the ecosystem.

2.0 Literature Review and Research Gap

Genomic selection and machine learning hold great promise in the context of climate-change adaptation in crop improvement. Genomic selection has revolutionized plant breeding by estimating the genetic potential for crop traits based on dense molecular markers throughout the genome (Meuwissen et al., 2001). GS has greatly accelerated maize breeding for drought tolerance, insect resistance, and nutrient use efficiency (Xu et al., 2020). GS uses genetic markers to reduce breeding cycles and select desired traits early (Crossa et al., 2017). GS and machine learning thrive at processing complex agricultural datasets. In agricultural genomics, where environmental and phenotypic variables interact with genomic information in complex ways, ML algorithms can find non-linear patterns (Wang et al., 2018). Random forests, support vector machines, and deep neural networks improve genomic selection predictions because they handle diverse data types and high-dimensional genomic data (Montesinos-López et al., 2018). ML applications increase genomic model prediction to optimize stress tolerance, yield stability, and other traits under changing environmental conditions (Jin et al., 2019). Recent studies illustrate the potential of GS and ML in enhancing climate resilience in agriculture. According to Heslot et al. (2015), the use of genetic predictions coupled with environmental data improves the forecast of wheat production across different conditions. On the other hand, ML algorithms are applicable in maize genomic predictions for abiotic stress tolerance and thus a hybrid GS-ML model is proposed for dynamic climatic solutions. Breeding of maize for climate resilience by using GS and ML has been nearly impossible recently though significant progress has been made. Due to the complexity of environmental interaction, more precise predictions from models require phenotypic data along with environmental and multi-omics data (Crossa et al., 2019). Data availability is another challenge, particularly the high-quality environmental and phenotypic data that span many seasons required to train robust predictive models (Liu et al., 2020). Model interpretability and

ML algorithm processing may also limit breeding program use (Spindel & McCouch, 2016). These challenges may be solved by ML in predictive agriculture. Deep learning and reinforcement learning improve model predictions by integrating diverse inputs (Krizhevsky et al., 2017). Therefore, hybrid GS-ML methods may improve breeding program resilience, efficiency, and climatic adaptation.

However, Genomic selection and machine learning improve climatic stress resistance, but corn breeding has yet to incorporate them. Most research on GS-ML integration has used theoretical models or small-scale trials, failing to depict maize's global environmental variability (Xu et al., 2021). Research by Crossa et al. (2017) highlights the need for models that more adequately capture genotype-by-environment (G×E) interactions, which are essential for the development of climate-resilient varieties but difficult to monitor in data that are very complex. Another problem is that the genetic, phenotypic, and environmental datasets that are accessible or of good quality are inadequate. Few studies have combined transcriptomics, proteomics, climate, and soil data to develop a comprehensive breeding model. Scalable computational frameworks are often overlooked in the literature, limiting prolonged breeding initiatives that may lack advanced machine learning infrastructure (Spindel & McCouch, 2016). Interpretable genomic forecasting machine learning models are another unexplored area. Deep learning models are accurate, but their complexity may obscure gene-phenotype relationships, which are crucial for breeding decisions (Montesinos-López et al., 2018). A comprehensive GS-ML model for corn breeding is created in this study to address these issues. This study integrates genomic selection with machine learning algorithms that can manage genotype-environment interactions and various data inputs to anticipate crop resilience under climate change more accurately, robustly, and interpretably. This system uses multi-omics data and environmental variables to help corn breeding programs generate climate-resilient crops.

3.0 Genomic Selection in Corn Breeding

3.1 Principles and Process of Genomic Selection

Genomic selection (GS) is an advanced approach in crop breeding that determines the breeding value of an individual for a particular trait using high-density genetic markers spread throughout the genome (Meuwissen et al., 2001). Unlike traditional marker-assisted selection, which focuses on a limited set of specific markers that are associated with traits considered desirable, GS employs all available markers and thus allows for a more accurate selection within complex phenotypes governed by many genes (Crossa et al., 2017). The initial phase of the GS process involves identifying a training population that has undergone genotyping and phenotyping (Figure 1). Models constructed from this data are utilized to ascertain genomic estimated breeding values (GEBVs), which indicate an individual's genetic potential for favorable traits (Goddard & Hayes, 2007). These GEBVs considerably accelerate the breeding cycle, enabling breeders to make decisions prior to the complete development of phenotypic expression. Selection considerations in GS include the precision of GEBV projections and the genetic advancement per breeding cycle. By choosing individuals with the highest GEBVs for resilience traits, breeders can expedite the creation of crop varieties that exhibit enhanced resistance to climate stressors such as heat and drought (Jannink et al., 2010). Genomic selection (GS) is an effective method for enhancing complex traits as it encompasses both minor and significant genetic influences, especially for traits exhibiting low heritability (Hatfield and Prueger 2015).

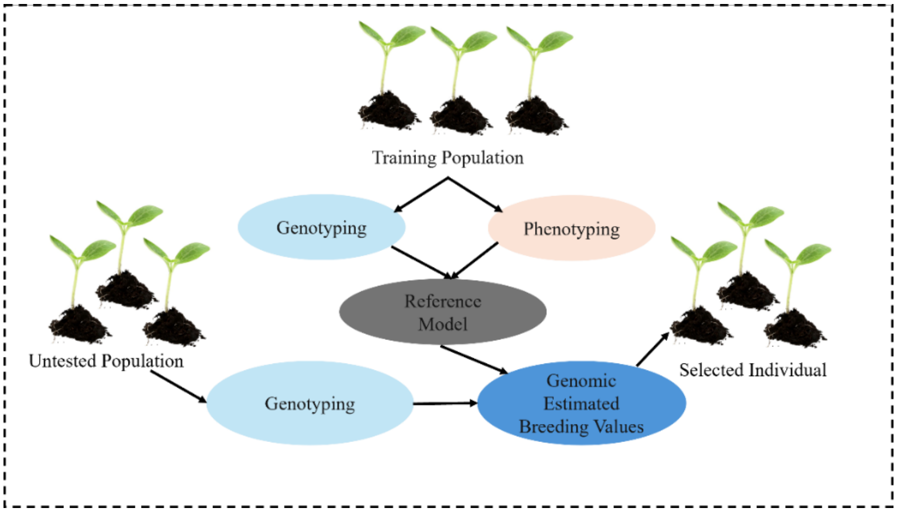

Figure 1. Schematic diagram of a pipeline for genomic selection.

3.2 Genomic Data in Corn and Relevant Traits for Resilience

Important resilience traits in maize include resistance to heat, dryness, and insects. Drought resistance is essential since climate change is expected to make the water scarcity worse. Cooper et al. (2014) state that the characteristics of drought-resistant maize varieties are controlled by a large number of genes. These characteristics include lower transpiration rates, deeper root systems, and improved water-use efficiency. Being able to tolerate high temperatures is essential, particularly during the reproductive stage, as too much heat during blooming may result in fewer and smaller kernels. Genomic studies on maize have shown that genes responsive to heat influence both the quantity of viable pollen grains and the plant's capacity to tolerate heat stress (Lobell et al., 2014). Another important resilience trait is pest resistance. Given that insect ranges are shifting as a result of climate change, maize types must be genetically resistant to a wider range of pests, such as the European corn borer and fall armyworm (Deutsch et al., 2018). Genomic research has identified QTLs for pest resistance; GS models can now incorporate such markers for more effective breeding of pest-resistant crops (Nelson et al., 2018). With the identification and use of such genetic markers, GS helps in the selection of maize varieties with a proper blend of resistance traits. These traits enable the maize to resist pests and environmental stresses.

3.3 Challenges in Traditional Genomic Selection

There are disadvantages to GS despite its many potential benefits. One of the biggest issues is managing the large and complex datasets required for accurate genetic predictions. The high-throughput genotyping techniques used in GS generate vast amounts of data, which must be efficiently handled and understood. Such large datasets are computationally expensive and also require complex statistical methods for effective analysis (Crossa et al., 2017). One of the crucial issues in GS is environmental interactions, or G×E interactions. Environmental factors significantly influence maize performance, and a relationship between these factors and genetic markers may obscure the relationship identified between those genetic markers and observable traits (Jarquín et al., 2014). The ability of a genotype selected for drought resistance to behave differently in different environmental conditions hinds the accuracy of genomic predictions. For this purpose, complex models are required that incorporate environmental parameters like weather and soil conditions to account for G×E interactions and strengthen the reliability of GS models (López-Cruz et al., 2015). Another major barrier is the heritability of complex factors, which may be difficult to accurately evaluate with GS alone. Since many resilience traits are influenced by multiple genes with minimal individual effects, it is more challenging to ascertain the full genetic contribution to these traits (Hatfield and Prueger 2015). While GS offers advantages over traditional breeding, its effectiveness for complex traits depends on the quality of the genetic and environmental data used, as well as the accuracy of the prediction models.

3.4 Machine Learning Techniques for Predictive Breeding

A lot of complex genetic data is used by ML systems to guess what breeding features will be. This is why they are very important for predictive breeding. Machine learning methods such as NN, SVM, and RF are used for genome screening. Firstly, there are Random Forests (RF), which use more than one decision tree to make better guesses. It works well with genetic data because it can handle large amounts of data without fitting too well (Breiman, 2001). When it comes to RF, the "forest" of decision trees can guess difficult features and find links that don't go in a straight line. Heslot et al. say that a number of genes can change how it makes predictions. Secondly, Getting help to learn SVM finds the best hyperplane to sort the data. SVM effectively separates genetic markers linked to good traits, such as being able to handle disease or drought (Cristianini & Shawe-Taylor, 2000). With completed missing values, appropriate feature selection, and normalized data, predictions are much more accurate (Daetwyler et al., 2013). Thirdly, Neural Networks (NN): Deep learning models can show that genetic marker links are complicated and not linear, which is useful for genomic selection (LeCun et al., 2015). Recurrent neural networks and convolutional neural networks can both work with genetic patterns and time series external data. NNs can show how genomic markers and environmental factors work together in a complicated way to improve breeding predictions (Montesinos-López et al., 2018). Genomics selection is helped by machine learning methods that try to guess how traits will change in future generations when genes and the environment interact in complicated ways.

4.0 Integrating Genomic Selection and Machine Learning

Combining genomic selection (GS) and machine learning (ML) can forecast crop output and resilience, especially in maize. The GS-ML approach uses genomic data from GS and powerful prediction algorithms in ML to generate robust models that account for complicated genetic-environmental interactions (Crossa et al., 2017). This approach uses ML algorithms to predict genomic estimated breeding values (GEBVs) from a genotyped and phenotyped training population (Goddard & Hayes, 2007). A practical GS-ML approach collects high-density genetic data and environmental factors like soil quality and climate, then preprocesses them to eliminate noise and fill in missing values. ML methods, which are ideal for genomics, train the model on non-linear and high-dimensional data. Ensemble methods like random forests or neural network topologies like multi-layer perceptrons can manage breeding's many markers and complex trait interactions, making them ideal for GS (Montesinos-López et al., 2018). This integrated approach captures genetic and environmental influences on crop performance to improve yield estimates (Jin et al., 2019).

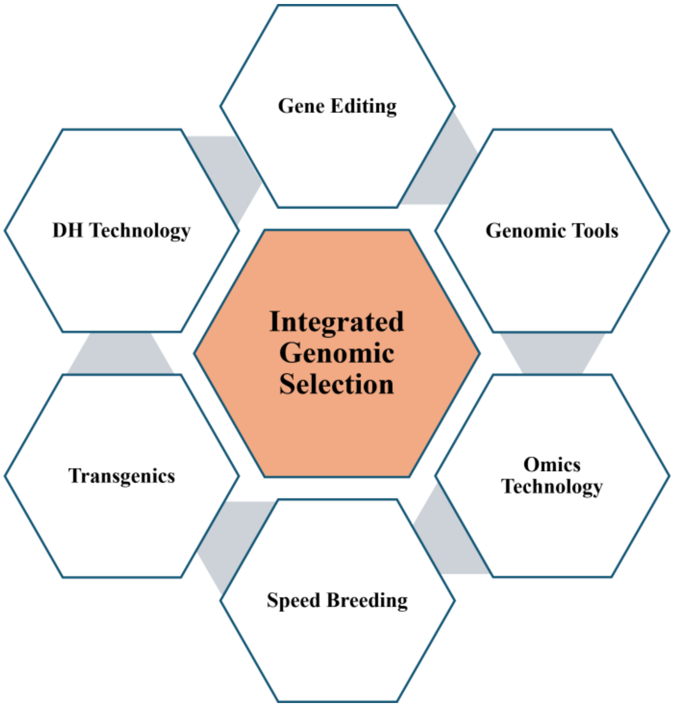

Figure 2. Integration of modern genomic tools with genomic selection

Genomic selection with machine learning, especially for complicated corn breeding traits, has many benefits (Figure 2). First, ML algorithms that manage non-linear relationships and interactions between markers increase prediction accuracy for multi-gene traits like yield and stress tolerance (Heslot et al., 2012). This accuracy enhances both the selection and genetic gain, which is crucial for breeding programs that are time sensitive. Another benefit is that ML algorithms are flexible to new data. In fact, ML models can be updated with genotypic, phenotypic, and environmental data, thus becoming more adaptable. This is an important requirement in corn breeding as the pressures from drought, temperature, and insects vary every year (Wang et al., 2018). The size and complexity of genomic selection datasets go hand in hand; ML systems excel at managing such datasets. Indeed, traditional analytical approaches fail with genomic datasets containing hundreds of markers and environmental factors. High-dimensional data structures that support application robustness in prediction are efficiently managed by support vector machines and neural networks (Gianola et al., 2006). Finally, ML allows multi-trait and multi-environment analysis. Bayesian models and ensemble techniques enable breeders to consider genotype-by-environment (G×E) interactions, crucial for forecasting corn performance in varied contexts (Jarquín et al., 2014). This allows the selection of high-yielding, environmentally resilient corn cultivars. ML models like MLPs and ensemble models show promise in genomic selection for predictive breeding. Multi-Layer Perceptrons (MLPs), MLPs are artificial neural networks with numerous layers of neurons that can capture complex, non-linear genetic data correlations. MLPs can simulate complex marker interactions and learn from big datasets, making them useful for GS. MLPs are suitable for multi-gene maize traits like yield and drought tolerance because they improve complex trait prediction accuracy (LeCun et al., 2015). MLPs' hierarchical nature makes them successful for high-dimensional genomic data but requires a lot of training data and computer capacity (Montesinos-López et al., 2018). In Ensemble Models (Random Forests, Gradient Boosting), Ensemble models improve prediction and reduce overfitting by using different methods. Random forests (RF) combine decision tree predictions to improve model stability and accuracy. RF can forecast corn production in GS using many genetic markers, even with complicated, noisy data (Breiman, 2001). Gradient boosting models, which generate sequential decision trees that correct each other's errors, improve genomic selection accuracy with less training data (Heslot et al., 2015). In Bayesian Neural Networks, Useful for genomic selection in corn, where G×E interactions are significant, as they incorporate uncertainty into predictions. These models are useful for yield prediction in different situations with changeable data. Bayesian networks can improve breeding decisions by providing confidence ranges for predictions by integrating probabilistic frameworks (Crossa et al., 2017). These models are implemented using a conventional pipeline. Raw genetic and environmental data are normalized, and feature selected to minimize dimensionality and focus on key markers (Poland & Rife, 2012). Data is separated into training and testing sets. To improve accuracy, training involves parameter adjustment and model selection via cross-validation. To verify robustness before applying the model to new breeding populations, independent datasets or multi-environment trials test its predictions (Goddard & Hayes, 2007). These adaptable, scalable ML models allow maize breeders to accurately forecast yield resilience and other complex traits under climate change conditions by merging genomic selection with machine learning.

5.0 Challenges and Future Directions

ML breakthroughs like deep learning and reinforcement learning can assess complex and high-dimensional data, changing genetic selection. CNNs and RNNs may model complex trait interactions using non-linear genetic links (LeCun et al., 2015). CNNs excel in image-based phenotyping, which quantifies yield-related plant traits, while RNNs excel at time-series data, such as climate patterns affecting crop development (Montesinos-López et al., 2018). In ML's reinforcement learning (RL) area, agents learn optimal actions by interacting with an environment, adaptive breeding tactics are fascinating. RL can model breeding conditions and develop long-term genetic gain strategies (Kaelbling et al., 1996). RL could assist breeders evaluate and optimize breeding methods in simulated environmental and climate conditions for genomic selection in uncertain climates (Wang et al., 2018). These approaches should increase forecast accuracy, manage complex trait interactions, and shorten breeding cycles for climate-resilient breeding (Jin et al., 2019). Genomic selection and ML integration worked in maize, but they could work in climate-vulnerable crops like wheat, rice, and soybeans. ML models can find heat tolerance and yield stability genes in wheat, which is sensitive to temperature increases throughout critical growth stages (Lobell et al., 2015). Rice and soybeans, susceptible to drought and insect stress, can benefit from ML-assisted genomic selection for resilience (Jarquín et al., 2014).

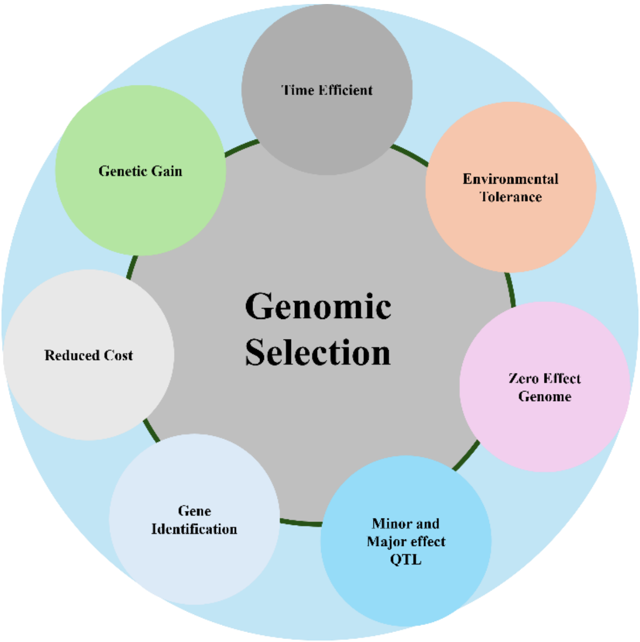

Figure 3. Application and future direction of genomic selection

Moreover, ML and genomic selection frameworks must adjust to different climate scenarios beyond crop-specific adaptations as climate change affects global agricultural zones. Breeders can employ MET and ML to predict G×E interactions in diverse climates, providing consistent selection under changing environmental stresses (Crossa et al., 2017). By adding regional climate forecasts to ML models, breeding programs can simulate how various varieties will perform under future conditions like drought frequency or higher temperatures, customizing selection for resilience to specific climate risks (Xu et al., 2020). Genomic selection and ML in agriculture require strong policy support for data sharing, infrastructure, and collaboration. Big genetic, phenotypic, and environmental data drives data-driven agriculture. Data exchange is hindered by privacy, proprietary rights, and lack of procedures. Policies must promote data standardization, privacy, and ethics to share data across institutions and regions (Poland & Rife, 2012). Also crucial is digital infrastructure investment. Breeding operations, especially in poor countries, may lack strong computers and data storage for innovative ML and genomic selection. Cloud and data storage infrastructure funding could democratize these technologies (Spindel & McCouch, 2016). Integration of ML-genomic selection requires research collaboration. Cross-disciplinary plant, genetic, data, and policymaker collaborations yield unique agricultural solutions. Research collaborations and policy-backed multinational partnerships can accelerate climate-resilient crop development (Goddard & Hayes, 2007). Governments and organizations should encourage private sector data-driven agriculture to promote and create sustainable agriculture (Jin et al., 2019).

6.0 Conclusion

Our integrated approach of GS and ML successfully addresses the non-linear and multidimensional data problem by accurately predicting the response of different maize varieties to environmental stresses such as heat and drought. This will not only allow for stronger attributes in prediction accuracy but also versatility with the genetic data machine learning will help breed. Comforting climate-proof maize varieties are thus a possibility while breeders may benefit from a faster selection process along with better consistency in output under varying conditions. The agricultural implications of this combined approach extend far into the future, especially in the context of climate change. Through the synergistic relationship between genetic selection and machine learning, flexible breeding practices have emerged. Never before has it been possible to develop crops with such precise environmental tolerance, as breeders now consider the intricate interplay between genes and their surroundings. This accuracy in crop production optimizes the use of resources, reduces excessiveness, thus reinforcing food security while having a lesser impact on the environment. In the face of data-driven predictive breeding technologies becoming increasingly vital for sustainably ensuring global food systems, climate change is exerting intense pressures. The marriage of GS and ML has paved the way for a new era in predictive breeding, where data-driven technologies and sustainable agriculture converge to boost crop resilience against climate change. Furthermore, the research and development funding in this field is crucial. The integration of GS and ML has demonstrated promising outcomes, yet prediction models must keep evolving to maintain their relevance and accuracy, given the dynamic nature of climate factors. Not only do these methodologies need to grow to encompass other crops and regions, but simpler, more user-friendly models must also emerge. By doing so, breeding operations across the globe can benefit from these enhanced tools. Success in data-driven agriculture hinges on governmental support, interdisciplinary cooperation, and investments in digital infrastructure. Just as GS and ML integration has proven successful, farmers who pour their efforts stand to gain improved agricultural productivity and enhanced crop resilience against climate change's unexpected challenges. If this momentum continues, we may witness more environmentally sustainable agricultural practices in the future.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Acknowledgement: We would like to express our gratitude to all the co-authors for their contribution and critical reviews from the anonymous reviewers.

ORCID ID:
Abu Saleh Muhammad Saimon: https://orcid.org/0009-0006-3147-1755
Mohammad Moniruzzaman: https://orcid.org/0009-0006-5981-4473
Md Shafiqul Islam: https://orcid.org/0009-0008-9067-4987
Md Kamal Ahmed: https://orcid.org/0009-0004-1003-6207
Md Mizanur Rahaman: https://orcid.org/0009-0004-5915-7114
Sazzat Hossain: https://orcid.org/0009-0008-6325-5496
Mia Md Tofayel Gonee Manik: https://orcid.org/0009-0005-6098-5213

References

[1] Cooper, M., Messina, C. D., Podlich, D., Totir, L. R., Baumgarten, A., Hausmann, N. J., ... & Chapman, S. C. (2014). Predicting the productivity of maize in drought environments: An integrated physiological and genetic approach. Agronomy Journal, 106(2), 455–473.
[2] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.
[3] Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press.
[4] Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., ... & Burgueño, J. (2017). Genomic selection in plant breeding: methods, models, and perspectives. Trends in Plant Science, 22(11), 961–975.
[5] Daetwyler, H. D., Calus, M. P., Pong-Wong, R., de los Campos, G., & Hickey, J. M. (2013). Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. Genetics, 193(2), 347–365.
[6] Deutsch, C. A., Tewksbury, J. J., Tigchelaar, M., Battisti, D. S., Merrill, S. C., Huey, R. B., & Naylor, R. L. (2018). Increase in crop losses to insect pests in a warming climate. Science, 361(6405), 916–919.
[7] Gianola, D., Fernando, R. L., & Stella, A. (2006). Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics, 173(3), 1761–1776.
[8] Goddard, M. E., & Hayes, B. J. (2007). Genomic selection. Journal of Animal Breeding and Genetics, 124(6), 323–330.
[9] Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: Effect on plant growth and development. Weather and Climate Extremes, 10, 4–10.
[10] Heslot, N., Akdemir, D., Sorrells, M. E., & Jannink, J. L. (2015). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theoretical and Applied Genetics, 127(2), 463–480.
[11] Jannink, J. L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: From theory to practice. Briefings in Functional Genomics, 9(2), 166–177.
[12] Jarquín, D., Crossa, J., Lacaze, X., Cheyron, P. D., Daucourt, J., Lorgeou, J., ... & Rincent, R. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theoretical and Applied Genetics, 127(3), 595–607.
[13] Jin, Z., Lee, B. H., & Chang, H. Y. (2019). Application of machine learning in precision agriculture. Agronomy, 9(10), 568.
[14] Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. Journal of Artificial Intelligence Research, 4, 237–285.
[15] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90.
[16] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
[17] Liu, X., Huang, M., Fan, X., & Wang, J. (2020). Genomic selection in crops: from research to practice. Frontiers in Plant Science, 11, 1356.

[18] Lobell, D. B., Hammer, G. L., McLean, G., Messina, C., Roberts, M. J., & Schlenker, W. (2014). The critical role of extreme heat for maize production in the United States. Nature Climate Change, 4(7), 615–619.

[19] Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2014). Climate trends and global crop production since 1980. Science, 333(6042), 616–620.

[20] López-Cruz, M. A., de los Campos, G., Pérez-Rodríguez, P., & Gianola, D. (2015). Predicting complex traits in plant breeding: An assessment of deep learning and Bayesian regression models. Plant Science, 10(1), 1–9.

[21] Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157(4), 1819–1829.

[22] Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Toledo, F. H., Pérez-Hernández, O., Eskridge, K. M., & Rutkoski, J. (2018). A genomic Bayesian multi-trait and multi-environment model. G3: Genes, Genomes, Genetics, 8(12), 3875–3888.

[23] Nelson, R., Wiesner-Hanks, T., Wisser, R., & Balint-Kurti, P. (2018). Navigating complexity to breed disease-resistant crops. Nature Reviews Genetics, 19(1), 21–33.

[24] Poland, J., & Rife, T. (2012). Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome, 5(3), 92–102.

[25] Spindel, J. E., & McCouch, S. R. (2016). When more is better: how data sharing would accelerate genomic selection of crop plants. New Phytologist, 212(4), 814–826.

[26] Wang, Y., Wang, X., Singh, P., Xu, Y., & Zhao, Y. (2018). Precision agriculture with advanced big data: The future of agronomy. Agronomy Journal, 110(2), 1778–1796.

[27] Xu, Y., Xu, C., & Xu, S. (2020). Prediction and association analysis of multi-environmental trials. The Plant Journal, 103(4), 1195–1209.