

Design and Implementation of a Machine Learning Simulation System for Crop Yield Prediction Under Climate Variability

Azeez Rahman Diekola and *Wakili Muazu Umar

¹Department of Computer Science, Faculty of Computing, Engineering and Technology, Al-Hikmah University, Ilorin, Kwara State, Nigeria.

*Corresponding authors' email: umarwakili44@gmail.com

ORCID Id: <https://orcid.org/0009-0006-2540-429X>

ABSTRACT

Climate variability continues to threaten agricultural productivity, especially in regions where farming depends heavily on rainfall and temperature conditions. This study presents the design and implementation of a web-based computational simulation system for predicting crop yield under climate variability. The system uses rainfall, temperature, pesticide usage, crop type, country, and land size as input variables to estimate crop yield and total production. Historical crop yield and climate data were processed, filtered, and used to train crop-specific machine learning models. Feature engineering was applied to generate additional predictors, including temperature-rainfall interaction, pesticide intensity, growing degree days, heat stress, cold stress, rainfall categories, logarithmic variables, and squared variables. Random Forest, Gradient Boosting, Ridge Regression, and ensemble models were trained and evaluated. The system was implemented using a Fast API backend, a web-based frontend, interactive visualization components, and an AI-based recommendation module that explains completed prediction results. The results showed strong predictive performance for several crops, with rice achieving the highest test R^2 value of 0.9617, followed by maize with 0.9080, soybeans with 0.8524, wheat with 0.8157, and potatoes with 0.7841. The study demonstrates that machine learning-based simulation can support crop yield prediction and agricultural decision-making under changing climate conditions, especially in data-limited regions.

Received: 06 May 2026

Accepted: 19 Oct. 2026

Published: 20 Oct. 2026

Keywords:

Agricultural simulation, Climate variability, Crop yield prediction, Decision support system, Machine learning

INTRODUCTION

Agriculture is essential for food security and poverty reduction in developing regions (Pawlak & Kołodziejczak, 2020), requiring scalable, adaptive farming approaches under growing climate pressures (Ewert, Baatz, & Finger, 2023). Shifting temperatures, rainfall instability, and extreme events increasingly disrupt global food systems (Farah, Mohamed, Musse, & Nor, 2025; Saleem, Anwar, Nawaz, Fahad, & Saud, 2025). This environmental variability heightens production uncertainty, while risk models often underestimate synchronized low-yield events (Kornhuber et al., 2023). Consequently, accurate yield forecasting is critical for early planning, food security, and risk mitigation (International Food Policy Research Institute, 2023), especially as unmitigated climate change risks sudden extreme shocks (Hultgren et al., 2025; Kornhuber et al., 2023). Traditional process-based crop simulation models assess climate impacts (Timlin, Paff, & Han, 2024) and support management decisions (Banerjee, Dutta, Das, & Sadhukhan, 2025). However, their reliance on extensive soil, cultivar, and management calibration limits deployment in data-scarce areas. Machine learning (ML) offers a flexible alternative by learning complex nonlinear climate-yield patterns directly from data without strict mechanistic assumptions (Hernández, Gómez, & Jiménez-Cabas, 2025), with interpretable models enhancing decision clarity (Hu, Zhang, Bohrer, Liu, & Zhou, 2023). Furthermore, because model performance varies significantly across crop types (Seck, Ngom, Ngom, & Noba, 2025) and predictive tools still struggle in data-scarce regions (Betew et al., 2026), crop-

specific ML modeling is essential.

To address these challenges, this study designs and implements a web-based computational simulation system for predicting crop yield under climate variability. Using historical agricultural and climate data, engineered features, crop-specific ML models, and an AI recommendation module, the system estimates yield per hectare and total production. The specific objectives are to evaluate climate-yield relationships, develop and evaluate crop-specific ML models, and deliver an accessible decision-support framework tailored for data-limited environments

Related Work

Process-Based Crop Simulation Models

Process-based models simulate crop growth using biological and environmental mechanisms. Studies demonstrate their utility in evaluating crop phenology, yield response, and water demand under climate change (Yang et al., 2014), assessing rainfall variability adaptations (Barati, Soundharajan, & Nikoo, 2024), projecting irrigation requirements (Koukoulis, Georgiou, & Karpouzios, 2025), and evaluating conservation practices (Yadav, Krishnan, Parihar, & Banerjee, 2025). While valuable for climate impact assessments (Timlin, Paff, & Han, 2024) and regional yield projections (Chisanga, Phiri, & Chinene, 2017), their heavy calibration requirements limit practical decision-support deployment in data-scarce settings (Banerjee, Dutta, Das, & Sadhukhan, 2025) and hinder extreme weather simulation (Zare, Azam, & Sauchyn, 2023).

Machine Learning Approaches

Machine learning (ML) derives empirical relationships directly from historical data. Comprehensive reviews highlight that AI-driven predictive modelling is increasingly moving toward multi-variable data integration and complex algorithms to improve yield estimation accuracy (Hernández Hernández, Gómez Gómez, & Jiménez-Cabas, 2025). IoT-enabled ML frameworks achieve high predictive performance for precision agriculture (Talaat, 2023), though explainability remains critical to avoid black-box risks in climate evaluations (Hu, Zhang, Bohrer, Liu, & Zhou, 2023). Ensemble models consistently capture nonlinear climate-yield dynamics (Hortizuela & Palaoag, 2025; Islam et al., 2024). Predictive performance varies significantly by crop type and regional dataset quality; evaluating multiple machine learning models across groundnut, millet, and cotton yields in West Africa demonstrates variable predictive performance across distinct crop types (Seck, Ngom, Ngom, & Noba, 2025), while algorithms like Naïve Bayes and Random Forest show strong general capability (Maazallahi et al., 2024; Tasneem et al., 2025). Other deployments include mobile prediction tools (Venugopal et al., 2021), hybrid ML/deep learning methods (Agarwal & Tarar, 2021), reusable baseline forecasting workflows (Paudel et al., 2021), and semiparametric neural networks (Crane-Droesch, 2018). Overall, contemporary ML methods outperform older models, yet data scarcity and regional generalization remain persistent challenges (Betew et al., 2026).

Hybrid Frameworks and System Integration

Hybrid models combine mechanistic structures with ML to optimize predictions. Pretraining ML models on synthetic

crop data improves efficiency (Kallenberg et al., 2023), while integrating models like WOFOST, DSSAT, and AquaCrop with Random Forest, XGBoost, or Gaussian Processes refines phenology estimation, climate driver identification, and drought risk mapping (Alimagham et al., 2025; El-Mahroug et al., 2025; Baluch et al., 2025). To address climate-induced variability, Bayesian hierarchical ensemble frameworks account for both climate scenarios and technological impacts to generate probabilistic yield projections (Li, Kitsios, Newth, & O’Kane, 2025). Furthermore, advancing multimodal remote sensing applications, vision transformer architectures like CMAViT effectively fuse climate, field management, and satellite data for high-accuracy yield estimation (Kamangir, Sams, Dokoozlian, Sanchez, & Earles, 2024). This study bridges this gap by delivering a web-based simulation platform that pairs crop-specific ML models and engineered climate features with an advisory AI module that explains prediction outputs without altering numerical results

MATERIALS AND METHODS

This study adopted a computational design to process historical agricultural data, engineer features, train crop-specific regression models, and deploy a web-based simulation platform. The overall framework encompasses system design, data preparation, feature engineering, model development, and implementation.

System Design

The modular platform collects crop type, country, and land size from the user to query climate values, compute engineered features, execute saved models, and present yield predictions alongside production metrics.

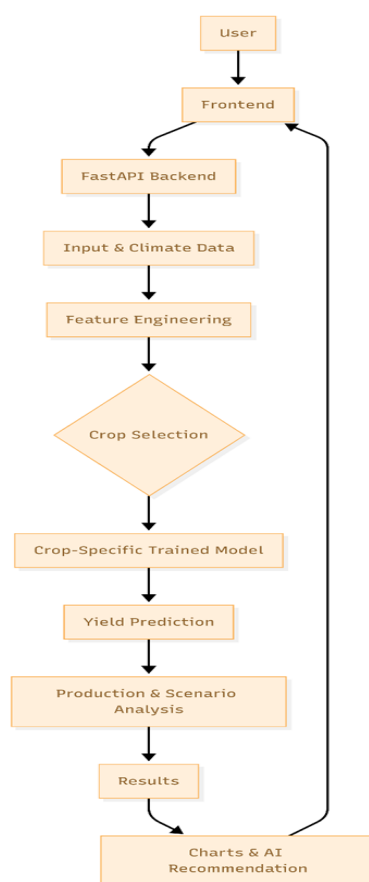


Figure 1: System Architecture of the Crop Yield Prediction Simulation System

The system architecture consists of five core components. The input component captures target crop selection, country, and land area to calculate total expected production. The climate data component stores historical rainfall and temperature values for 10 African countries, such as Ghana, Kenya, and South Africa. The feature engineering component transforms raw input variables into derived agronomic indicators. The prediction component loads crop-specific models to process features, forecast yield per hectare, and scale metrics by land size. Finally, the output component renders yield metrics, comparative dynamic charts, and natural language AI explanations.

Data Source and Sample Size

The dataset contains historical agricultural records spanning 1990–2013. The original dataset of 28,242 records across 101 countries was filtered to focus on 12 African nations, yielding 2,776 localized records. Target variables were recorded in hectograms per hectare (hg/ha).

Feature Engineering

To capture complex non-linear climate dynamics, additional features were constructed. Climate interactions and extreme thresholds were captured using temperature-rainfall interaction, growing degree days, heat stress, and cold stress indicators. Agronomic management and non-linear terms were incorporated via pesticide intensity, discrete rainfall categories, squared variables, and logarithmic transformations. This multi-modal feature expansion aligns with literature advocating remote sensing indicators and comprehensive environmental inputs for predictive stability (Kira et al., 2024; Pathak et al., 2023).

Model Development

Crop-specific regression models including Random Forest, Gradient Boosting, Ridge Regression, and Ensembles were trained for maize, rice, sorghum, cassava, wheat, potatoes, and soybeans. Models were selected based on individual evaluation performance using a temporal validation split, using earlier years for training and later years for testing. Model performance was assessed using R^2 , RMSE, MAE, and MAPE evaluation metrics. Baseline, best-case, and

worst-case scenario outputs were calculated to define uncertainty ranges.

System Implementation

The backend application was developed in Python using Fast API for API endpoint handling, with Joblib applied for model serialization and NumPy used for numerical array operations. The user interface and frontend were built using HTML, CSS, JavaScript, and ApexCharts to render real-time dynamic visualizations. Additionally, the AI recommendation layer integrates the Gemini API strictly as an advisory component to explain model outputs without modifying numerical predictions.

RESULTS AND DISCUSSION

This section presents the results obtained from the development, training, evaluation, and implementation of the crop yield prediction system. The results are organized into three main parts. The first part presents the evaluation performance of the crop-specific machine learning models using standard regression metrics. The second part presents the 2026 climate scenario prediction results, including baseline prediction, best-case prediction, worst-case prediction, and uncertainty range. The third part presents the system output, showing how the trained models were integrated into a web-based simulation interface for user interaction, visualization, and AI-assisted recommendation.

Model Evaluation Results

Crop-specific models were evaluated using R^2 , RMSE, MAE, and MAPE. Model selection varied by crop: Gradient Boosting was selected for maize, rice, and sorghum; Random Forest for wheat and cassava; and Ensemble modeling for potatoes and soybeans. Rice achieved the highest predictive accuracy ($R^2 = 0.9617$, MAPE = 9.09%), followed by maize ($R^2 = 0.9080$) and soybeans ($R^2 = 0.8524$). Conversely, cassava ($R^2 = 0.6728$) and sorghum ($R^2 = 0.6067$) exhibited moderate performance, reflecting higher relative error or unmodeled local variance.

Table 1: Training and Evaluation Performance of Crop Yield Prediction Models

Crop	Selected Method	Train R^2	Test R^2	RMSE	MAE	MAPE
Maize	Gradient Boosting	0.9999	0.9080	6,602	4,008	20.68%
Rice	Gradient Boosting	0.9999	0.9617	5,209	3,320	9.09%
Wheat	Random Forest	0.9846	0.8157	7,065	5,700	17.55%
Cassava	Random Forest	0.9350	0.6728	31,078	25,254	60.89%
Sorghum	Gradient Boosting	0.9999	0.6067	20,140	4,898	12.81%
Potatoes	Ensemble	0.9710	0.7841	45,158	33,131	26.68%
Soybeans	Ensemble	0.9483	0.8524	3,258	2,733	19.24%

Rice and maize achieved the highest evaluation performance using Gradient Boosting ($R^2 > 0.90$), while soybeans, wheat, and potatoes demonstrated strong predictive power ($R^2 > 0.78$). Cassava and sorghum exhibited moderate accuracy ($R^2 < 0.68$), indicating that these crops may require additional local environmental predictors.

Climate Scenario Prediction Results

The system generated 2026 yield scenario forecasts, detailing baseline, best-case, worst-case, and uncertainty metrics. Cassava produced the highest baseline yield (132,526 hg/ha) but also displayed the widest uncertainty range (68,860 hg/ha). Potatoes yielded the second-highest baseline (78,203 hg/ha), while sorghum recorded the narrowest uncertainty range (1,717 hg/ha), indicating high scenario stability under projected conditions.

Table 2: 2026 Climate Scenario Prediction Results

Crop	2026 Baseline Prediction	Best Case	Worst Case	Uncertainty Range
Maize	16,154 hg/ha	17,297 hg/ha	14,626 hg/ha	2,671 hg/ha
Rice	20,521 hg/ha	22,842 hg/ha	18,606 hg/ha	4,236 hg/ha
Wheat	17,127 hg/ha	22,074 hg/ha	16,948 hg/ha	5,126 hg/ha
Cassava	132,526 hg/ha	140,274 hg/ha	71,413 hg/ha	68,860 hg/ha
Sorghum	11,346 hg/ha	12,034 hg/ha	10,317 hg/ha	1,717 hg/ha
Potatoes	78,203 hg/ha	86,643 hg/ha	69,660 hg/ha	16,984 hg/ha
Soybeans	10,734 hg/ha	13,568 hg/ha	10,416 hg/ha	3,152 hg/ha

Cassava recorded the highest projected baseline yield (132,526 hg/ha) but also the largest uncertainty spread (68,860 hg/ha), showing high sensitivity to scenario variations. In contrast, sorghum yielded the narrowest uncertainty range (1,717 hg/ha), demonstrating the most stable forecast across projected conditions.

Comparative Ranking of Model Performance

Ranking crops by test R^2 performance demonstrates that five of the seven models achieved strong to very strong predictive accuracy ($R^2 > 0.78$), with rice and maize outperforming all other categories.

Table 3: Ranking of Crops by Test R^2 Performance

Rank	Crop	Selected Method	Test R^2	Performance Level
1	Rice	Gradient Boosting	0.9617	Very strong
2	Maize	Gradient Boosting	0.9080	Very strong
3	Soybeans	Ensemble	0.8524	Strong
4	Wheat	Random Forest	0.8157	Strong
5	Potatoes	Ensemble	0.7841	Good
6	Cassava	Random Forest	0.6728	Moderate
7	Sorghum	Gradient Boosting	0.6067	Moderate

Ranking by test R^2 confirms that five out of seven crop models achieved strong to very strong explanatory performance ($R^2 \geq 0.7841$) on unseen data. Rice and maize led overall performance, while cassava and sorghum ranked lowest.

Comparative Ranking of Scenario Uncertainty

Ranking by uncertainty range confirms that root and tuber crops exhibit greater variation between best-case and worst-case scenarios, whereas cereal and legume crops maintain substantially narrower prediction spreads.

Table 4: Ranking of Crops by Uncertainty Range

Rank	Crop	Best Case	Worst Case	Uncertainty Range
1	Cassava	140,274 hg/ha	71,413 hg/ha	68,860 hg/ha
2	Potatoes	86,643 hg/ha	69,660 hg/ha	16,984 hg/ha
3	Wheat	22,074 hg/ha	16,948 hg/ha	5,126 hg/ha
4	Rice	22,842 hg/ha	18,606 hg/ha	4,236 hg/ha
5	Soybeans	13,568 hg/ha	10,416 hg/ha	3,152 hg/ha
6	Maize	17,297 hg/ha	14,626 hg/ha	2,671 hg/ha
7	Sorghum	12,034 hg/ha	10,317 hg/ha	1,717 hg/ha

Uncertainty rankings highlight that root and tuber crops (cassava and potatoes) experience greater variance between best-case and worst-case outcomes. Conversely, cereals and legumes (maize, soybeans, and sorghum) displayed tighter prediction bounds.

System Output

The trained models were deployed into an interactive web application. The frontend workflow encompasses an input configuration interface, dynamic climate trend visualizations, interactive yield comparison charts, and land-scaled production summaries. The integrated Gemini AI module provides natural language explanations of prediction outputs without altering underlying numerical results.

Summary of Results

The evaluation metrics and scenario projections confirm that crop-specific machine learning models effectively forecast crop yield and production values under climate variability, delivering both quantitative outputs and automated narrative interpretation within a unified web platform.

Discussion

The strong evaluation performance for rice ($R^2 = 0.9617$), maize ($R^2 = 0.9080$), soybeans ($R^2 = 0.8524$), wheat ($R^2 = 0.8157$), and potatoes ($R^2 = 0.7841$) demonstrates that machine learning effectively models climate-yield relationships in data-limited contexts. The success of Gradient Boosting, Random Forest, and Ensemble models confirms that tree-based algorithms excel at capturing non-linear feature interactions and heat/rainfall stress thresholds (Hortizuela & Palaoag, 2025; Islam et al., 2024; Morales & Villalobos, 2023). Conversely, moderate performance in cassava ($R^2 = 0.6728$) and sorghum ($R^2 = 0.6067$) highlights that certain crops remain strongly influenced by unmodeled localized factors such as soil fertility, planting density, irrigation, and pest management.

Scenario projections further illustrate crop-specific climate sensitivity. Cassava achieved the highest projected baseline yield (132,526 hg/ha) alongside the largest uncertainty spread (68,860 hg/ha), indicating high yield potential coupled with strong scenario volatility. In contrast, sorghum recorded the narrowest uncertainty range (1,717 hg/ha),

demonstrating scenario stability despite moderate explanatory performance. This dichotomy emphasizes the necessity of evaluating both statistical accuracy metrics and projection spreads when assessing climate risk.

Compared to data-intensive, process-based models (e.g., DSSAT, APSIM), this system provides an accessible, low-overhead alternative for agricultural planning. Crucially, decoupling numerical ML predictions from the AI recommendation module preserves model integrity, aligning with decision-support governance and explainable AI principles by explaining outputs without altering predictions (Mehryar, 2024; Longo, 2024; Najjar et al., 2024).

Primary system limitations include reliance on historical training data, coarse country-level climate inputs, potential overfitting in select models, and the omission of key field-level variables like soil chemistry, disease pressure, and real-time weather integration (Ahmed, Das, & Zubair, 2024). Incorporating satellite-derived vegetation indices and real-time weather APIs represents a critical direction for future regional scaling (Kira et al., 2024; Hu, Zhang, Bohrer, Liu, & Zhou, 2023).

Conclusion

This study successfully designed, evaluated, and deployed a web-based computational simulation system for crop yield prediction under climate variability in data-scarce environments (Pawlak & Kołodziejczak, 2020; Betew et al., 2026). By engineering agronomic features and leveraging crop-specific machine learning models, the system effectively captures non-linear climate-yield dynamics (Hortizuela & Palaog, 2025; Islam et al., 2024). Model evaluation demonstrated high predictive capability for rice ($R^2 = 0.9617$), maize ($R^2 = 0.9080$), soybeans ($R^2 = 0.8524$), wheat ($R^2 = 0.8157$), and potatoes ($R^2 = 0.7841$), while identifying moderate accuracy for cassava ($R^2 = 0.6728$) and sorghum ($R^2 = 0.6067$) (Seck et al., 2025; Tasneem et al., 2025). Scenario projections for 2026 further highlighted key trade-offs between yield potential and climate risk, exemplified by cassava's high yield potential coupled with wide scenario uncertainty, compared to sorghum's strong scenario stability (Kornhuber et al., 2023; Baluch et al., 2025).

The primary contribution of this work lies in bridging the gap between complex machine learning workflows and accessible decision support (Hernández Hernández et al., 2025; Banerjee et al., 2025). By coupling quantitative yield forecasts with an advisory Gemini AI layer that explains predictive outputs without altering numerical calculations, the system maintains model integrity while enhancing usability for non-technical agricultural stakeholders (Hu et al., 2023; Longo et al., 2024; Mehryar et al., 2024). Future work will focus on integrating real-time weather APIs, sub-national soil parameters, and satellite-derived vegetation metrics to refine localized forecasts and extend the platform's regional scalability (Ahmed et al., 2024; Kira et al., 2024; Pathak et al., 2023).

Data Availability Statement

The data used for this study were obtained from historical agricultural and climate datasets used for model development. Processed data and source code may be made available by the corresponding author upon reasonable request.

Acknowledgments

The author acknowledges the Department of Computer Science, Faculty of Computing, Engineering and Technology, Al-Hikmah University, Ilorin, Nigeria, for academic support. The author also appreciates the project

supervisor, Azeez Rahman Dickola, for his guidance throughout the research.

REFERENCES

- Agarwal, S., & Tarar, S. (2021). A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. *Journal of Physics: Conference Series*, 1714(1), 012012. <https://doi.org/10.1088/1742-6596/1714/1/012012>
- Ahmed, F. U., Das, A., & Zubair, M. (2024). A machine learning approach for crop yield and disease prediction integrating soil nutrition and weather factors. *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*, Dhaka, Bangladesh. <https://arxiv.org/abs/2403.19273>
- Alimagham, S., van Loon, M. P., Ramirez-Villegas, J., Berghuijs, H. N. C., Rosenstock, T. S., & van Ittersum, M. K. (2025). Integrating crop models and machine learning for projecting climate change impacts on crops in data-limited environments. *Agricultural Systems*, 228, 104367. <https://doi.org/10.1016/j.agry.2025.104367>
- Baluch, S. M., Wang, L., Faiz, M. A., Li, H., Chen, Y., Wang, L., & Li, M. (2025). Adaptation simulation and planning for crop yield under climate change: Integrating AquaCrop and DSSAT to project drought-induced yield risks in the Sanjiang Plain. *Agricultural Water Management*, 319, 109818. <https://doi.org/10.1016/j.agwat.2025.109818>
- Banerjee, K., Dutta, S., Das, S., & Sadhukhan, R. (2025). Crop simulation models as decision tools to enhance agricultural system productivity and sustainability: A critical review. *Technology in Agronomy*, 5, e002. <https://doi.org/10.48130/tia-0024-0032>
- Barati, M. K., Soundharajan, B.-S., & Nikoo, M. R. (2024). Simulation of climate-adaptation responses to rainfall variability on rainfed yield anomalies. *Environmental and Sustainability Indicators*, 22, 100411. <https://doi.org/10.1016/j.indic.2024.100411>
- Betew, A. G., Gebresenbet, G., Gelaw, G. K., Mengistu, D. A., & Yibre, A. M. (2026). Historical and contemporary crop yield prediction models: Key lessons and innovations. *Smart Agricultural Technology*, 13, 101672. <https://doi.org/10.1016/j.atech.2025.101672>
- Chisanga, C. B., Phiri, E., & Chinene, V. R. N. (2017). Climate change impact on maize (*Zea mays* L.) yield using crop simulation and statistical downscaling models: A review. *Scientific Research and Essays*, 12(18), 167–187. <https://doi.org/10.5897/SRE2017.6521>
- Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environmental Research Letters*, 13(11), 114003. <https://doi.org/10.1088/1748-9326/aae159>
- El-Mahroug, S. E., Suleiman, A. A., Zoubi, M. M., Al-Omari, S., Abu-Afifeh, Q. Y., Al-Jawaldeh, H. F., Alta'any, Y. A., Al-Nawaiseh, T. M. F., Obeidat, N., Alsoud, S. H., Alshoshan, A. M., Al-Shibli, F. M., & Ta'any, R. (2025). Predictive modeling of climate-driven crop yield variability using DSSAT towards sustainable agriculture. *AgriEngineering*, 7(5), 156. <https://doi.org/10.3390/agriengineering7050156>

- Ewert, F., Baatz, R., & Finger, R. (2023). Agroecology for a sustainable agriculture and food system: From local solutions to large-scale adoption. *Annual Review of Resource Economics*, 15, 351–381. <https://doi.org/10.1146/annurev-resource-102422-090105>
- Farah, A. A., Mohamed, M. A., Musse, O. S. H., & Nor, B. A. (2025). The multifaceted impact of climate change on agricultural productivity: A systematic literature review of SCOPUS-indexed studies (2015–2024). *Discover Sustainability*, 6, 397. <https://doi.org/10.1007/s43621-025-01229-2>
- Hernández Hernández, G. C., Gómez Gómez, J., & Jiménez-Cabas, J. (2025). Predictive models based on artificial intelligence to estimate crop yield: A literature review. *Agriculture*, 15(23), 2438. <https://doi.org/10.3390/agriculture15232438>
- Hortizuela, R. D., & Palaoag, T. D. (2025). Predicting staple crop yields under climate variability using multiple regression techniques. *Indonesian Journal of Electrical Engineering and Computer Science*, 40(3), 1531–1538. <https://doi.org/10.11591/ijeecs.v40.i3.pp1531-1538>
- Hu, T., Zhang, X., Bohrer, G., Liu, Y., Zhou, Y., Martin, J., Li, Y., & Zhao, K. (2023). Crop yield prediction via explainable AI and interpretable machine learning: Dangers of black box models for evaluating climate change impacts on crop yield. *Agricultural and Forest Meteorology*, 336, 109458. <https://doi.org/10.1016/j.agrformet.2023.109458>
- Hultgren, A., Carleton, T., Delgado, M., Gergel, D. R., Greenstone, M., Houser, T., Hsiang, S., Jina, A., Kopp, R. E., Malevich, S. B., McCusker, K. E., Mayer, T., Nath, I., Rising, J., Rode, A., & Yuan, J. (2025). Impacts of climate change on global agriculture accounting for adaptation. *Nature*, 642, 644–652. <https://doi.org/10.1038/s41586-025-09085-w>
- International Food Policy Research Institute. (2023). *Global food policy report 2023: Rethinking food crisis responses*. International Food Policy Research Institute. <https://doi.org/10.2499/9780896294417>
- Islam, M. M., Alharthi, M., Alkadi, R. S., Islam, R., & Masum, A. K. M. (2024). Crop yield prediction through machine learning: A path towards sustainable agriculture and climate resilience in Saudi Arabia. *AIMS Agriculture and Food*, 9(4), 980–1003. <https://doi.org/10.3934/agrfood.2024053>
- Kallenberg, M. G. J., Maestrini, B., van Bree, R., Ravensbergen, P., Pylaniadis, C., van Evert, F., & Athanasiadis, I. N. (2023, July). *Integrating process-based models and machine learning for crop yield prediction*. Paper presented at the 1st Workshop on Synergy of Scientific and Machine Learning Modeling, International Conference on Machine Learning, Honolulu, HI, United States.
- Kamangir, H., Sams, B. S., Dokoozlian, N., Sanchez, L., & Earles, J. M. (2024). CMAViT: Integrating climate, management, and remote sensing data for crop yield estimation with multimodel vision transformers. *arXiv*. <https://arxiv.org/abs/2411.16989>
- Kira, O., Wen, J., Han, J., McDonald, A. J., Barrett, C. B., Ortiz-Bobea, A., Liu, Y., You, L., Mueller, N. D., & Sun, Y. (2024). A scalable crop yield estimation framework based on remote sensing of solar-induced chlorophyll fluorescence (SIF). *Environmental Research Letters*, 19(4), 044071. <https://doi.org/10.1088/1748-9326/ad3142>
- Kornhuber, K., Lesk, C., Schleussner, C. F., Jägermeyr, J., Pfleiderer, P., & Horton, R. M. (2023). Risks of synchronized low yields are underestimated in climate and crop model projections. *Nature Communications*, 14, 3528. <https://doi.org/10.1038/s41467-023-38906-7>
- Koukoulis, P., Georgiou, P., & Karpouzios, D. (2025). Assessment of the impacts of climate change scenarios on maize yield and irrigation water using the CropSyst model: An application in Northern Greece. *Agronomy*, 15(3), 638. <https://doi.org/10.3390/agronomy15030638>
- Li, D., Kitsios, V., Newth, D., & O’Kane, T. J. (2025). A Bayesian ensemble projection of climate change and technological impacts on future crop yields. *arXiv*. <https://arxiv.org/abs/2507.21559>
- Longo, L., Bric, M., Cabitza, F., Choi, J., Confalonieri, R., Del Ser, J., Guidotti, R., Hayashi, Y., Herrera, F., Holzinger, A., Jiang, R., Khosravi, H., Lecue, F., Malgieri, G., Páez, A., Samek, W., Schneider, J., Speith, T., & Stumpf, S. (2024). Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. *Information Fusion*, 106, 102301. <https://doi.org/10.1016/j.inffus.2024.102301>
- Maazallahi, A., Thota, S., Kondaboina, N. P., Muktineni, V., Annem, D., Rokkam, A. S., Amini, M. H., Salari, M. A., Norouzzadeh, P., Snir, E., & Rahmani, B. (2024). Naïve Bayes and Random Forest for crop yield prediction. *arXiv*. <https://arxiv.org/abs/2404.15392>
- Mehryar, S., Yazdanpanah, V., & Tong, J. (2024). AI and climate resilience governance. *iScience*, 27(6), 109812. <https://doi.org/10.1016/j.isci.2024.109812>
- Morales, A., & Villalobos, F. J. (2023). Using machine learning for crop yield prediction in the past or the future. *Frontiers in Plant Science*, 14, 1128388. <https://doi.org/10.3390/fpls.2023.1128388>
- Najjar, H., Miranda, M., Nuske, M., Roscher, R., & Dengel, A. (2024). Explainability of sub-field level crop yield prediction using remote sensing. *arXiv*. <https://arxiv.org/abs/2407.08274>
- Pathak, D., Miranda, M., Mena, F., Sanchez, C., Helber, P., Bischke, B., Habelitz, P., Najjar, H., Siddamsetty, J., Arenas, D., Vollmer, M., Charfuelan, M., Nuske, M., & Dengel, A. (2023). Predicting crop yield with machine learning: An extensive analysis of input modalities and models on a field and sub-field level. *arXiv*. <https://arxiv.org/abs/2308.08948>
- Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylaniadis, C., & Athanasiadis, I. N. (2021). Machine learning for large-scale crop yield forecasting. *Agricultural Systems*, 187, 103016. <https://doi.org/10.1016/j.agsy.2020.103016>
- Pawlak, K., & Kołodziejczak, M. (2020). The role of agriculture in ensuring food security in developing countries: Considerations in the context of the problem of sustainable

- food production. *Sustainability*, 12(13), 5488. <https://doi.org/10.3390/su12135488>
- Saleem, A., Anwar, S., Nawaz, T., Fahad, S., & Saud, S. (2025). Securing a sustainable future: The climate change threat to agriculture, food security, and sustainable development goals. *Journal of Umm Al-Qura University for Applied Sciences*. <https://doi.org/10.1007/s43994-024-00147-w>
- Seck, N. K. G., Ngom, A., Ngom, P., & Noba, K. (2025). Crop yield forecasting in Senegal: Application of machine learning methods. *Discover Agriculture*, 3, 192. <https://doi.org/10.1007/s44279-025-00381-7>
- Talaat, F. M. (2023). Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes. *Neural Computing and Applications*, 35(23), 17281–17292. <https://doi.org/10.1007/s00521-023-08619-5>
- Tasneem, K. T., Shahzad, M. U., Rashid, J., Othman, K. M., Zafar, T., & Faheem, M. (2025). Predicting rice yield and impact of climate change on rice production using machine learning models. *Theoretical and Applied Climatology*, 156, 665. <https://doi.org/10.1007/s00704-025-05912-2>
- Timlin, D., Paff, K., & Han, E. (2024). The role of crop simulation modeling in assessing potential climate change impacts. *Agrosystems, Geosciences & Environment*, 7(1), e20453. <https://doi.org/10.1002/agg2.20453>
- Venugopal, A., Aparna, S., Mani, J., Mathew, R., & Williams, V. (2021). Crop yield prediction using machine learning algorithms. *International Journal of Engineering Research & Technology*, 9(13), 87–91.
- Yadav, B., Krishnan, P., Parihar, C. M., & Banerjee, K. (2025). Modelling crop growth and soil hydrothermal regimes under conservation agriculture using APSIM-wheat. *Scientific Reports*, 15, 36362. <https://doi.org/10.1038/s41598-025-20211-6>
- Yang, Y., Yang, Y., Han, S., Macadam, I., & Liu, D. L. (2014). Prediction of cotton yield and water demand under climate change and future adaptation measures. *Agricultural Water Management*, 144, 42–53. <https://doi.org/10.1016/j.agwat.2014.06.001>
- Zare, M., Azam, S., & Sauchyn, D. (2023). Simulation of climate change impacts on crop yield in the Saskatchewan Grain Belt using an improved SWAT model. *Agriculture*, 13(11), 2102. <https://doi.org/10.3390/agriculture13112102>

