Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://asr.nsps.org.ng/index.php/asr

African Scientific Reports 5 (2026) 590

African Scientific Reports

Crop recommendation in precision agriculture: a systematic literature review of methods, trends, and challenges

Ebenezer O. Oladipe[a,b,*], Sunday E. Adewumi[b], Taiwo Kolajo[b], Joshua B. Agbogun[b]

[a]Department of Information Technology, Confluence University of Science and Technology, Osara, Kogi State, Nigeria
[b]Department of Computer Science, Federal University Lokoja, Kogi State, Nigeria

Abstract

Crop recommendation methods have become an essential component of modern agriculture, helping farmers identify the most suitable crops based on soil, climatic, and environmental conditions. As key applications of precision agriculture, these methods increasingly employ artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), to improve crop prediction, recommendations, productivity, and farming decision-making. This study presents a systematic literature review (SLR) of ML and DL techniques applied to crop recommendation and related agricultural applications. From 183 identified studies, 129 articles published between 2020 and 2026 were carefully selected through a structured screening process for comprehensive analysis. Relevant studies were retrieved from major academic literature databases and publishing platforms, including ScienceDirect, Scopus, SpringerLink, MDPI, Nature, Frontiers, IEEE, Wiley, and Google Scholar. The review used the PRISMA protocol and showed that ensemble learning-based approaches, particularly Random Forest and Extreme Gradient Boosting (XGBoost), are powerful for predictive performance on various agricultural datasets. Traditional ML approaches such as Support Vector Machines, Decision Trees, and k-Nearest Neighbors are still commonly used. At the same time, Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are used for remote sensing and time-dependent agricultural analysis. The most frequently used dataset source is Kaggle, and typical inputs include soil nutrients (NPK), soil pH, weather conditions, and satellite indices such as NDVI and EVI. Many studies have achieved high accuracy, but most are based on static datasets, which reduces their reliability in real-world scenarios. The main research gaps are limited real-time deployment, low integration of multiple data sources, low cross-regional validation, and low model interpretability. The review shows the importance of scalable and explainable AI systems for real applications in agriculture.

DOI: 10.46481/asr.2026.5.3.590

Keywords: Precision agriculture, Crop recommendation, Machine learning, Autonomous decision-making

Article History:
Received: 19 June 2026
Received in revised form: 06 July 2026
Accepted for publication: 07 July 2026
Available online: 20 August 2026

1. Introduction

Agriculture is a basic component of rural development, economic stability, and global food security [1, 2]. However, the sector faces major challenges, including climate variability, soil degradation, and inefficiencies in traditional farming methods [2]. Traditional practices rely on the wisdom of generations but lack the precision and adaptability needed in today's complex agricultural

*Corresponding Author Tel. No.: +234-806-929-3332.
Email address: oladipebenezer@gmail.com (Ebenezer O. Oladipe)

systems [2]. The world's population is projected to reach 9.8 billion by 2050, and data-driven approaches are needed to optimize crop selection and maximize productivity on limited land [3]. Precision agriculture has been presented as a game-changing approach that integrates broad environmental data with actionable farming decisions using modern technologies [2, 3].

With rapid developments in machine learning (ML), remote sensing (RS), and the Internet of Things (IoT), significant changes have occurred in agricultural decision-making [3–5]. Landsat 8 and 9 satellite platforms provide tools for extracting spatial and temporal variation in climate and soil factors [2]. At the same time, IoT-based sensors provide real-time farm-level data on different factors, including soil pH, moisture, and the nitrogen–phosphorus–potassium (NPK) triplet [3–5]. Data obtained from different sources are processed by algorithms ranging from Random Forest and Support Vector Machine to deep neural networks to recommend an optimal crop choice for a region [5–7]. Recently, there has been growing attention to nature-inspired optimization models, such as hybrid gravitational search algorithms, for improving crop recommendation results [2]. However, restrictions remain in the field use of this technology, even though some state-of-the-art approaches achieve high accuracy. One practical challenge is the black-box nature of models, which can make farmers hesitant to trust automatic suggestions [6, 8]. Researchers have started adopting explainable artificial intelligence (XAI) frameworks, such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), to explain which factors contribute to model predictions [5, 8]. Other challenges include temporal uncertainty and the ability to generalize across different agro-ecological zones [2, 6, 7]. Therefore, this systematic literature review (SLR) summarizes recent advances in precision agriculture for crop recommendation, feature selection, and novel algorithmic developments toward sustainable agriculture.

The relevant literature searched in electronic databases is presented and discussed based on the study questions. The remaining sections are organized as follows. Related work is reviewed in Section 2. The survey methodology is presented in Section 3. Section 4 presents the SLR results for the research questions and discusses these results. Section 5 presents future research directions, and Section 6 concludes the paper.

2. Mapping studies

Several SLRs and review studies have examined the role of artificial intelligence (AI) in agriculture. Most current reviews focus on particular crop production issues, specific crops, or AI approaches; therefore, broader synthesized overviews of crop recommendation are still needed. For instance, in an SLR of 25 studies on machine learning approaches in crop yield forecasting conducted under the PRISMA protocol, the dominant input features included soil parameters, weather elements, and crop data [9].

The study revealed that Support Vector Machine (SVM), k-nearest neighbors (KNN), and Extreme Gradient Boosting (XGBoost) were the most extensively used algorithms.

Although the article concluded that AI is increasingly relevant in agricultural decision-making, it was limited to yield prediction and did not address generalized crop recommendation applications. Similarly, another review [10] analyzed 82 publications on crop yield prediction, specifically maize and soybeans, and examined machine learning, deep learning, and ensemble learning techniques. That review noted that temperature, precipitation, historical yields, and normalized difference vegetation index (NDVI) were among the most common input factors. It highlighted Random Forest as the most robust machine learning technique and long short-term memory (LSTM) networks as efficient for modeling temporal patterns; however, it was limited to two crop varieties and therefore provided restricted insight into the scope of AI-enabled crop recommenders. Remote sensing techniques in crop prediction have also drawn attention from several researchers.

In evaluating the integration of deep learning (DL) with remote sensing, Ref. [11] investigated publications from the previous 10 years and identified vegetation indices, such as NDVI, and satellite imagery from sensors such as MODIS as instrumental in predicting crop yield.

The results indicated that DL models, including convolutional neural networks (CNNs) and LSTMs, were effective for capturing spatial and temporal patterns, respectively. However, this study was limited to remote sensing and did not comprehensively cover machine learning approaches, data sources, and measurement tools. A few reviews were also tailored to specialized agricultural areas.

For example, Ref. [12] presented a systematic review on palm oil yield prediction based on 223 studies and proposed an integrated model for predicting yield by combining three successful algorithms: Random Forest, artificial neural network, and support vector regression. They concluded that an integrated model combining the strengths of different machine learning models leads to better performance than a single algorithm. In plant protection, Ref. [13] systematically reviewed 60 articles focusing on banana disease detection and fruit ripeness assessment. That review suggested SVM in combination with GLCM and SIFT for plant disease detection, with accuracy ranging from 80.00% to 99.61%, whereas CNN-based methods were predominantly used for image-based detection and assessment.

Similar to the other studies mentioned above, this work focused on a particular crop within a narrow scope and therefore could not directly support the development of comprehensive general crop recommenders.

In summary, existing studies show that machine learning and DL have been extensively applied across diverse agricultural domains, such as yield forecasting, disease detection, and remote sensing. However, most reviews are specific to a crop, problem type, machine learning or DL technique, feature set, or data source. Relatively few comprehensive reviews address crop recommendation

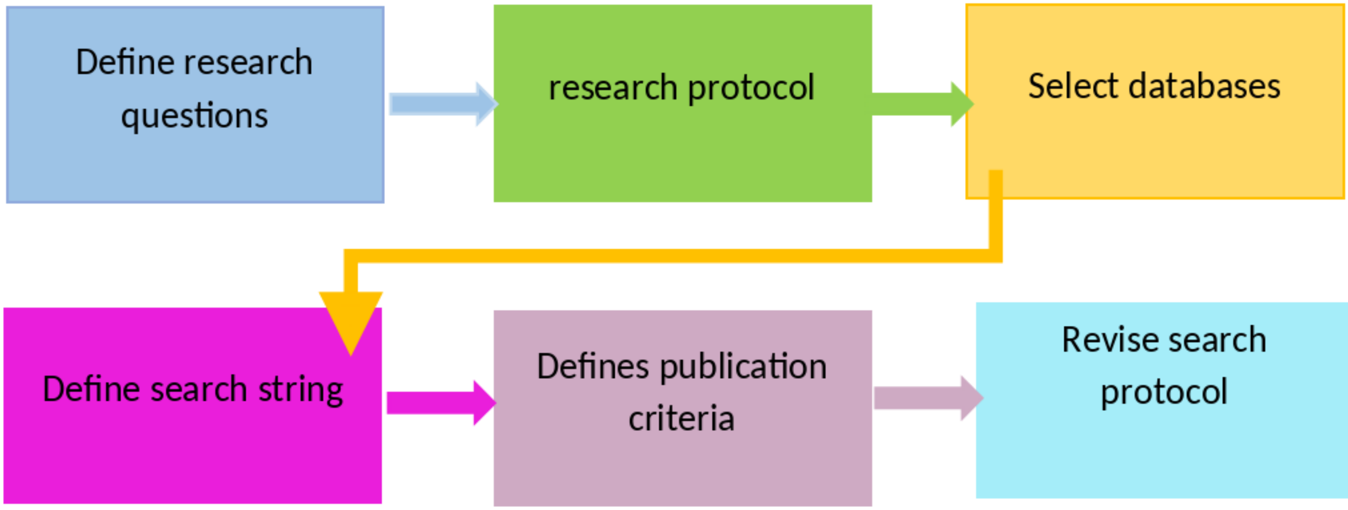

Figure 1: Detailed schedule of the review method.

techniques while considering data sources, environmental and soil inputs, machine learning and DL approaches, performance measures, deployment scenarios, publication trends, challenges, and future directions in a systematic fashion. In particular, recent work published between 2020 and 2026 has not yet been comprehensively synthesized. This gap motivates the current study, which offers a holistic perspective on AI-based crop recommendation.

3. Research methodology

3.1. Protocol for review

Before carrying out a systematic review, a review protocol has to be produced. This systematic review was conducted using the PRISMA guidelines. The first process was the specification of the research questions. Once the research questions were stated, databases were searched to gather appropriate studies. The publication search sources considered for search queries included Frontiers, Nature, ScienceDirect, Scopus, IEEE, MDPI, SpringerLink, Wiley, and Google Scholar to access relevant articles across a diverse range of outlets. The review process using the PRISMA protocol was organized into three stages: planning, executing, and reporting. In the conducting stage, studies meeting the inclusion and exclusion criteria were obtained through the search strategy. Identified studies were evaluated for quality, and data related to authors, year of publication, methodologies, and results relevant to the research questions were extracted, summarized, tabulated, and organized systematically. The summarized data were then reported to answer the research questions.

3.2. Selection strategy

This research followed a systematic literature search, which enabled the most appropriate, high-quality, accurate, and scientifically valid articles to be identified and reviewed. Major academic databases and publishing platforms, including Frontiers, Nature, Science, ScienceDirect, Scopus, IEEE, MDPI, SpringerLink, Wiley, and Google Scholar, were thoroughly searched using systematically derived queries. The search included words and phrases such as "crop recommendation system", "machine learning", "deep learning", "precision agriculture", and "intelligent agriculture", along with their variants.

The scope was restricted to peer-reviewed articles published in English between 2020 and 2026 to promote consistency and accessibility.

All retrieved articles were screened as follows. First, duplicates were excluded from the articles obtained through the systematic search. Second, papers whose titles and abstracts showed no relevance to machine learning and deep learning crop recommendation systems, as well as conference papers, opinion/editorial/review papers, and articles with poorly documented methodologies, were excluded. Third, full texts of eligible papers were reviewed according to the pre-set inclusion and exclusion criteria to check whether they addressed the review questions. Fourth, eligible papers were assessed for quality according to an adjusted list based on accepted systematic review guidelines to ensure reliable results. Fifth, only high-quality papers were retained in the synthesis.

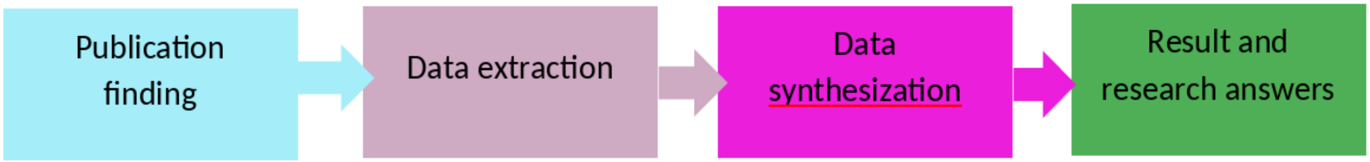

Figure 2: Procedures for carrying out and reporting the review.

3.3. Research question formulation

The following research questions were developed to pursue the aim of this SLR and produce a complete survey of crop recommender technologies.

RQ1. What machine learning algorithms are proposed in the literature for crop prediction?
RQ2. What relevant input features are being exploited?
RQ3. What evaluation methods and metrics are applied for performance assessment?
RQ4. What limitations are found in previous studies?
RQ5. What practical challenges are faced by crop recommender systems?
RQ6. How can farmers deal with the black-box problem of crop recommender systems?

3.4. Search strategy

We searched for foundational ideas relevant to the objectives of this review. Because ML can be applied in numerous domains, many publications fall outside the scope of this review.

The relevant studies were selected through a literature search in major scientific databases, including Frontiers, Nature, ScienceDirect, Scopus, IEEE, MDPI, SpringerLink, Wiley, and Google Scholar.

Figure 1 shows the structured process used to perform the SLR. First, a clear research question was raised. Second, a research protocol was created, stating the systematic approach and methods of the review. Third, the major research databases were selected, a search string was developed to identify relevant studies, and publication inclusion and exclusion criteria were specified. Finally, the search protocol was refined as needed to maximize the accuracy of relevant retrieval.

Figure 2 shows the overall procedure for implementation and reporting. The procedures included publication finding, through which relevant studies were systematically identified based on specific search methods and inclusion criteria; data extraction, through which important data such as research details, methods, characteristics, algorithms, and evaluation metrics were documented; data synthesis, in which patterns, themes, and research gaps were organized and compared; and reporting, in which the study results and research answers were presented.

In this SLR, we considered several terms for crop recommender systems, including “crop selection” and “crop suggestion”. We used the terms “artificial intelligence”, “machine learning”, and “deep learning” to identify articles based on newer approaches using agricultural data. An automatic search was performed. Initially, the keywords “precision agriculture” and “crop prediction” were used. Synonyms of the terms were identified through abstract screening and information extraction from available articles. A database search using “precision agriculture AND crop prediction” generated a comprehensive set of studies on the subject.

3.5. Exclusion and inclusion criteria

To increase the validity of the SLR, studies without an examination component were excluded. The following exclusion criteria (EC) were used:

EC1. Letters, notes, and patents were excluded from the examination.
EC2. Papers submitted by graduates were not examined.
EC3. Studies published in languages other than English were excluded.
EC4. Studies in which the author failed to clarify the methodological approach were excluded.
EC5. Papers published before 2020 were excluded.

The following inclusion criteria (IC) were used:

IC1. Articles that used machine learning methods for crop recommendation or prediction were included.
IC2. English-language articles were included.
IC3. Articles published from 2020 to 2026 were included.
IC4. Open-access articles were included.
IC5. Peer-reviewed papers/proceedings were included.

Table 1: Database distribution based on final selected papers.

Journal Database	Initial papers retrieved	Frequency of occurrence
ScienceDirect	57	36
Scopus	33	27
SpringerLink	24	17
MDPI	29	16
Nature	18	13
Google Scholar	13	12
Frontiers	4	3
IEEE	3	3
Wiley	2	2
Total	183	129

3.6. Data collection process

Data from the selected studies were collected and collated to answer the research questions. The extracted information included publication type and year, search library, models addressing the problem, characteristics used in each study, performance metrics used to evaluate proposed models, and limitations in each study. After applying the initial exclusion criteria, 141 of 183 studies were included for analysis. After applying all exclusion criteria, 129 studies were selected for analysis.

Table 1 reports the number of initially retrieved papers and final selected papers from each database. Initially, 183 papers were collected across the nine databases, and 129 papers satisfied the inclusion criteria and were chosen for the SLR. ScienceDirect had the largest number of final selected studies, with 36 articles, followed by Scopus (27), SpringerLink (17), and MDPI (16). Nature, Google Scholar, and Frontiers contributed 13, 12, and 3 studies, respectively. IEEE and Wiley provided 3 and 2 papers, respectively. These findings indicate that ScienceDirect and Scopus were the two main databases supplying information about machine learning and DL applications in crop recommendation. Together, they accounted for approximately half of the selected studies. Although some databases, such as Frontiers, IEEE, and Wiley, contributed fewer papers, the studies obtained from them supplemented the evidence base and increased diversity.

The studies selected for detailed analysis are presented in the Appendix. During data synthesis, all extracted data were aggregated to answer the research questions. The synthesis results are presented in Section 4.

Figure 3 presents the flow of information from initial search results to the final set of studies included from the nine selected academic databases. A total of 183 articles were searched from the online databases, of which 129 studies fulfilled the inclusion and quality-check criteria and were retained for final analysis. The decrease in studies in every database after screening indicates the effectiveness of the search and review process.

Of the 183 studies searched from these databases, ScienceDirect contained the largest number of relevant documents searched and retained (57; 36), followed by Scopus (33; 27). SpringerLink (24; 17), MDPI (29; 16), Nature (18; 13), and Google Scholar (13; 12) also contributed substantial numbers of articles. Frontiers (4; 3), IEEE (3; 3), and Wiley (2; 2) contributed fewer articles but remained valuable for broadening the research base. This multi-database research process included peer-reviewed studies from several research streams and reduced potential publication bias.

4. Results and discussion

There is a clear upward trend in research on the application of machine learning in agriculture after 2022. Table 2 summarizes the number of publications between 2020 and 2026, with a total of 129 unique articles analyzed.

The quantity of machine learning applications for crop forecasting in the literature has grown substantially in recent years. After remaining relatively stable during the early 2020s, at between 11 and 14 articles per year, the body of work increased to 20 articles in 2022. This trend continued through the latter years of the analysis period, with 2025 and 2026 contributing close to 44% of the total number of unique research papers (n = 129), with 29 and 28 articles, respectively. This growth is associated with increased data availability and processing power. There is also a notable increase in the sophistication of computational methodologies, with many 2025 and 2026 studies using more specialized model types, such as dual-stacking ensembles, transformer models, and robust regression, rather than only simple model configurations. This indicates an increasingly mature field focused on developing practical agricultural models. For consistent analysis, papers covering more than one year were attributed to the later year.

4.1. RQ1: Technique-related analysis

Based on the literature reviewed, Table 3 shows the frequency of occurrence of the primary machine learning algorithms used in crop prediction. Figure 4 shows the distribution of selected publications per year.

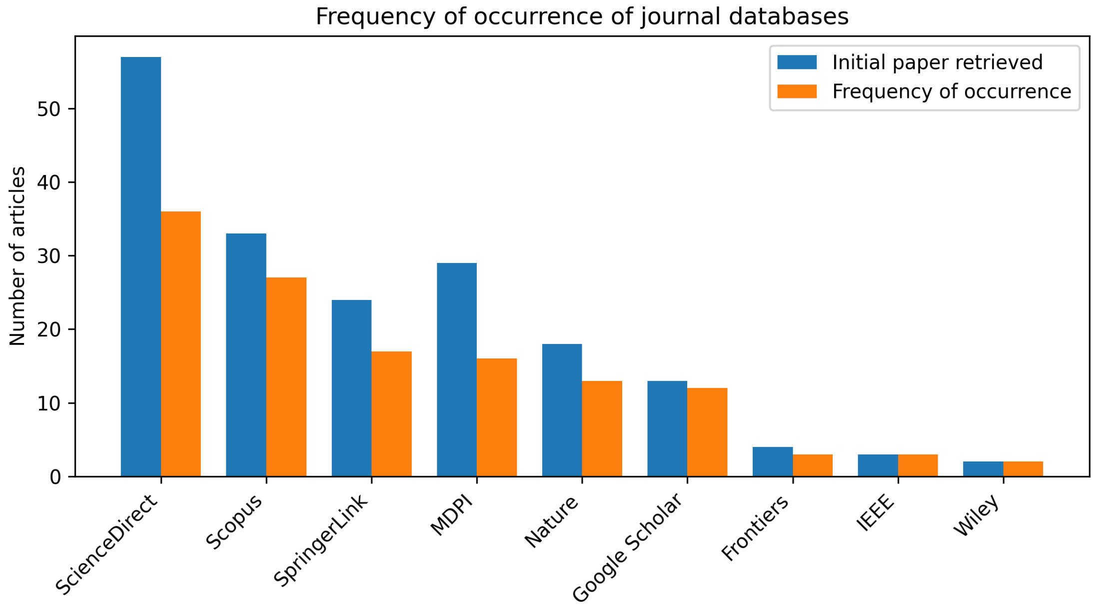

Figure 3: Frequency of occurrence of journal databases.

Table 2: Distribution of publications per year.

Publication Year	Frequency of Occurrence
2020	11
2021	14
2022	20
2023	12
2024	15
2025	29
2026	28
Total	129

As shown in Figure 5, Random Forest is the most extensively implemented machine learning algorithm in the reviewed studies. It appears in more than 70% of the publications and is associated with reported accuracies of approximately 97%–99.9%. Ensemble and boosting methods, such as XGBoost and Gradient Boosting, were preferred for their stable and robust performance. Beyond classical machine learning algorithms, DL methods such as CNN, LSTM, and artificial neural network (ANN) models are increasingly implemented for processing complex images and time-series data in agriculture. Various ensemble and hybrid methods, including CNN–LSTM and particle swarm optimization–SVM (PSO–SVM), along with XAI methods such as SHAP and LIME, are also being used to improve understanding of model prediction results.

4.2. RQ2: Feature-related analysis

Identifying optimal input features is essential for developing accurate and trustworthy crop prediction models. From the literature examined, three major classes of inputs were identified: climate, soil, and remote sensing/management data. Table 4 indicates the number of times that specific input features were identified as important across the corpus. Figure 6 visualizes these inputs as a bar chart. Of the three input types considered, climate and soil characteristics had the greatest influence on agricultural productivity and therefore had the greatest presence in the literature. Climatic variables such as temperature (79 occurrences) and precipitation (73 occurrences) were key in modeling growth stages and water availability, while relative humidity (70 occurrences) was important for understanding disease risk and estimating evapotranspiration. Soil inputs, particularly pH (66 occurrences) and combined NPK inputs (60 occurrences each), provided foundational information on chemical characteristics related to growth potential. Remote sensing/management data were less represented but provided real-time monitoring and contextual depth for crop production decisions. These inputs included vegetation indices (18 occurrences) obtained from satellite/unmanned aerial vehicle (UAV) remote sensing, as well as management factors (9 occurrences), such as fertilizer inputs, pesticide treatments, and field history.

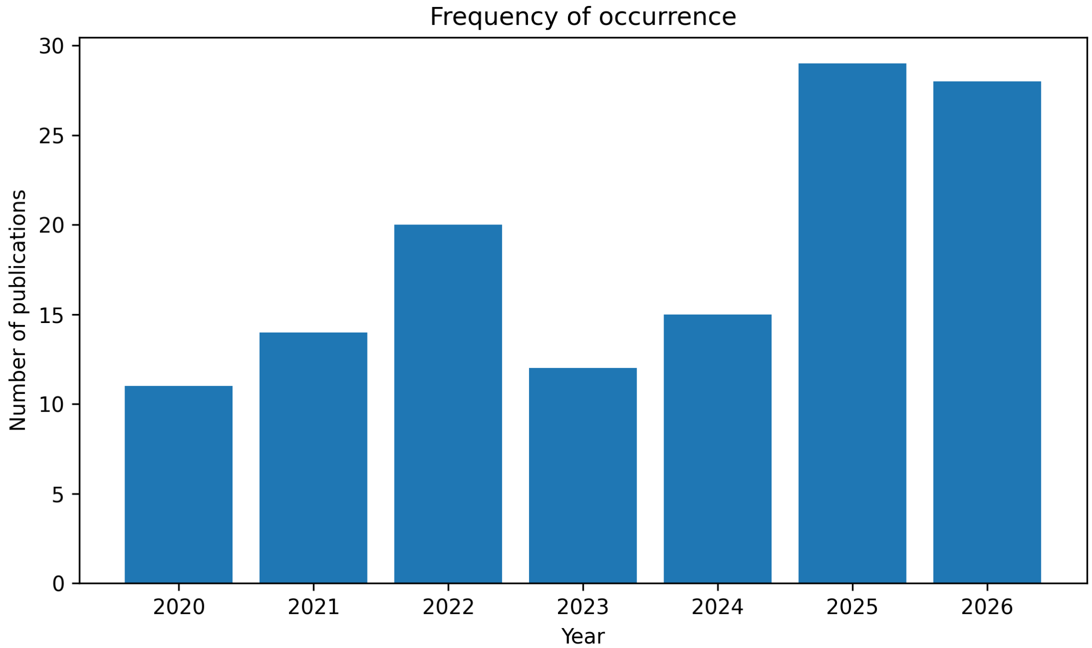

Figure 4: Distribution of selected publications per year.

Table 3: Frequency of machine learning algorithms in the literature.

Category	Algorithm Type	Frequency
Machine Learning	Random Forest (RF)	82
	Support Vector Machine (SVM/SVR)	48
	Decision Tree (DT)	29
	K-Nearest Neighbors (KNN)	24
	Linear/Logistic Regression & LASSO	26
	Naive Bayes (NB)	9
Deep Learning	LSTM / RNN / GRU	44
	Convolutional Neural Networks (CNN)	29
	Artificial Neural Networks (ANN/MLP)	23
	DNN / TabNet / Transformers	7
Ensemble Learning	XGBoost (XGB)	31
	General Ensembles (Bagging, Voting, Stacking)	24
	Gradient Boosting (GB / GBDT)	14
	LightGBM / CatBoost / AdaBoost	15
Others	Explainable AI (SHAP / LIME)	12
	Optimization & Fuzzy Logic (PSO, etc.)	6

4.3. *RQ3: Evaluation-technique-related analysis*

Several techniques are used to evaluate machine learning models applied to agricultural purposes. For crop recommendation, Table 5 and Figure 7 show that accuracy is one of the most frequently used metrics for assessing classification models because it indicates the proportion of correct recommendations. Precision and recall are also considered in crop recommendation, while the F1-score evaluates robustness across crop classes. Many authors reported accuracy values close to 1.

The primary metrics used for crop prediction are root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). RMSE and MAE describe the distance between model predictions and actual values, whereas R^2 indicates how well the model explains variability in the data. Modern crop prediction studies also incorporate cross-validation to estimate model robustness and use explainability tools such as SHAP or LIME to support interpretation. However, high prediction rates are often calculated using clean datasets and may not fully characterize real-world farm settings.

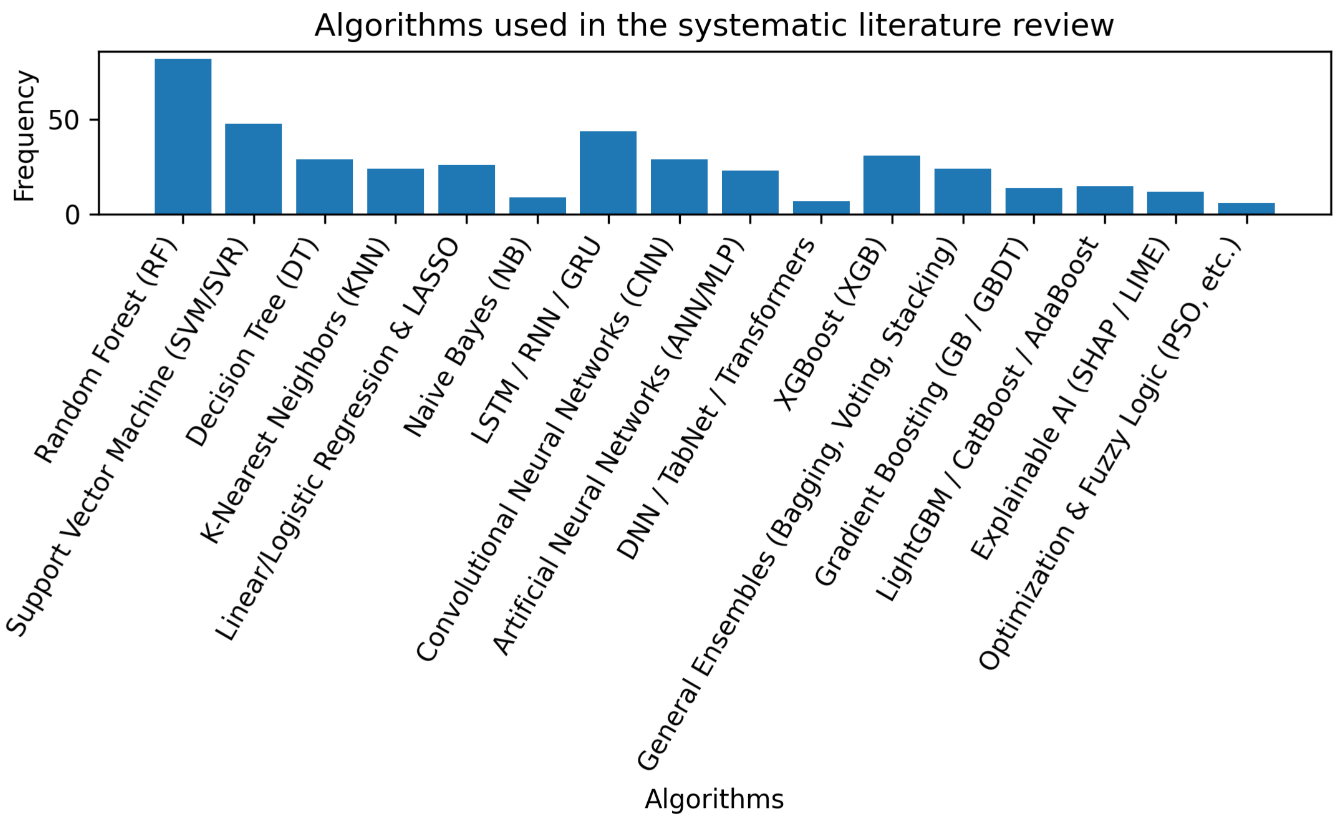

Figure 5: Algorithms used in the systematic literature review.

Table 4: Frequency of input features in crop prediction literature.

Feature Category	Input Feature	Total Occurrences
Climatic Variables	Temperature (Average, Min, Max)	79
	Rainfall / Precipitation	73
	Relative Humidity	70
	Solar Radiation / Irradiance	11
	Wind Speed	5
Soil Properties	Soil pH	66
	Nitrogen (N)	60
	Phosphorus (P)	60
	Potassium (K)	60
	Soil Moisture / Water Content	27
	Soil Type / Texture / Color	12
Advanced Features	Vegetation Indices (NDVI, EVI, NDWI)	18
	Satellite / UAV Imagery	16
	Historical Yield Records	14
	Management (Pesticides, Fertilizer Volume)	9

4.4. RQ4: Limitation analysis

Current crop prediction literature demonstrates a gap between high-performing experimental models and the unpredictable, noisy reality of field deployment. Most architectures report accuracy values close to 100% on clean published benchmarks but remain affected by long-term uncertainty and limited adaptability to climatic variation because of small or potentially biased datasets. The black-box problem is also important: network abstractions such as CNNs and LSTMs can produce recommendations that farmers cannot readily interpret. This limitation becomes more significant when models are overparameterized or overgeneralized, as networks trained on specific datasets may underperform in other agro-ecological zones.

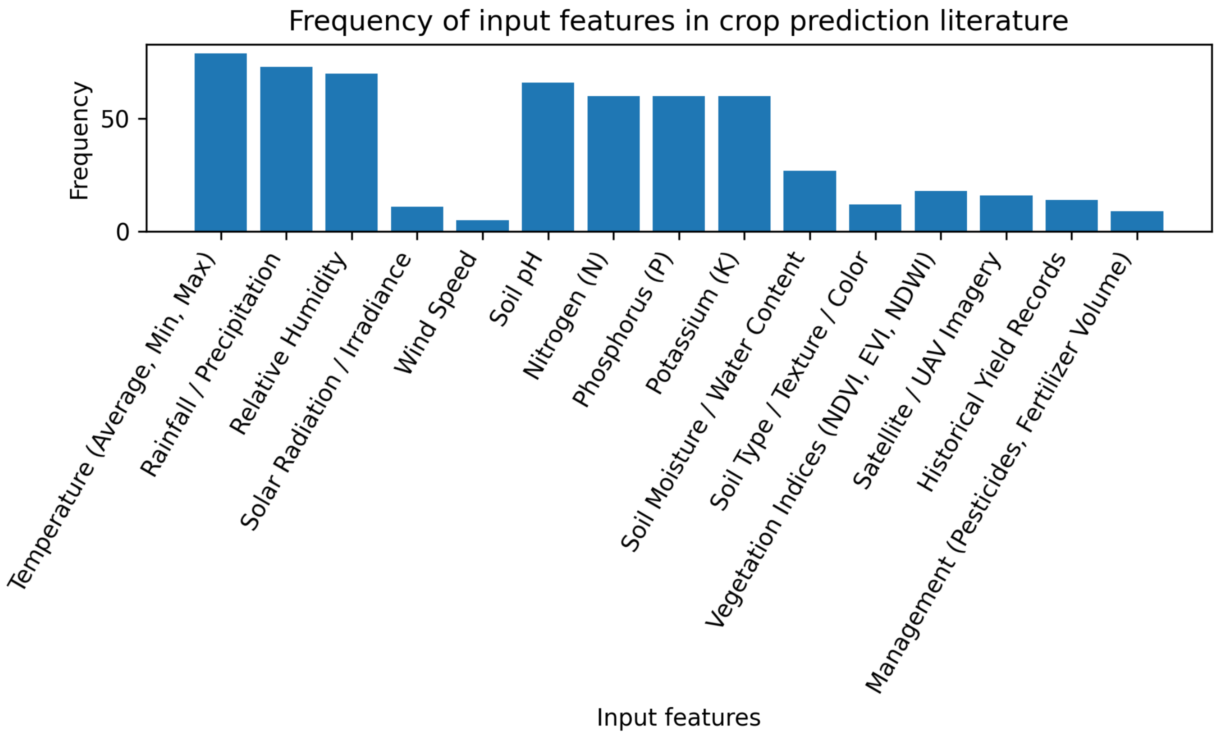

Figure 6: Bar chart representation of crop prediction features.

Table 5: Evaluation techniques used in the literature.

Evaluation Technique	Total Occurrences	Context of Application
Accuracy	49	Primarily used for crop recommendation systems to measure the percentage of correct classifications.
RMSE (Root Mean Square Error)	31	The standard metric for prediction; it measures the average magnitude of the prediction error.
R^2 Score (Coefficient of Determination)	28	Used to determine how well the model explains the variability within the agricultural data.
MAE (Mean Absolute Error)	18	Measures the average absolute difference between predicted and actual yield values.
Precision, Recall, and F1-Score	15	Used to evaluate the balance between precision and sensitivity, especially in classification models.
MSE (Mean Squared Error)	9	Used in forecasting to penalize larger errors more heavily than smaller ones.
Kappa Coefficient	3	Measures the agreement between predicted and observed classifications while accounting for chance.

Beyond feature selection, as shown in Figure 8, few applications include economic features such as commodity prices and production costs; many focus only on ecological and climatic information. In addition, static values based on historical averages, rather than real-time IoT data, reduce real-time applicability. Real-time deployment constraints (30%) and scalability constraints (26%) are therefore major concerns hindering adoption, especially when combined with infrastructure deficits in rural areas, including unstable electricity and low digital awareness. To develop practical tools from experimental crop prediction studies, more attention should be given to interpretable multi-source applications that account for environmental and socio-economic factors.

4.5. RQ5: Challenge analysis

Several major challenges remain in transforming theoretically designed crop recommendation systems into real-world agricultural tools. These challenges can be classified as data constraints, technical constraints, and real-world application constraints.

Data quality and availability challenges include data bias, small and imbalanced datasets, data acquisition difficulties, and data

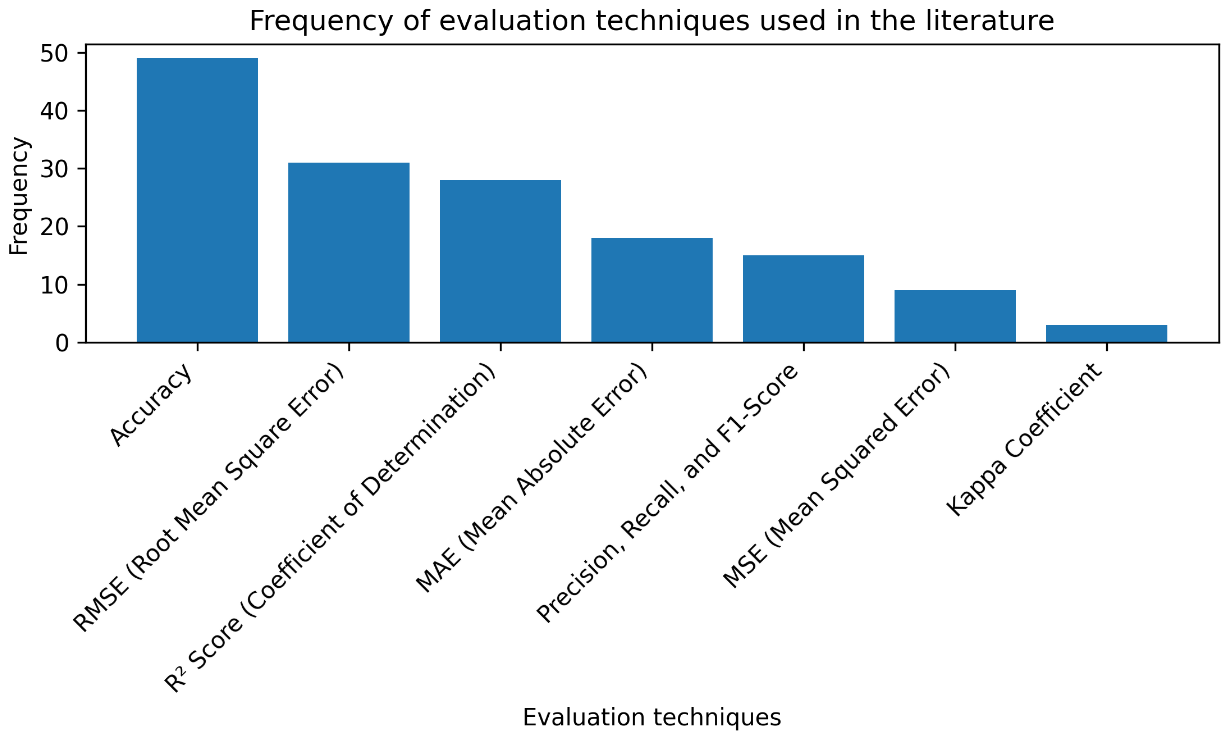

Figure 7: Frequency of evaluation techniques used in the literature.

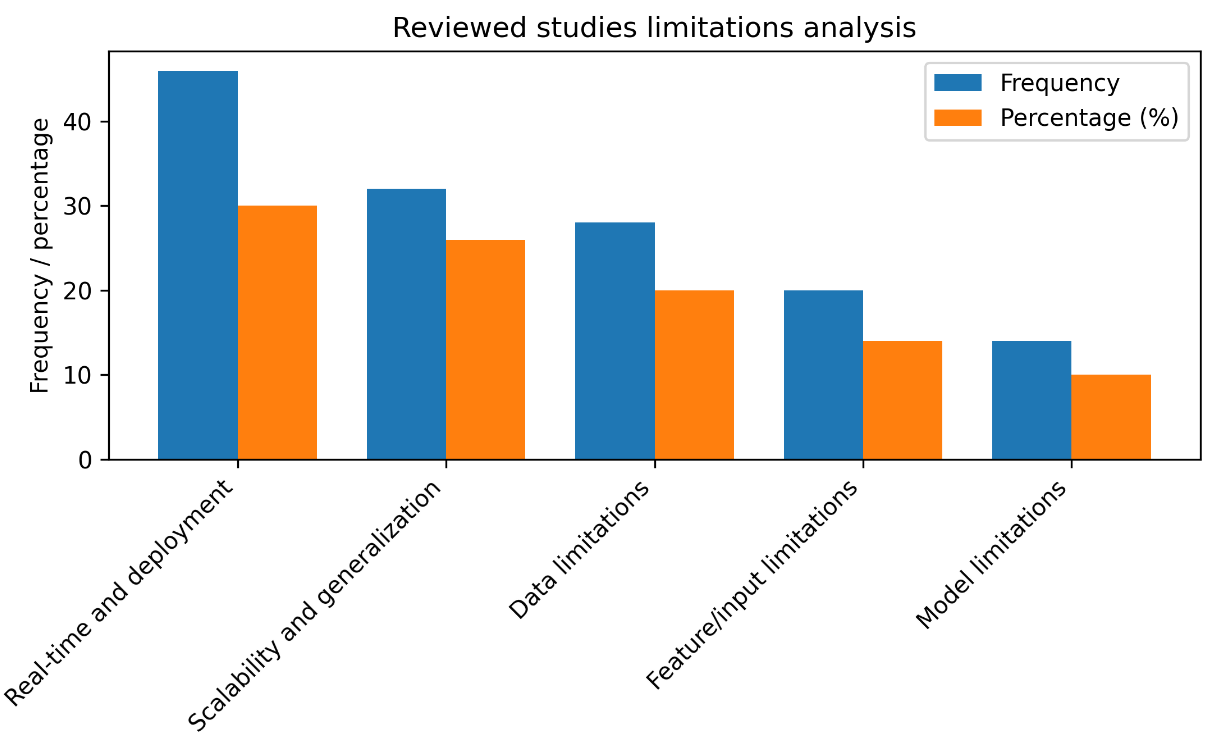

Figure 8: Reviewed studies limitations analysis.

security concerns. Systems are often trained on clean public benchmarks such as Kaggle datasets, which may not reflect noisy farm realities. Some datasets contain as few as 2,200 entries, increasing the risk of overfitting and reducing the ability to handle uncertainty over time, especially during years of extreme weather. Collection of primary data is difficult, and privacy and security are often insufficiently addressed.

Most models also rely on historical data that are not standardized to the degree of accuracy needed for modern precision farming.

Technical and algorithmic constraints persist even though many datasets report accuracy above 99%. Complex models, including neural networks and ensembles, are not readily explainable. When a model cannot explain why a particular crop is recommended, farmer confidence can decrease. Powerful architectures such as hybrid CNN–LSTM models or complex ensembles also require substantial computing power and training time and may not run effectively on ordinary mobile devices. In addition, many models are geographically restricted. Models trained in one region may lose accuracy when tested in new regions or climatic zones unless retrained.

Feature-scope limitations also occur because existing systems usually fail to incorporate crucial elements of the farmer's decision context. Many models ignore production costs, profit needs, and market prices. They often consider soil type and weather conditions but do not directly model profitability. Most systems also assume static conditions based on long-term averages and do not integrate actual IoT data. In some datasets, nutrient levels such as NPK were simulated rather than collected from real-world sensor values, potentially reducing reliability.

Infrastructure and rural barriers affect village-level deployment. The costs of implementing hardware and IoT sensors may be prohibitive for small-scale and marginal farmers. Rural applications are also affected by unreliable power, limited connectivity, and the high cost of servicing complex electronic sensors. In addition, end users may lack sufficient digital skills to use machine learning-based recommendation tools efficiently.

The challenges stated in the selected papers are explicit, but additional unreported challenges may also exist. The main challenges of crop recommender systems include data scarcity and regional specificity; integration of heterogeneous sources such as soil, climate, remote sensing, and socio-economic data; poor generalization across seasons and agro-ecological zones; data imbalance; noisy sensor input; and weather volatility. Limited local infrastructure hinders real-time operations in rural regions, while social factors such as interpretability, farmer trust, and data privacy impede broader adoption.

4.6. RQ6: Overcoming black-box challenges

Modern crop forecast models rely heavily on stacked inputs, including climate data (temperature and rainfall) and soil data (the NPK triplet), as principal biological indicators influencing predicted yield. Newer features such as satellite imagery and remote sensing have supported real-time and context-specific information. However, increasing the dimensionality of multi-source data can make these models black-boxes, reducing the interpretability of predictions even when high technical accuracy is achieved.

Efforts to address the black-box nature of AI have led researchers to explore XAI models such as SHAP and LIME, which help explain the importance of environmental variables, such as nitrogen levels or rainfall, in prediction. Attention is also turning toward intrinsically interpretable algorithms such as decision trees. The if–then structure of decision trees is close to farmers' decision processes and can therefore be explained and understood more easily. Recent developments in glass-box frameworks, such as XAI-CROP and TabNet, aim to build transparency into model architecture.

Adding interpretability techniques introduces challenges related to computational cost and model robustness. Complex frameworks such as SHAP KernelExplainer can be time- and resource-intensive, limiting their use in on-device mobile applications. LIME has been flagged as potentially unstable, and applying explanation layers to a model can sometimes cause a small decrease in maximum accuracy. For these systems to become reliable, future work should move beyond metrics toward trustworthy and interpretable systems.

4.7. Discussion

This review highlights a growing trend in agricultural AI research for crop prediction over the past six years. Traditional models such as SVM, decision trees, and KNN remain relevant, while ensemble methods such as Random Forest and XGBoost dominate because of their superior performance. DL models, including CNNs and LSTMs, are increasingly used, especially for image and time-series data. Most studies rely on soil and climate features, with limited integration of IoT, socio-economic, and genomic data. Although many studies report accuracy above 90%, concerns remain about overfitting, poor generalizability, and deployment challenges. Future work should focus on scalable, interpretable, and real-world-applicable solutions.

5. Future research directions

Although crop recommendation has achieved notable success, several limitations remain in field application and may restrict its effectiveness. Future work should address these obstacles to improve reliability, flexibility, and usability for practical intelligent crop recommendation.

First, future studies should integrate real-time data. Most work depends on static historical data and does not capture changing field conditions. Future studies should integrate real-time IoT sensor data, satellite data, and weather APIs to produce dynamic recommendations in real environments.

Second, future studies should enhance model generalizability. Most models have performed well only within the regions in which they were trained. Future work should examine models in different regions and agro-ecological zones to support robust and realistic crop recommendations.

Third, future work should build explainable AI models. Because most DL models function as black-boxes, farmers cannot always understand how recommendations are produced. Future research should develop more understandable and interpretable crop recommendation systems that explain the reasons for recommendations without reducing prediction performance.

Fourth, future work should support deployment in rural farming environments. Most high-performing models consume significant computing power that many small-scale farmers cannot afford. Future research should develop lightweight and affordable models that can run effectively on portable devices without requiring constant internet access or stable power supply.

Fifth, future work should improve data quality and availability. Several studies rely on small, imbalanced, or regionally bounded datasets. Future studies should create larger and more inclusive agricultural datasets and establish methods for handling noise and missing data to support more dependable recommendation systems.

6. Conclusion

Current research trends in agricultural AI for recommendation systems show that traditional ML and ensemble methods, including Random Forest, XGBoost, SVM, decision trees, and KNN, are the most adopted methods for soil and climate datasets. CNNs and LSTMs are increasingly adopted for image and time-series data. These methods generally use soil and climate variables and sometimes rely on satellite information, although not all systems use every information source or make their inputs transparent. Current models report high accuracy rates, but important challenges remain, including limited availability of regionally representative datasets, inadequate data merging, and limited generalization. Related issues include dependence on technological infrastructure, the need for interpretable recommendations, privacy concerns, and farmer acceptance. These limitations show the need for scalable and interpretable agricultural recommendation systems that can be implemented in real-world settings.

Data availability

The data are available with the corresponding author upon request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Funding

Financial support was received from Confluence University of Science and Technology, Osara, Nigeria.

Acknowledgment

The authors are grateful to the respondents for their participation and input into this study.

References

[1] Y. Kang, M. Ozdogan, X. Zhu, Z. Ye, C. Hain & M. Anderson, "Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest", Environmental Research Letters **15** (2020) 064003. https://doi.org/10.1088/1748-9326/ab7df9.

[2] L. Goel, N. Kumar, S. Jain & Y. Agrawal, "Improving precision agriculture using integrated bio-inspired optimisation models for crop recommendation in Rajasthan, India", Scientific Reports **16** (2026) 6925. https://doi.org/10.1038/s41598-026-37863-7.

[3] S. Poornima, P. Mishra, R. R. Mahishi, D. C. Chaganti, M. Pushpalatha & R. B. Jana, "Optimising crop selection for sustainable agriculture: A compound ensemble approach integrating machine learning and IoT-based sensors", Scientific Reports **16** (2026) 11350. https://doi.org/10.1038/s41598-026-40772-4.

[4] X. Fan & C. Li, "A simulated framework for China's rural revitalisation enabled by big data and smart agriculture under the perspective of new quality productivity", Discover Artificial Intelligence **6** (2026) 25. https://doi.org/10.1007/s44163-025-00714-x.

[5] N. L. P. Swati, S. V. Gupta, N. S. Duddela & L. R. Parvathy, "Agentic AI-driven autonomous decision support system for smart agriculture", Scientific Reports **16** (2026) 9972. https://doi.org/10.1038/s41598-026-39472-w.

[6] M. Jahan & M. Nassiri-Mahallati, "An interdisciplinary machine learning-based approach for predicting corn grain yield under biofertiliser applications", Scientific Reports **16** (2026) 9912. https://doi.org/10.1038/s41598-026-40919-3.

[7] A. Raza, M. A. Shahid, M. Zaman, Y. Miao, Y. Huang, M. Safdar, S. Maqbool & N. E. Muhammad, "Improving wheat yield prediction with multi-source remote sensing data and machine learning in arid regions", Remote Sensing **17** (2025) 774. https://doi.org/10.3390/rs17050774.

[8] T. Kehkashan, M. Abdelhaq, A. S. Al-Shamayleh, M. Hamza, M. A. Khan, S. Z. Alharthi & A. Akhunzada, "Sustainable regional corn yield prediction for the United States through interpretable machine learning approach", Scientific Reports **16** (2026) 13942. https://doi.org/10.1038/s41598-026-43213-4.

[9] C. R. Screpnik, E. Zamudio & L. I. Gimenez, “Artificial Intelligence in Agriculture: A Systematic Review of Crop Yield Prediction and Optimisation”, IEEE Access **13** (2025) 70691. https://doi.org/10.1109/ACCESS.2025.3560631.
[10] R. K. Sharma, J. Kaur, G. Feng, Y. Huang, C. Kumar, Y. Wang, S. Sharma, J. Jenkins & J. Dhillon, “Maize and soybean yield prediction using machine learning methods: a systematic literature review”, Discover Agriculture **3** (2025) 64. https://doi.org/10.1007/s44279-025-00215-6.
[11] P. Muruganantham, S. Wibowo, S. Grandhi, N. H. Samrat & N. Islam, “A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing”, Remote Sensing **14** (2022) 1990. https://doi.org/10.3390/rs14091990.
[12] M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin & N. Khan, “A Comprehensive Review of Crop Yield Prediction Using machine learning Approaches with Special Emphasis on Palm Oil Yield Prediction”, IEEE Access **9** (2021) 63406. https://doi.org/10.1109/ACCESS.2021.3075159.
[13] A. P. Singh, P. Sahu, A. Chug & D. Singh, “A Systematic Literature Review of machine learning Techniques Deployed in Agriculture: A Case Study of Banana Crop”, IEEE Access **10** (2022) 87333. https://doi.org/10.1109/ACCESS.2022.3199926.
[14] A. Gupta, D. Nagda, P. Nikhare & A. Sandbhor, “Smart crop prediction using IoT and machine learning”, International Journal of Engineering Research & Technology (IJERT) **9** (2020) 18. https://doi.org/10.17577/IJERTCONV9IS03005.
[15] A. Gutiérrez, K. Bakthavatchalam, B. Karthik, V. Thiruvengadam, S. Muthal, D. Jose, K. Kotecha & V. Varadarajan, “IoT Framework for Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms”, Technologies **10** (2022) 13. https://www.mdpi.com/2227-7080/10/1/13.
[16] A. K. K. Kouame, G. B. M. Heuvelink & P. S. Bindraban, “Unravelling drivers of maize (Zea mays L.) yield variability in Ghana: A machine learning approach”, Computers and Electronics in Agriculture **237** (2025) 110647. https://doi.org/10.1016/j.compag.2025.110647.
[17] A. K. Pal, B. D. K. Patro & S. Chaube, “Design and implementation of a deep learning framework for automated crop classification and health diagnosis in precision agriculture”, Scientific Reports **16** (2026) 11436. https://doi.org/10.1038/s41598-026-42151-5.
[18] A. K. Soni, P. Singh & A. A. Waoo, “Machine Learning Approaches for Climate-Smart Crop Recommendation”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology **12** (2026) 175. https://doi.org/10.32628/CSEIT2612327.
[19] A. Oikonomidis, C. Catal & A. Kassahun, “Hybrid Deep Learning-based Models for Crop Yield Prediction”, Applied Artificial Intelligence **36** (2022) 2031823. https://doi.org/10.1080/08839514.2022.2031823.
[20] A. Punitha & V. Geetha, “Gorilla troops optimisation with deep learning based crop recommendation and yield prediction”, International Journal of Cognitive Computing in Engineering **5** (2024) 494. https://doi.org/10.1016/j.ijcce.2024.09.006.
[21] A. Suruliandi, G. Mariammal & S. P. Raja, “Crop prediction based on soil and environmental characteristics using feature selection techniques”, Mathematical and Computer Modelling of Dynamical Systems **27** (2021) 117. https://doi.org/10.1080/13873954.2021.1882505.
[22] A. Umera, Z. Khalid, S. Qamar, Z. Iqbal, M. Saeed & M. Yousif, “Smart Crop Recommendation Using Ensemble Machine Learning Models”, Journal of Computing and Biomedical Informatics **9** (2025) 1016. https://doi.org/10.56979/901/2025.
[23] A. Wolanin, G. Mateo-Garciá, G. Camps-Valls, L. Gómez-Chova, M. Meroni, G. Duveiller, Y. Liangzhi & L. Guanter, “Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt”, Environmental Research Letters **15** (2020) 024003. https://doi.org/10.1088/1748-9326/ab68ac.
[24] A. Yaganteeswarudu, S. K. Biswas, V. Aruna & D. Tripathi, “Enhancing Transparency in Smart Farming: Local Explanations for Crop Recommendations Using LIME”, Procedia Computer Science **258** (2025) 1993. https://doi.org/10.1016/j.procs.2025.04.450.
[25] Ayushi & P. K. Buttar, “Satellite Imagery Analysis for Crop Type Segmentation Using U-Net Architecture”, Procedia Computer Science **235** (2024) 3418. https://doi.org/10.1016/j.procs.2024.04.322.
[26] B. Dey, J. Ferdous & R. Ahmed, “machine learning-based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables”, Heliyon **10** (2024) e25112. https://doi.org/10.1016/j.heliyon.2024.e25112.
[27] B. M. Lionel, R. Musabe, O. Gatera & C. Twizere, “A comparative study of machine learning models in predicting crop yield”, Discover Agriculture **3** (2025) 151. https://doi.org/10.1007/s44279-025-00335-z.
[28] B. Panigrahi, K. C. R. Kathala & M. Sujatha, “A machine learning-Based Comparative Approach to Predict the Crop Yield Using Supervised Learning with Regression Models”, Procedia Computer Science **218** (2023) 2684. https://doi.org/10.1016/j.procs.2023.01.241.
[29] B. Peng, K. Guan, W. Zhou, C. Jiang, C. Frankenberg, Y. Sun, L. He & P. Köhler, “Assessing the benefit of satellite-based Solar-Induced Chlorophyll Fluorescence in crop yield prediction”, International Journal of Applied Earth Observation and Geoinformation **90** (2020) 102126. https://doi.org/10.1016/j.jag.2020.102126.
[30] B. S. Anami, N. N. Malvade & S. Palaiah, “Deep learning approach for recognition and classification of yield affecting paddy crop stresses using field images”, Artificial Intelligence in Agriculture **4** (2020) 12. https://doi.org/10.1016/j.aiia.2020.03.001.
[31] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze & I. Kabandana, “Data-Driven Analysing and Machine Learning-Based Crop and Fertiliser Recommendation System for Revolutionising Farming Practices”, Agriculture (Switzerland) **13** (2023) 2141. https://doi.org/10.3390/agriculture13112141.
[32] D. Batool, M. Shahbaz, H. S. Asif, K. Shaukat, T. M. Alam, I. A. Hameed, Z. Ramzan, A. Waheed, H. Aljuaid & S. Luo, “A Hybrid Approach to Tea Crop Yield Prediction Using Simulation Models and Machine Learning”, Plants **11** (2022) 1925. https://doi.org/10.3390/plants11151925.
[33] D. Dahiphale, P. Shinde, K. Patil & V. Dahiphale, “Smart farming: Crop recommendation using machine learning with challenges and future ideas”, Artificial Intelligence and Applications **4** (2026) 323. https://doi.org/10.47852/bonviewAIA52026214.
[34] D. Elavarasan & P. M. D. Vincent, “Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications”, IEEE Access **8** (2020) 86886. https://doi.org/10.1109/ACCESS.2020.2992480.
[35] D. Garg & M. Alam, “An effective crop recommendation method using machine learning techniques”, International Journal of Advanced Technology and Engineering Exploration **10** (2023) 498. https://doi.org/10.19101/IJATEE.2022.10100456.
[36] D. Gosai, C. Raval, R. Nayak, H. Jayswal & A. Patel, “Crop Recommendation System using machine learning”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology **7** (2021) 554. https://doi.org/10.32628/CSEIT2173129.
[37] D. Li, Y. Miao, S. K. Gupta, C. J. Rosen, F. Yuan, C. Wang, L. Wang & Y. Huang, “Improving potato yield prediction by combining cultivar information and uav remote sensing data using machine learning”, Remote Sensing **13** (2021) 3322. https://doi.org/10.3390/rs13163322.
[38] D. Paudel, H. Boogaard, A. de Wit, M. van der Velde, M. Claverie, L. Nisini, S. Janssen, S. Osinga & I. N. Athanasiadis, “machine learning for regional crop yield forecasting in Europe”, Field Crops Research **276** (2022) 108377. https://doi.org/10.1016/j.fcr.2021.108377.
[39] D. Taremwa, E. Ahishakiye, A. Obbo, P. K. Kisozi & F. Kaggwa, “Prediction of maize yield in Uganda using CNN-LSTM architecture on a multimodal climate and remote sensing dataset”, Discover Artificial Intelligence **6** (2026) 164. https://doi.org/10.1007/s44163-026-00855-7.
[40] E. Cheng, B. Zhang, D. Peng, L. Zhong, L. Yu, Y. Liu, C. Xiao, C. Li, X. Li, Y. Chen, H. Ye, H. Wang, R. Yu, J. Hu & S. Yang, “Wheat yield estimation using remote sensing data based on machine learning approaches”, Frontiers in Plant Science **13** (2022) 1090970. https://doi.org/10.3389/fpls.2022.1090970.
[41] E. Elbasi, N. Mostafa, C. Zaki, Z. AlArnaout, A. E. Topcu & L. Saker, “Optimising Agricultural Data Analysis Techniques through AI-Powered Decision-Making Processes”, Applied Sciences (Switzerland) **14** (2024) 8018. https://doi.org/10.3390/app14178018.
[42] E. S. M. El-Kenawy, A. A. Alhussan, N. Khodadadi, S. Mirjalili & M. M. Eid, “Predicting Potato Crop Yield with machine learning and Deep Learning for Sustainable Agriculture”, Potato Research **68** (2025) 759. https://doi.org/10.1007/s11540-024-09753-w.
[43] F. Abbas, H. Afzaal, A. A. Farooque & S. Tang, “Crop yield prediction through proximal sensing and machine learning algorithms”, Agronomy **10** (2020) 1046. https://doi.org/10.3390/AGRONOMY10071046.
[44] F. Imtiaz, A. A. Farooque, G. S. Randhawa, X. Wang, T. J. Esau, S. E. H. Garmdareh & B. Acharya, “Optimising potato yield mapping and prediction:

Integrating satellite-based remote sensing and machine learning for sustainable agriculture", Computers and Electronics in Agriculture **237** (2025) 110636. https://doi.org/10.1016/j.compag.2025.110636.

[45] F. M. Talaat, "Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes", Neural Computing and Applications **35** (2023) 17281. https://doi.org/10.1007/s00521-023-08619-5.

[46] G. Alp & F. Soygazi, "Enhancing Crop Recommendation with SMOTE-Augmented Machine Learning Algorithms", Turkish Journal of Engineering **9** (2025) 612–620. https://www.researchgate.net/publication/396252511_Enhancing_Crop_Recommendation_with_SMOTE-Augmented_Machine_Learning_Algorithms.

[47] G. Rajendiran, J. Rethnaraj, S. Zade, R. Guttula & K. K. Pandey, "UniTriRob: A robust machine learning regression model for predicting lettuce yields in aeroponic vertical farming", Scientific Reports **16** (2026) 15791. https://doi.org/10.1038/s41598-026-44564-8.

[48] G. Singh & S. Sharma, "Enhancing precision agriculture through cloud-based transformative crop recommendation model", Scientific Reports **15** (2025) 9138. https://doi.org/10.1038/s41598-025-93417-3.

[49] G. V. Reddy, M. V. K. Reddy, K. Spandana, Y. Subbarayudu, A. Albawi, R. Chandrashekar, A. Singla & Praveen, "Precision farming practices with data-driven analysis and machine learning-Based Crop and Fertiliser Recommendation System", E3S Web of Conferences **507** (2024) 01078. https://doi.org/10.1051/e3sconf/202450701078.

[50] H. Afzal, M. Amjad, A. Raza, K. Munir, S. G. Villar, L. A. D. Lopez & I. Ashraf, "Incorporating soil information with machine learning for crop recommendation to improve agricultural output", Scientific Reports **15** (2025) 8560. https://doi.org/10.1038/s41598-025-88676-z.

[51] H. C. de C. Filho, O. A. de C. Júnior, O. L. F. de Carvalho, P. P. de Bem, R. dos S. de Moura, A. O. de Albuquerque, C. R. Silva, P. H. G. Ferreira, R. F. Guimarães & R. A. T. Gomes, "Rice crop detection using LSTM, Bi-LSTM, and machine learning models from Sentinel-1 time series", Remote Sensing **12** (2020) 2655. https://doi.org/10.3390/RS12162655.

[52] H. Sahu, P. K. Garg, S. Vijay & A. Dasgupta, "Predictive drivers and transferability of multi-scale machine learning-based crop yield prediction under drought across European and Asian climates", Agricultural Water Management **327** (2026) 110254. https://doi.org/10.1016/j.agwat.2026.110254.

[53] H. Wang, E. Cimen, N. Singh & E. Buckler, "Deep learning for plant genomics and crop improvement", Current Opinion in Plant Biology **54** (2020) 34. https://doi.org/10.1016/j.pbi.2019.12.010.

[54] J. Ansarifar, L. Wang & S. v. Archontoulis, "An interaction regression model for crop yield prediction", Scientific Reports **11** (2021) 97221. https://doi.org/10.1038/s41598-021-97221-7.

[55] J. P. Bharadiya, N. T. Tzenios & M. Reddy, "Predicting Crop Yield Using Deep Learning and Remote Sensing", Journal of Engineering Research and Reports **24** (2023) 29. https://doi.org/10.9734/jerr/2023/v24i12858.

[56] J. Shook, T. Gangopadhyay, L. Wu, B. Ganapathysubramanian, S. Sarkar & A. K. Singh, "Crop yield prediction integrating genotype and weather variables using deep learning", PLoS ONE **16** (2021) e0252402. https://doi.org/10.1371/journal.pone.0252402.

[57] K. Alibabaei, P. D. Gaspar & T. M. Lima, "Crop yield estimation using deep learning based on climate big data and irrigation scheduling", Energies **14** (2021) 3004. https://doi.org/10.3390/en14113004.

[58] K. Janani, R. Alageswaran & R. Amirtharajan, "Dynamic SG-SKRDX hybrid framework for precision weather forecasting and crop suitability in the Cauvery Delta", Scientific Reports **16** (2026) 2042. https://doi.org/10.1038/s41598-025-31717-4.

[59] K. Jhajharia, P. Mathur, S. Jain & S. Nijhawan, "Crop Yield Prediction using machine learning and deep learning techniques", Procedia Computer Science **218** (2023) 406. https://doi.org/10.1016/j.procs.2023.01.023.

[60] K. Kindra, N. G. B. Amma & N. G. N. Amma, "Towards enhancing the performance of crop prediction system for precision agriculture using feature correlation square-based nearest neighbor classifier", Scientific Reports **16** (2026) 8807. https://doi.org/10.1038/s41598-026-35212-2.

[61] K. Nischitha, D. Vishwakarma, Ashwini, N. Mahendra & M. R. Manjuraju, "Crop prediction using machine learning approaches", International Journal of Engineering Research & Technology (IJERT) **9** (2020) IJERTV9IS080029. https://doi.org/10.17577/IJERTV9IS080029.

[62] K. Rangarajan Aravind & P. Raja, "Automated disease classification in (Selected) agricultural crops using transfer learning", Automatika **61** (2020) 260. https://doi.org/10.1080/00051144.2020.1728911.

[63] L. Gong, M. Yu, S. Jiang, V. Cutsuridis & S. Pearson, "Deep learning-based prediction on greenhouse crop yield combined TCN and RNN", Sensors **21** (2021) 4537. https://doi.org/10.3390/s21134537.

[64] L. Stracqualursi, "A Lightweight Exponential-Weighted Ensemble for Crop Recommendation", Journal of Agricultural, Biological, and Environmental Statistics (2025). https://doi.org/10.1007/s13253-025-00694-6.

[65] L. Wang, J. Wang, Z. Liu, J. Zhu & F. Qin, "Evaluation of a deep-learning model for multispectral remote sensing of land use and crop classification", Crop Journal **10** (2022) 1435. https://doi.org/10.1016/j.cj.2022.01.009.

[66] M. A. Razavi, A. P. Nejadhashemi, B. Majidi, H. S. Razavi, J. Kpodo, R. Eeswaran, I. Ciampitti & P. V. V. Prasad, "Enhancing crop yield prediction in Senegal using advanced machine learning techniques and synthetic data", Artificial Intelligence in Agriculture **14** (2024) 99. https://doi.org/10.1016/j.aiia.2024.11.005.

[67] M. Bouni, B. Hssina, K. Douzi & S. Douzi, "Interpretable machine learning Techniques for an Advanced Crop Recommendation Model", Journal of Electrical and Computer Engineering **2024** (2024) 7405217. https://doi.org/10.1155/2024/7405217.

[68] M. Bouni, B. Hssina, K. Douzi & S. Douzi, "Towards an Efficient Recommender Systems in Smart Agriculture: A deep reinforcement learning approach", Procedia Computer Science **203** (2022) 825. https://doi.org/10.1016/j.procs.2022.07.124.

[69] M. Herrero-Huerta, P. Rodriguez-Gonzalvez & K. M. Rainey, "Yield prediction by machine learning from UAS-based multi-sensor data fusion in soybean", Plant Methods **16** (2020) 78. https://doi.org/10.1186/s13007-020-00620-6.

[70] M. K. Senapaty, A. Ray & N. Padhy, "A Decision Support System for Crop Recommendation Using Machine Learning Classification Algorithms", Agriculture (Switzerland) **14** (2024) 1256. https://doi.org/10.3390/agriculture14081256.

[71] M. K. Senapaty, A. Ray & N. Padhy, "IoT-Enabled Soil Nutrient Analysing and Crop Recommendation Model for Precision Agriculture", Computers **12** (2023) 61. https://doi.org/10.3390/computers12030061.

[72] M. Kuradusenge, E. Hitimana, D. Hanyurwimfura, P. Rukundo, K. Mtonga, A. Mukasine, C. Uwitonze, J. Ngabonziza & A. Uwamahoro, "Crop Yield Prediction Using machine learning Models: Case of Irish Potato and Maize", Agriculture (Switzerland) **13** (2023) 225. https://doi.org/10.3390/agriculture13010225.

[73] M. Maimaitijiang, V. Sagan, P. Sidike, A. M. Daloye, H. Erkbol & F. B. Fritschi, "Crop monitoring using satellite/UAV data fusion and machine learning", Remote Sensing **12** (2020) 1357. https://doi.org/10.3390/RS12091357.

[74] M. Marshall, M. Belgiu, M. Boschetti, M. Pepe, A. Stein & A. Nelson, "Field-level crop yield estimation with PRISMA and Sentinel-2", ISPRS Journal of Photogrammetry and Remote Sensing **187** (2022) 191. https://doi.org/10.1016/j.isprsjprs.2022.03.008.

[75] M. N. Navidi, E. Fazli, R. Kharazmi & J. Seyedmohammadi, "Wheat yield prediction using integrated optical and radar remote sensing with machine learning across key phenological stages", Scientific Reports **16** (2026) 10470. https://doi.org/10.1038/s41598-026-41501-7.

[76] M. O. Abdullahi, A. D. Jimale, Y. A. Ahmed & A. Y. Nageye, "Revolutionising Somali agriculture: harnessing machine learning and IoT for optimal crop recommendations", Discover Applied Sciences **6** (2024) 77. https://doi.org/10.1007/s42452-024-05739-y.

[77] F. Zhang, S. Liang, H. Ma, W. Li, Y. Chen, T. He, F. Tian, J. Xu, H. Fang, H. Liang, Y. Ma, A. Jia & Y. Zhang, "A review of crop yield estimation on pixel and field scales from remotely sensed data", in *Science of Remote Sensing* (Vol. 13), Elsevier B.V., 2026. https://doi.org/10.1016/j.srs.2025.100342.

[78] M. R. Islam, K. Oliullah, M. M. Kabir, M. Alom & M. F. Mridha, "machine learning enabled IoT system for soil nutrients monitoring and crop recommenda-

tion", Journal of Agriculture and Food Research **14** (2023) 100880. https://doi.org/10.1016/j.jafr.2023.100880.

[79] M. S. Naik, P. Raguraman, P. Ramisetty, K. Leha & M. S. Gayathri, *Crop yield prediction and fertiliser recommendation using machine learning and remote sensing devices*, in *Challenges in Information, Communication and Computing Technology* 1st ed. CRC Press (2024). https://doi.org/10.1201/9781003559092-50.

[80] M. S. Rao, A. Singh, N. V. S. Reddy & D. U. Acharya, "Crop prediction using machine learning", Journal of Physics: Conference Series **2161** (2022) 012033. https://doi.org/10.1088/1742-6596/2161/1/012033.

[81] M. Shahhosseini, G. Hu & S. v. Archontoulis, "Forecasting Corn Yield With machine learning Ensembles", Frontiers in Plant Science **11** (2020) 01120. https://doi.org/10.3389/fpls.2020.01120.

[82] M. Shahhosseini, G. Hu, I. Huber & S. v. Archontoulis, "Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt", Scientific Reports **11** (2021) 1606. https://doi.org/10.1038/s41598-020-80820-1.

[83] M. Snehalatha & A. Patil, "AgriYieldAI an AI driven decision support system for crop yield prediction using YieldFusionNet", Discover Computing **29** (2026) 222. https://doi.org/10.1007/s10791-026-10096-y.

[84] M. T. Ahmed, M. S. Hossen, F. R. Annasha, M. S. Moni & M. A. A. Mridha, "Crop Recommendation System Using Machine Learning Classifiers", UCICS (2025). https://ucics.org/.

[85] M. Tamilselvi, S. Vishnupriya, K. Ushanandhini & A. Dhanamathi, "A Robust Lemuria Framework for efficient crop prediction", Scientific Reports **16** (2026) 9615. https://doi.org/10.1038/s41598-025-33811-z.

[86] M. Y. Shams, S. A. Gamel & F. M. Talaat, "Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making", Neural Computing and Applications **36** (2024) 5695. https://doi.org/10.1007/s00521-023-09391-2.

[87] N. Bali & A. Singla, "Deep Learning Based Wheat Crop Yield Prediction Model in Punjab Region of North India", Applied Artificial Intelligence **35** (2021) 1304. https://doi.org/10.1080/08839514.2021.1976091.

[88] N. Emmanuel, N. Gabriel, R. Gad, D. Bwiza, J. K. Manzi & C. Kennet, "Crop Recommendation Model Using Machine Learning", International Journal of Life Science and Agriculture Research **04** (2025) 707. https://doi.org/10.55677/ijlsar/V04I11Y2025-11.

[89] N. Farmonov, K. Amankulova, J. Szatmari, A. Sharifi, D. Abbasi-Moghadam, S. M. M. Nejad & L. Mucsi, "Crop Type Classification by DESIS Hyperspectral Imagery and Machine Learning Algorithms", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **16** (2023) 1576. https://doi.org/10.1109/JSTARS.2023.3239756.

[90] N. N. Thilakarathne, M. S. A. Bakar, P. E. Abas & H. Yassin, "A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming", Sensors **22** (2022) 6299. https://doi.org/10.3390/s22166299.

[91] N. R. Sawant, A. Kumar, S. Pant & K. Kotecha, "An IoT-driven machine learning system for real-time smart crop recommendation and optimisation in precision agriculture", Discover Artificial Intelligence **6** (2026) 194. https://doi.org/10.1007/s44163-026-00896-y.

[92] P. Chaudhary, P. Gulia, N. S. Gill, N. Alduaiji, P. K. Shukla & S. Vishnoi, "An evaluation of machine learning for soil analysis in internet of things-enabled smart farming", Scientific Reports **16** (2026) 10318. https://doi.org/10.1038/s41598-026-36017-z.

[93] P. D. Reddy, K. S. S. Reddy, P. Jayanth, B. P. Kakarla & R. M. Balakrishnan, "Agri Assist: An AI Integrated Farmer Assistant", Procedia Computer Science **258** (2025) 3510. https://doi.org/10.1016/j.procs.2025.04.607.

[94] P. Nevavuori, N. Narra, P. Linna & T. Lipping, "Crop yield prediction using multitemporal UAV data and spatio-temporal deep learning models", Remote Sensing **12** (2020) 4000. https://doi.org/10.3390/rs12234000.

[95] P. Sharma, P. Dadheech, N. Aneja & S. Aneja, "Predicting Agriculture Yields Based on Machine Learning Using Regression and Deep Learning", IEEE Access **11** (2023) 111255. https://doi.org/10.1109/ACCESS.2023.3321861.

[96] Q. Li, X. Liang, M. Gao, Y. Li, I. R. Willett, H. Suter, A. Pang, M. Zhang, P. Xangsayasane, Z. Wang & D. Chen, "A hybrid machine-learning and process-based regional crop modelling framework to improve rice production strategies under data constraints", Computers and Electronics in Agriculture **249** (2026) 111885. https://doi.org/10.1016/j.compag.2026.111885.

[97] R. A. Ahmed, W. El-Shafai, Z. A. Ahmed, E.-S. M. El-Rabaie & F. E. Abd El-Samie, "High-precision crop recommendation system with stacking ensemble classifiers for optimising agricultural productivity", Scientific Reports **15** (2025) 43377. https://doi.org/10.1038/s41598-025-09640-5.

[98] R. Akhter & S. A. Sofi, "Precision agriculture using IoT data analytics and machine learning", Journal of King Saud University - Computer and Information Sciences **34** (2022) 5602. https://doi.org/10.1016/j.jksuci.2021.05.013.

[99] R. Aworka, L. S. Cedric, W. Y. H. Adoni, J. T. Zoueu, F. K. Mutombo, C. L. M. Kimpolo, T. Nahhal & M. Krichen, "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries", Smart Agricultural Technology **2** (2022) 100048. https://doi.org/10.1016/j.atech.2022.100048.

[100] R. K. Dhanaraj, M. Maragatharajan, A. Sureshkumar & S. P. Balakannan, "On-device AI for climate-resilient farming with intelligent crop yield prediction using lightweight models on smart agricultural devices", Scientific Reports **15** (2025) 31195. https://doi.org/10.1038/s41598-025-16014-4.

[101] S. A. Mohamed, O. O. A. Maksoud, A. Fathy, A. S. Mohamed, K. Hosny, H. M. Keshk & S. A. Mohamed, "A hybrid deep learning and rule-based model for smart weather forecasting and crop recommendation using satellite imagery", Scientific Reports **15** (2025) 36102. https://doi.org/10.1038/s41598-025-21506-4.

[102] S. Agarwal & S. Tarar, "A hybrid approach for crop yield prediction using machine learning and deep learning algorithms", Journal of Physics: Conference Series **1714** (2021) 012012. https://doi.org/10.1088/1742-6596/1714/1/012012.

[103] S. C. Ramu, D. Raman, K. Ramana, A. V. Krishna, A. A. Khan, S. Khan & S. M. Abdu, "A lightweight deep learning and whale optimisation framework for sustainable precision agriculture", Discover Computing **29** (2026) 73. https://doi.org/10.1007/s10791-026-09952-8.

[104] S. Chor, B. Dandekar, S. Prerna, S. Shelke & S. Divekar, "IOT-based crop recommendation using machine learning", International Journal of Scientific Research in Science, Engineering and Technology **12** (2025) 272. https://ijsrset.com/home/article/view/IJSRSET2512307.

[105] S. Das & M. Ambika, "Revolutionising Indian Farming: machine learning-Powered NLP for Optimal Crop Recommendations", Procedia Computer Science **258** (2025) 2040. https://doi.org/10.1016/j.procs.2025.04.454.

[106] S. Fathololoumi, A. Gillespie & A. Biswas, "Enhancing crop yield predictions through spectral-temporal integration across multiple agricultural fields", Computers and Electronics in Agriculture **245** (2026) 111552. https://doi.org/10.1016/j.compag.2026.111552.

[107] S. Gupta, A. Geetha, K. S. Sankaran, A. S. Zamani, M. Ritonga, R. Raj, S. Ray & H. S. Mohammed, "machine learning-and Feature Selection-Enabled Framework for Accurate Crop Yield Prediction", Journal of Food Quality **2022** (2022) 6293985. https://doi.org/10.1155/2022/6293985.

[108] S. Jeong, J. Ko & J. M. Yeom, "Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea", Science of the Total Environment **802** (2022) 149726. https://doi.org/10.1016/j.scitotenv.2021.149726.

[109] S. K. S. Durai & M. D. Shamili, "Smart farming using machine learning and Deep Learning techniques", Decision Analytics Journal **3** (2022) 100041. https://doi.org/10.1016/j.dajour.2022.100041.

[110] S. Khaki, L. Wang & S. v. Archontoulis, "A CNN-RNN Framework for Crop Yield Prediction", Frontiers in Plant Science **10** (2020) 1750. https://doi.org/10.3389/fpls.2019.01750.

[111] S. Koparde, A. Behare, S. Kasare, J. Patil & K. Nadar, *Crop Yield Prediction for Cereals using machine learning*, in *Advancements in Communication and Systems*, Soft Computing Research Society (2024) pp. 407-418. https://doi.org/10.56155/978-81-955020-7-3-36.

[112] S. M. M. Nejad, D. Abbasi-Moghadam, A. Sharifi, N. Farmonov, K. Amankulova & M. Laszlz, "Multispectral Crop Yield Prediction Using 3D-Convolutional Neural Networks and Attention Convolutional LSTM Approaches", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **16** (2023) 254. https://doi.org/10.1109/JSTARS.2022.3223423.
[113] S. Murgod, T. Kabbur, B. Matte, V. Mujumdar & M. M. Raikar, "IoT-Driven Smart Farming with machine learning for Sustainable Food Systems", Procedia Computer Science **260** (2025) 552. https://doi.org/10.1016/j.procs.2025.03.233.
[114] S. P. Raja, B. Sawicka, Z. Stamenkovic & G. Mariammal, "Crop prediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers", IEEE Access **10** (2022) 23625. https://doi.org/10.1109/ACCESS.2022.3154350.
[115] S. Ramzan, Y. Y. Ghadi, H. Aljuaid, A. Mahmood & B. Ali, "An Ingenious IoT Based Crop Prediction System Using ML and EL", Computers, Materials and Continua **79** (2024) 183. https://doi.org/10.32604/cmc.2024.047603.
[116] S. Rani, A. K. Mishra, A. Kataria, S. Mallik & H. Qin, "machine learning-based optimal crop selection system in smart agriculture", Scientific Reports **13** (2023) 15997. https://doi.org/10.1038/s41598-023-42356-y.
[117] S. S. Rajest, S. S. Priscila, R. Regin T. Shynu & R. Steffi, "Application of machine learning to the Process of Crop Selection Based on Land Dataset", International Journal on Orange Technology **5** (2023) 1. https://journals.researchparks.org/index.php/IJOT.
[118] S. Sambare, R. Patil & T. Sambare, "Investigation and classification of soil using distance method and Crop Prediction", Procedia Computer Science **259** (2025) 473. https://doi.org/10.1016/j.procs.2025.03.349.
[119] S. Solanki & S. Verma, "DFI-ADR: Fuzzy Logic-Driven Information Retrieval and machine learning for Environmental and Crop Prediction to Optimise Farming Decisions", International Journal of Information Engineering and Electronic Business **17** (2025) 48. https://doi.org/10.5815/ijieeb.2025.06.04.
[120] S. Sukriadi, A. Adawiah & A. M. N. Afdhal, "Integrating Big Data and machine learning to support smart village decisions for agricultural productivity improvement", Discover Global Society **4** (2026) 33. https://doi.org/10.1007/s44282-026-00360-3.
[121] S. Sukriadi, A. Adawiah & I. Ismail, "Leveraging machine learning for sustainable smart farming in rural landscapes", Discover Food **6** (2026) 168. https://doi.org/10.1007/s44187-026-00888-y.
[122] S. V. Joshua, A. S. M. Priyadharson, R. Kannadasan, A. A. Khan, W. Lawanont, F. A. Khan, A. U. Rehman & M. J. Ali, "Crop Yield Prediction Using Machine Learning Approaches on a Wide Spectrum", Computers, Materials and Continua **72** (2022) 5663–5679. https://doi.org/10.32604/cmc.2022.027178.
[123] T. R. Mahesh & G. S. Madhuri, "Prediction of Crop Yield Based on Soil Moisture using machine learning Algorithms", International Journal of Information Technology, Research and Applications **2** (2023) 33. https://doi.org/10.59461/ijitra.v2i1.30.
[124] T. Zhang, Q. Yang, X. Tong, L. Hu & J. Shao, "DualSpinNet: A crop yield prediction model based on LSTM and GRU", Cognitive Robotics **6** (2026) 32. https://doi.org/10.1016/j.cogr.2025.12.001.
[125] U. Bhimavarapu, G. Battineni & N. Chintalapudi, "Improved Optimisation Algorithm in LSTM to Predict Crop Yield", Computers **12** (2023) 10. https://doi.org/10.3390/computers12010010.
[126] V. Pandith, H. Kour, S. Singh, J. Manhas & V. Sharma, "Performance Evaluation of machine learning Techniques for Mustard Crop Yield Prediction from Soil Analysis", Journal of Scientific Research **64** (2020) 394. https://doi.org/10.37398/jsr.2020.640254.
[127] W. Xu, P. Chen, Y. Zhan, S. Chen, L. Zhang & Y. Lan, "Cotton yield estimation model based on machine learning using time series UAV remote sensing data", International Journal of Applied Earth Observation and Geoinformation **104** (2021) 102511. https://doi.org/10.1016/j.jag.2021.102511.
[128] W. Zhou, Y. Liu, S. T. Ata-Ul-Karim, Q. Ge, X. Li & J. Xiao, "Integrating climate and satellite remote sensing data for predicting county-level wheat yield in China using machine learning methods", International Journal of Applied Earth Observation and Geoinformation **111** (2022) 102861. https://doi.org/10.1016/j.jag.2022.102861.
[129] Y. A. Mari, F. Talpur, R. Kumar, H. Shafi, A. H. Soomro & S. H. Talpur, "Crop Recommendation System Using Machine Learning", Research Journal for Social Affairs **4** (2026) 149. https://doi.org/10.71317/RJSA.004.02.0779.
[130] Y. Guo, Y. Fu, F. Hao, X. Zhang, W. Wu, X. Jin, C. R. Bryant & J. Senthilnath, "Integrated phenology and climate in rice yields prediction using machine learning methods", Ecological Indicators **120** (2021) 106935. https://doi.org/10.1016/j.ecolind.2020.106935.
[131] Y. Mahale, N. Khan, K. Kulkarni, S. A. Wagle, P. Pareek, K. Kotecha, T. Choudhury & A. Sharma, "Crop recommendation and forecasting system for Maharashtra using machine learning with LSTM: a novel expectation-maximisation technique", Discover Sustainability **5** (2024) 134. https://doi.org/10.1007/s43621-024-00292-5.
[132] Z. Sun, L. Di, H. Fang & A. Burgess, "Deep learning classification for crop types in North Dakota", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **13** (2020) 2200. https://doi.org/10.1109/JSTARS.2020.2990104.

Appendix A. Summary of the reviewed work

Table A.1 begins the summary of the reviewed work. The summary is split into multiple standard tables for ASR production, preserving the original columns: serial number, journal database, author and year, algorithms used, input parameters, performance metrics, and limitations. The bibliography includes all sources supplied with the manuscript; the detailed reviewed-work summary below reflects the authors' extracted data. For completeness, the retained source set is cited here in production order: [1–20], [21–40], [41–60], [61–80], [81–91, 93–100], [101, 103–120], [121**? ?** –132].

Table A.1: Summary of the reviewed work (part 1).

	Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
1	Frontiers	Khaki *et al*., (2020)	CNN-RNN, LASSO, Deep Feedforward Neural Network (DFNN), RF	Soil conditions, weather components	RMSE (8–9% of average yield), Correlation Coefficient	LASSO models failed to capture the nonlinear effects of environmental components.
2	Frontiers	Shahhosseini *et al*. (2020)	Optimized weighted ensemble, LASSO, XGBoost, LightGBM, RF	In-season weather data	RMSE (1,138 kg/ha), RRMSE (9.5%), MBE, MDA	Cross-validation may not be an appropriate estimator of true error for non-IID yield datasets.
	Frontiers	Cheng *et al*., (2022)	LSTM, RF, GBDT, SVR	Remote sensing data (EVI, SR, NDWI, REP)	MAE, RMSE (0.219 t/ha), R^2 (0.93)	Hyperspectral band selection is computationally intensive with unclear physical meanings.

Table A.2: Summary of the reviewed work (part 2).

	Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
	Google Scholar	Mari *et al*. (2026)	Decision Tree, SVM, Logistic Regression, Random Forest, Naive Bayes	N, P, K, pH, temperature, humidity, rainfall	Random Forest Classifier achieved 99.725% accuracy	Selection errors lead to low yield and income
	Google Scholar	Chor *et al*. (2025)	Machine Learning via IoT Cloud Engine	Real-time IoT sensor data (environmental)	Near real-time updated crop recommendations	Needs better scaling and weather integration
	Google Scholar	Dahiphale *et al*., (2025)	LR, DT, RF, KNN, Naïve Bayes, SVM, 4-layer Neural Network,	Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH, Rainfall	97.73% validation accuracy (Neural Network); 99.5% accuracy (Random Forest),	Lacks real-time inputs; ignores economic factors; low interpretability of Neural Networks

Table A.3: Summary of the reviewed work (part 3).

	Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
	Google Scholar	Soni *et al*. (2026)	Random Forest, SVM, XGBoost	Soil characteristics and multi-source weather data	Hybrid models exceeded 90–95% accuracy	Lack of real-time data integration and model interpretability
	Google Scholar	Emmanuel N. *et al*. (2025)	Random Forest, XGBoost	Rwanda-specific soil, climate, and geographic data	Contextually relevant recommendations for Rwanda	External source reliability varies
	Google Scholar	Garg & Alam, (2023)	Wrapper-PART-Grid, MLP, IBk, CDT, REP tree	N, P, K, Temperature, Humidity, pH, Rainfall	Accuracy (99.31%), Kappa, Precision, Recall, MAE, Time	Study was limited to only five specific machine learning models.

Table A.4: Summary of the reviewed work (part 4).

	Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
	Google Scholar	Gosai *et al*., (2021)	Naïve Bayes, SVM, Logistic Regression, RF, XGBoost	N, P, K, pH, Temperature, Humidity, Rainfall	Accuracy (ranging from 10% to 95%)	Reliance on historical data precision and standardization challenges.
	Google Scholar	Solanki & Verma, (2025)	DFI-ADR (Fuzzy Logic Information Retrieval)	Price, season, humidity, soil type, irrigation method	DFI-ADR reached 95% accuracy and 100% precision for similar cases	Relies heavily on historical similarity and complex vector embeddings
	Google Scholar	Mahesh T R & Sindhu Madhuri G (2023)	Decision Tree (DT), ANN, Linear Regression, K-NN	Soil moisture (N, P, K, pH)	Accuracy and Error	High time and space complexity for K-NN; needs more parameters like weather forecasts.

Table A.5: Summary of the reviewed work (part 5).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Google Scholar	Gupta *et al*., (2021)	Decision Tree, KNN, SVM	Location, soil type, rainfall, live temp/humidity, pH,	Decision Tree Accuracy: 91.037%	Market requirements ignored; more specific sensors (moisture, pH) needed for accuracy
IEEE	Elavarasan & Vincent. (2020)	DRQN (Deep Recurrent Q-Network), GB, RF, ANN	Climatic conditions, soil macronutrients, groundwater properties, fertilizer volume	Proposed DRQN model reached 93.7% accuracy	Fabrication is tedious due to dynamic non-linear behavior; requires massive data
IEEE	Nagender Aneja, (2023)	Random Forest (RF), XGBoost, CNN, LSTM	Rainfall, meteorological conditions, area, production, yield	RF achieved 98.96% accuracy; CNN had minimum loss of 0.00060	Machine learning algorithms act as "black boxes" with low interpretability

Table A.6: Summary of the reviewed work (part 6).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
IEEE	S. P. Raja, (2021)	KNN, DT, NB, SVM, RF, Bagging	Potato tubers yield, dry matter, starch, temp, soil moisture	RF with feature selection achieved 87.43% accuracy	Irrelevant extra features increase time and space complexity
Google Scholar	Yanghui Kang *et al*. (2020)	XGBoost, LASSO, SVR, RF, LSTM, CNN	Daily air temp, precip, soil moisture, ET, harvest area, yield mean	R^2, RMSE	CNN architecture does not fit tabular data well; LSTM slightly better due to sequential features.
Google Scholar	Aleksandra Wolanin *et al*. (2020)	CNN, Ridge Regression, Random Forest	Satellite observations (NDVI, NDWI, NIRv), Meteorological data, District info	R^2, RMSE	Performance and variable ranking were inconsistent across different years.

Table A.7: Summary of the reviewed work (part 7).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
MDPI	Gong *et al*., (2021)	LSTM-RNN, TCN, TCN-RNN hybrid, RF, SVR, DT, GBR,	Historical yield, CO2, temperature, HD, relative humidity, radiation,	TCN-RNN hybrid: achieved lowest RMSE across all datasets	Adding layers/blocks to process long sequences risks overfitting
MDPI	Thilakarathne *et al*., (2022)	KNN, DT, RF, XGBoost, SVM,	N, P, K, air temperature, humidity, soil pH, rainfall,	RF: 97.18% Accuracy, 97% F1-Score	Small dataset (2200 entries); geographically limited to Indian climatic conditions
MDPI	Elbasi *et al*. (2024)	Decision Tree, RF, BN, NBC, MP, HT, LWL	N, P, K, pH, rainfall, humidity, temperature	DT (89.38%) and RF (87.61%) optimised data analysis	High variability of environmental conditions and imbalanced data

Table A.8: Summary of the reviewed work (part 8).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
MDPI	Senapaty *et al*. (2023)	MSVM-DAG-FFO (Hybrid Multi-class SVM)	Real-time N, P, K, moisture, pH, temperature	97.3% accuracy achieved	High complexity in mineral inter-dependency relationships
MDPI	Alibabaei *et al*., (2021)	BLSTM, LSTM, GRU, CNN, MLP, RF	Big climate data (Temp, Humidity, Irradiance, Wind, Precipitation), Irrigation	MSE (best 0.017), R^2 (best 0.99)	DL models require long training times and large datasets; removing cell state (GRU) reduces performance.
MDPI	Suruliandi *et al*., (2021)	Recursive Feature Elimination (RFE), Boruta, SFFS with Bagging	12 soil attributes, 4 environmental characteristics	Accuracy (94.73%), F1 Score, MAE, Log Loss, Kappa	Feature selection performance is heavily dependent on data splitting ranges.

Table A.9: Summary of the reviewed work (part 9).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
MDPI	Bali & Singla, (2021)	RNN, Random Forest (RF), Multiple Linear Regression (MLR)	Historical yield and environmental data	RMSE (540.88 kg/ha), MAE, MSE	Research relied on confidential data, limiting third-party verification.
MDPI	Bhimavarapu. *et al*. (2023)	IOFLSTM (New optimizer IOF + LSTM), Adam, SGD, RMSprop	Temperature and production datasets	RMSE, MSE, R^2 score	Some state-of-the-art models like GRU performed worse than basic LSTM in yield tasks.
MDPI	Agriculture Review, 2024	SMOTE with LR, DT, KNN, SVC, RF, XGBoost, CatBoost	Soil mineral dataset	Accuracy, Recall, Precision, ROC AUC	Challenges persist with imbalanced datasets and reliance on sensors/drones.

Table A.10: Summary of the reviewed work (part 10).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
MDPI	Agronomy Study, 2020	Support Vector Regression (SVR), Linear Regression, Elastic Net, KNN	Soil and topographic properties	MAE, RMSE, R^2	Linear methods are inferior to nonlinear ones for agrological data.
MDPI	Dania Batool *et al*. (2022)	XGBoost, SVR, AdaBoost, ARD, Decision Tree, MLP, MLR, RANSAC, SLR	Meteorological factors, observed tea yield plucking samples	MAE, MSE, RMSE	AdaBoost is vulnerable to outliers and noisy data; ARD involves complex explanatory subsets.
MDPI	Yumiao Wang *et al*. (2020)	AdaBoost, OLS, LASSO, SVM, RF, DNN	Satellite imagery, climate data, soil maps, historical yield records	RMSE, R^2, MAE, Global Moran's I	Yields were underestimated in some high-yield counties due to small sample size.

Table A.11: Summary of the reviewed work (part 11).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
MDPI	Maitiniyazi Maimaitijiang *et al*. (2020)	Extreme Learning Machine (ELM), PLSR, SVR, RFR	Satellite (spectral), UAV (structural), AGB, LAI, Nitrogen	R^2, RMSE, RMSE%	PLSR limited by collinearity; SVR less sensitive to sample size but highly parameter-sensitive.
MDPI	Nevavuori *et al*., (2020)	CNN-LSTM, ConvLSTM, 3D-CNN	Multitemporal UAV RGB and weather data (15-week sequences),	3D-CNN: MAE 218.9 kg/ha, MAPE 5.51%	Performance degrades with shorter frame sequences; methods still under development,
MDPI	Li *et al*., (2021)	Random Forest (RFR), SVR, Normalized Vegetation Indices (NVI),	UAV remote sensing (VIs), potato cultivar info, GDDs,	RFR: Rˆ2p 0.75, RMSEp 4.81 t/ha	Needs more detailed weather, soil, landscape, and management variables

Table A.12: Summary of the reviewed work (part 12).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
MDPI	Bakthavatchalam *et al*., (2022)	MLP, JRip, Decision Table,	N, P, K, pH, Temperature, Humidity, Rainfall,	MLP Accuracy: 98.2273%; Weighted Avg. ROC: 1,	High data redundancy; normalisation needed to reduce model build time,
Nature	Dhanaraj *et al*. (2025)	BBO-BDC (Bernoulli Deep Belief Network)	Pesticide, temperature, yield, and rainfall	Accuracy of 88.25%; Sensitivity of 91.19%	Failure in analyzing diverse geographies for yield prediction
Nature	Dhanaraj *et al*. (2025)	Random Forest Classifier, LoRa, ANN, XGBoost	Temp, soil moisture, humidity, area, rainfall	Irrigation prediction accuracy of 90.1%	High initial investment in hardware/infrastructure

Table A.13: Summary of the reviewed work (part 13).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Nature	Mohamed *et al*. (2025)	Hybrid CNN, RNN-LSTM, Rule-based	NDVI, EVI, and meteorological criteria	Suitability classification at 93.5% accuracy	Complex architecture increases computational complexity
Nature	Shahhosseini *et al*., (2021)	LR, LASSO, XGBoost, LightGBM, RF, Stacking Ensembles,	Yield, weather (weekly), soil (180 features), management, APSIM outputs,	Hybrid model: up to 27% RMSE decrease	Simulations leverage full weather data, which adds uncertainty in forecasting modes
Nature	Ansarifar *et al*., (2021)	Interaction Regression, Stepwise, Lasso, RF, XGBoost, NN,	Weather, soil (9 depths), management acreages, genetic trends	Interaction Regression: RRMSE < 8%	Heuristic selection algorithms don't guarantee global optimality; limited interaction kernels

Table A.14: Summary of the reviewed work (part 14).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Nature	Srivastava *et al.*, (2022)	1D-CNN, DNN, XGBoost, RF, KNN, Lasso, Ridge, SVR,	Weekly meteorology, soil moisture, phenology (sowing, flowering, harvest dates),	CNN: 7–14% lower RMSE than best performing baseline	Overfitting in KNN/SVR/RF models; inconsistent production/management input data,
Nature	Rani *et al.*, (2023)	LSTM RNN, Random Forest Classifier, ANN, Decision Tree,	Weather (temp, rainfall), Soil (pH, fertility, type, drainage),	RF Accuracy: 97.235% (crop selection); RNN RMSE (Min Temp): 5.023%,	LSTM context window limit; RF requires high computational power and lacks interpretability
Nature	Dhanaraj *et al.*, (2025)	BBO-BDC, Bernoulli Deep Belief Network	Pesticide, Rainfall, Temperature, Yield	88.25% Accuracy; 27% reduction in convergence speed	Failed to analyze diverse geographies involved in predicting yield

Table A.15: Summary of the reviewed work (part 15).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Nature	Ahmed *et al.*, (2025)	Dual-stacking Ensemble (18 classifiers), Voting Ensemble	N, P, K, T, humidity, pH, rainfall; Pesticide usage, hectare yield	99.54% Accuracy (Small dataset); 85.6% (Large dataset)	Voting ensemble underperforms on large datasets; neglected calculation times
Nature	Dhanaraj *et al.*, (2025)	Lightweight Random Forest (RF)	Temperature, Soil moisture, Humidity, Rainfall	90.1% Accuracy; 0.5W Energy efficiency	Challenges in data security, privacy, and initial infrastructure investment
Nature	Mohamed *et al.*, (2025)	Hybrid CNN, RNN-LSTM, Rule-based model	Multispectral imagery (NDVI, NDWI); meteorological variables	97% Accuracy (CNN); RMSE 0.19 (LSTM)	High computational complexity due to remote sensing processing

Table A.16: Summary of the reviewed work (part 16).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Nature	Venkateswara & Padmanaban, (2025)	TabNet (deep learning), SHAP	Soil color, N, P, K, pH, Rainfall, Temperature	96.21% Accuracy (Crop); 95.24% (Fertilizer)	Requires higher computing power compared to traditional tree-based models
Nature	Tamilselvi *et al.*, (2026)	Robust Lemuria Framework (DBN, RF, Naive Bayes, J48)	Temperature, moisture, rainfall (2010–2020 dataset)	98.99% Accuracy; 99.35% Specificity; R^2 0.9994	Complex implementation; requires parallel processing; black-box nature of ANNs
ScienceDirect	Sambare *et al.* (2025)	Bray-Curtis dissimilarity distance method	Clay, sand, lime, N, P, Potash, pH, humidity	Suggested satisfactory identification regardless of soil type	Accurate primary data acquisition remains a challenge

Table A.17: Summary of the reviewed work (part 17).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Yaganteeswarudu *et al*. (2025)	LIME (XAI), Logistic Regression	N, P, K, pH, temp, humidity, rainfall	Achieved 100% probability for specific crop instances	Complexity for end-users and high computational cost
ScienceDirect	Hou *et al*. (2025)	Linear Regression (LR), Random Forest (RF)	NDVI, EVI, RVI, GNDVI, and phenological parameters	Rˆ2 up to 0.8201; MAE of 0.2878 t/ha	Model model accuracy limited by small sample size
ScienceDirect	Kouame *et al*. (2025)	RF, Gradient Boosting (GB), Cubist (CUB)	Gridded climate, soil, terrain, and management data	MAE (0.63–0.65 t/ha); CCC (0.81–0.82)	Performance variations across different agro-ecological zones

Table A.18: Summary of the reviewed work (part 18).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Maduri Shripathi Rao *et al*. (2022)	Random Forest, Decision Tree, KNN	Soil parameters (N, P, K, pH), Climatic conditions (Temp, Humidity, Rainfall)	Accuracy, Precision, Recall, F1 Score	Limited by farmers repeating same crops and random fertiliser usage.
ScienceDirect	Razavi *et al*. (2024)	RF, XGBoost, CatBoost, LightGBM, VAE	38 spatiotemporal and physiographical features	CatBoost achieved RMSE of 0.294	Limited availability of raw historical data in specific regions
ScienceDirect	Imtiaz *et al*. (2025)	RF Regression, CART, Gradient Tree Boosting (GTB)	Sentinel-2A, PlanetScope, and Vegetation Indices	GTB achieved best accuracy (Rˆ2 0.62–0.78)	Manual harvesting for ground truth is labor intensive

Table A.19: Summary of the reviewed work (part 19).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Murgoda *et al*. (2025)	Decision Trees, Random Forests, LR, SVM, ANN	IoT sensor data (soil moisture, temp, humidity), motion	Random Forest achieved 99.99% accuracy	Data privacy and security parameters not incorporated
ScienceDirect	Bouni *et al*. (2022)	Deep Reinforcement Learning (DRL), NB, KNN	N, P, K, humidity, temperature, rainfall	DRL and Random Forest showed highest accuracy	Structure expansion needed to mitigate overfitting
ScienceDirect	Li *et al*. (2026)	Particle Swarm Optimization (PSO)	Remote sensing and crop model integration	Overall accuracy exceeded 90% in study areas	High complexity in modeling environmental data

Table A.20: Summary of the reviewed work (part 20).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Das *et al*. (2025)	RF, SVM, KNN, DT, NB, NLP (Chatbot)	N, P, K, pH level, temperature, humidity, rainfall	RF Classifier reached 99.88% accuracy	Traditional methods lack dynamic seasonal consideration
ScienceDirect	Zhang *et al*. (2026)	DualSpinNet (LSTM + GRU)	Crop type, area, rainfall, fertilizer/pesticide usage	Rˆ2 of 0.99993; RMSE of 3.7425E-04	Interpretability is not the primary focus of the structure
ScienceDirect	Ayushi *et al*. (2024)	U-Net architecture, FCN, SegNet	Sentinel-2 satellite imagery, NDVI	Accuracy of 95.3%; F1 score of 73.6%	Smallholder farms feature irregular field shapes and cloud cover

Table A.21: Summary of the reviewed work (part 21).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Bouni *et al.* (2022)	DRL, Naïve Bayes, KNN, Decision Tree, Random Forest	N, P, K, humidity, temperature, rainfall	DRL and Random Forest showed best performance	LSTM needed to lessen consequences of overfitting
ScienceDirect	Reddy *et al.* (2025)	Stacked Ensemble (RF + GB), FastText	N, P, K, pH, temperature, humidity, rainfall	99.32% Accuracy; F1-Score 99.26%	Needs multilingual support and mobile app development
ScienceDirect	Punitha & Geetha (2024)	GTODL-CRYPM (GTO + LSTM + DBN)	Rainfall, climate, and fertiliser data	99.88% Accuracy; 99.14% R2 score	Computational complexity of GTO and LSTM algorithms

Table A.22: Summary of the reviewed work (part 22).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Marshall *et al.* (2022)	Vegetation Indices (TBVIs), PLSR, RF	PRISMA/Sentinel-2 satellite spectral data	PRISMA improved variability explanation by 20%	Benefits/limitations depends on imagery availability/frequency
ScienceDirect	Akhter & Sofi (2022)	Survey of ML and IoT Data Analytics	Volume, velocity, and variety of complex IoT data	Verified ZigBee/LoRa as convenient for precision farming	Challenges in application layer and interoperability
ScienceDirect	Razavi *et al.* (2024)	XGBoost, RF, CatBoost, LightGBM, VAE	39 Spatiotemporal/physiographical features	Catboost Regressor RMSE of 0.294	Overcame lack of historical records using synthetic data

Table A.23: Summary of the reviewed work (part 23).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Jhajharia *et al.* (2023)	RF, SVM, Gradient Descent, LSTM, Lasso	Area, production, rainfall, soil type	RF R2 score of 0.963; RMSE 0.035	LSTM requires much larger databases for better accuracy
ScienceDirect	Fathololoumi *et al.* (2026)	Survey (SVM, ANN, RF, Hybrid)	Remote sensing, soil, and weather datasets	Automated ML for fast/robust yield prediction	Variability in climatic data affects precision consistency
ScienceDirect	Kavita & Mathur, (2023)	RF, SVM, Gradient Descent, LSTM, Lasso Regression	Rajasthan crop data, 27 unique soil types, annual rainfall	RF performed best with R^2 of 0.963 and RMSE of 0.035	LSTM requires a larger database than available; high errors in Gradient Descent

Table A.24: Summary of the reviewed work (part 24).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Paudel *et al.*, (2021)	Ridge, KNN, SVR, GBDT	WOFOST simulation outputs, weather, remote sensing, soil data	Early season (30 days) ML forecasts were comparable to operational MCYFS	ML forecasts lag behind operational human-led systems as the season progresses
ScienceDirect	Panigrahi *et al.*, (2023)	DT, RF, GBR, XGBoost, Voting Regression	Min/max temp, annual rainfall, area, production (Bengal Gram, Maize)	RF outperformed others with MAE of 468.16	XGBoost model suffered from overfitting with high R^2 but low cross-validation
ScienceDirect	Paudel *et al.*, (2022)	Ridge, KNN, SVR, GBDT	WOFOST, elevation, slope, field size, irrigation, remote sensing	Optimized ML significantly better than trend models; improved NRMSE in 71% of cases	Substantial room for improvement remains during extreme climate years

Table A.25: Summary of the reviewed work (part 25).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	S. Jeong *et al*., (2022)	FC, 1D-CNN, LSTM, hybrid CNN-LSTM	Spectral indices (8-day), solar radiation, county IDs, transplanting dates	Simulated yields fit reported data with R^2 of 0.805 and NSE of 0.770	Process-based models are limited by a lack of input variables from harvest to prediction
ScienceDirect	Islam *et al*., (2023)	CatBoost, Bagging, Voting, Random Forest	N, P, K, pH, temp, humidity, rainfall, soil moisture	CatBoost achieved 97.5% accuracy and 98% precision	Occasional classification errors can negatively impact fertilization decisions
ScienceDirect	Aworka *et al*., (2022)	CRF (Crop RF), CGBM, CSVM	Pesticides, Nitrous Oxide (N_2O) emissions, climate data	CRF (Crop Random Forest) reached R^2 of 92.272%	Increasing complexity results in high hardware resource consumption (CPU/RAM)

Table A.26: Summary of the reviewed work (part 26).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Hempushpa Sahu *et al*. (2026)	Random Forest (RF)	Climate data (Temp, Humidity, Wind, Rad, Precip), NDVI, LAI, Soil moisture	RMSE, MAPE	Difficulty in ensuring representation of drought years; spatial transferability challenges.
ScienceDirect	Dilli Paudel *et al*. (2023)	LSTM, 1DCNN, GBDT, Linear trend models	Seasonal indicators (biomass, temp, moisture), elevation, yield trends	Regression loss, MAE, RMSE	Small data sizes (300–1950 instances); DL requires much larger datasets for significant gains.
ScienceDirect	Weimo Zhou *et al*. (2022)	RF, SVM, LASSO	Climate data (Temp, Precip, Rad, Wind), Remote sensing (NDVI, EVI, SIF)	R^2, RMSE	Contribution of climate data decreases as crops mature; SIF better for soybean than maize.

Table A.27: Summary of the reviewed work (part 27).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Mengjia Qiao *et al*. (2021)	SSTNN, RNN (LSTM), 3D CNN, 2D CNN, DT, RF, SVM	Multi-spectral & multi-temporal MODIS imagery	RMSE, R^2, MAPE	Traditional ML cannot handle multi-time step images well; spectral information waste in flattening.
ScienceDirect	Weicheng Xu *et al*. (2021)	BP Neural Network, U-Net, MLR	UAV remote sensing (RGB/Multispectral), Vegetation Indices, BOP%	R^2, MSE, Kappa coefficient	Initial input sets contained redundant variables; MLR has poor generalisation ability.
ScienceDirect	Bin Peng *et al*. (2020)	LASSO, RIDGE, SVR, RF, ANN	Satellite-based SIF, climate variables, Vegetation Indices (NDVI, EVI, NIRv, LST)	R^2, RMSE	Relative performance using SIF might not change with more advanced algorithms.

Table A.28: Summary of the reviewed work (part 28).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
ScienceDirect	Lijun Wang *et al*. (2022)	Timm-RegNetY-320, UNet++, Deeplab V3+, RF	Multispectral remote sensing (S-2A, GF PMS), feature bands, patch sizes	OA, F1 score, mIoU, Recall, Precision	No unified standard for feature selection; imbalanced class training issues.
ScienceDirect	Shook *et al*., (2021)	Stacked LSTM, Temporal Attention model, SVR-RBF, LASSO	Weather variables, Maturity Group (MG), Genotype Cluster	RMSE, MAE (5.441 bu/acre), Rˆ2 (0.796)	Limited genotype-specific prediction due to LSTM complexity and lack of genomic data.
Scopus	Umera *et al*. (2025)	Logistic Regression, SVM, Decision Tree, KNN, XGBoost, ANN	N, P, K, humidity, rainfall, pH level	ANN and XGBoost reached 98% accuracy	Traditional mathematical models have low reliability

Table A.29: Summary of the reviewed work (part 29).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Dey *et al*., (2024)	SVM, XGBoost, RF, KNN, DT,	NPK, soil pH, temperature, rainfall, humidity,	XGBoost Precision: 99.09% (Agri), 99.3% (Horti)	Reliant on public dataset; mixed-crop modelling accuracy is lower than that of individual classes,
Scopus	Alp & Soygazi (2025)	TabNet, XGBoost, KNN, Decision Tree, LightGBM, RF, SMOTE	N, P, K, temperature, rainfall, humidity, pH	RF + SMOTE attained 99.81% accuracy	TabNet (deep learning) did not perform well on targeted dataset
Scopus	Raja *et al*. (2022)	RF, KNN, Bagging, NB, Decision Tree, SVM, MRFE/RFE	Felin dataset (soil and environmental factors)	RF with MRFE gave 97.29% accuracy	Requires robust preprocessing to balance datasets

Table A.30: Summary of the reviewed work (part 30).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Joshua *et al*. (2022)	BPNN, SVM, GRNN	Historical rice production data, rainfall	GRNN R2 score of 0.97; MAE 82.74	RMSE of 12% found when considering weather data
Scopus	Ramzan *et al*. (2024)	14 Algorithms (Voting, CatBoost, ExtraTrees, etc.)	Air temp/humidity, rainfall, soil pH, N, P, K	RF and Voting models showed near 100% accuracy	Traditional manual methods remain time-consuming/expensive
Scopus	Pal *et al*. (2026)	CNN, RNN	Satellite imagery, drone-based NIR, IoT (pH, moisture, NPK), weather	Accuracy >90%, Precision, Recall, Inference Time	Previous models lacked real-time IoT integration or broader environmental context.

Table A.31: Summary of the reviewed work (part 31).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Ramu *et al*. (2026)	CNN (spatial), LSTM (temporal), Whale Optimization (WOA)	Crop images, sensor data (Temp, RH, Soil Moisture, pH, N, P, K)	Accuracy (96.73%), Precision (95.80%), Recall (95.21%), F1-score (95.50%)	Low-cost sensors subject to drift; reliance on stable IoT infrastructure.
Scopus	Snehalatha & Patil (2026)	YieldFusionNet (MLP, LSTM, CNN), SHAP	Soil (pH, NPK, moisture, organic carbon), climate data, NDVI/EVI	RMSE (0.431), MAE (0.292), Rˆ2 (0.91)	Data heterogeneity and limited transferability across agro-climatic regions.
Scopus	Janani *et al*. (2025)	Dynamic SKRDX Ensemble, XGBoost, KNN, SVM, DT, RF	Soil (N, P, K, pH) and environmental (Temp, Humidity, Rainfall)	Accuracy (93%), Precision (94%), Recall (94%), F1-Score (94%)	Risk of overfitting with specific window size configurations.

Table A.32: Summary of the reviewed work (part 32).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Kindra *et al*. (2026)	FCSNN (Feature Correlation Square based Nearest Neighbor)	N, P, K, temperature, humidity, pH, rainfall	Accuracy (97.8%), Precision (0.97), Recall (0.98), F1-Score (0.97)	k-NN sensitivity to noise if the k-value is not properly calibrated.
Scopus	Chaudhary *et al*. (2026)	Review: SVM, RF, XGBoost, GBRT, CNN, hybrid CNN-LSTM	(SLR focus) Soil pH, NPK, moisture, texture, EC, temperature	Reported Accuracy: 80% to 97% across various studies	Sensor reliability, environmental noise, and maintenance costs in rural areas.
Scopus	Goel *et al*. (2026)	Hybrid NIO + ANN, Random Forest, SVM, DT, KNN	16 satellite image features (soil N/P), Humidity, Rainfall, Moisture	Accuracy (up to 0.9494), computational efficiency (seconds)	Requires robust IoT sensing infrastructure and high computational costs.

Table A.33: Summary of the reviewed work (part 33).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	L. R. P. *et al*. (2026)	Multi-Agent System (CNN, RF, XGBoost), SHAP, LIME	Soil images, N, P, K, pH, dynamic weather API	Accuracy: Crop Recommendation (92.4%), Fertilizer Suggestion (94.7%)	Multi-task complexity leads to modest reduction in peak accuracy.
Scopus	Poornima *et al*. (2026)	Compound Ensemble (12 weak classifiers), Intensified LSTM	Soil telemetry (moisture, EC, pH, NPK), rainfall forecasts, drought indices	Accuracy (99.8%), Precision (99.7%), Recall (99.7%), F1-score (99.7%)	Validation restricted to 50 crops in one district; contingent on sensor uptime.
Scopus	M. J. *et al*. (2026)	ANFIS, Transformer, ANN, SVR, LightGBM, XGBoost	73 features (Dry matter yield, SPAD, LAI, photosynthesis rate, height)	R^2 (ANFIS: 0.555, Transformer: 0.545), RMSE, MAE	Regression models act as black boxes; prone to overfitting with small datasets.

Table A.34: Summary of the reviewed work (part 34).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Raza *et al*. (2026)	Random Forest, SVM, SVR, CNN	Multi-source remote sensing (Optical/Radar), soil moisture	Predictive accuracy, RMSE	Short data spans and lack of longitudinal uncertainty quantification.
Scopus	Kehkashan & Hamza *et al*. (2026)	RF, Gradient Boosting, Voting Classifier ensemble, SHAP	Futures price, NOAA weather, USDA soil (NPK), satellite NDVI	Accuracy (0.97), Precision (0.92), Recall (0.92), F1-score (0.92)	U.S. Corn Belt conditions specific; requires retraining for international use.
Scopus	G. R. *et al*. (2026)	UniTriRob Robust Regression, Huber, RANSAC, Theil-Sen	Topographic, environmental, and growth-related variables	Adjusted R^2 (UniTriRob: 97.865), RMSE, MAE	Standard regression models are prone to underfitting or overfitting risks.

Table A.35: Summary of the reviewed work (part 35).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Fan & Li (2026)	BiLSTM (Bidirectional LSTM), MLP	N, P, K, pH, temperature, humidity, rainfall	Accuracy (98.88%), Precision (98.00%), Recall (98.22%), F1-score (98.32%)	Previous studies lacked targeted frameworks linking nutrients to environment.
Scopus	Taremwa *et al*. (2026)	Hybrid CNN-LSTM, RF, CNN-RF	Climatic and Vegetation Indices (NDVI, EVI, NDWI, LAI, GPP)	MSE (0.107), explains 78% of yield variation (R^2 = 0.78)	Short data span; need for higher-resolution satellite imagery.
Scopus	Sawant *et al*. (2026)	Enhanced Gaussian Naive Bayes (EGNB), Random Forest, Bagging	Nitrogen, phosphorus, potassium, pH, Temp, Humidity, Rainfall	Accuracy (99.55%), Precision (99.56%), Recall (99.55%), F1-score (99.54%)	NPK values were simulated rather than collected from field nutrient sensors.

Table A.36: Summary of the reviewed work (part 36).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Bhatnagar & Saxena (2026)	EMLF Ensemble (RF, AdaBoost, SVR, XGBoost), Stacking	Agro-climatic (TEMP, RH, VP, EVAP, RFall, SSD, WSpeed)	R^2 (0.9821), MAE (0.0112), MSLE (0.0001)	Heavy reliance on human intervention; data outlier sensitivity.
Scopus	Sukriadi *et al.* (2026)	Review: Supervised Learning (CNN, RF, SVM, LSTM)	(SLR focus) leaf/drone images, time-series weather	Bibliometric: normalized citation impact, publication trends	Scarcity of labeled rural datasets and limited digital literacy.
Scopus	Sukriadi *et al.* (2026)	Bibliometric: VOSviewer, Biblioshiny (ML: RF, SVM, LSTM)	Big Data (soil/crop sensors, RS imagery, climate/market data)	Bibliometric: Co-occurrence, collaboration networks	Limited to Scopus-indexed/Open Access journal articles.

Table A.37: Summary of the reviewed work (part 37).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
Scopus	Oikonomidis *et al.*, 2022	XGBoost, CNN-DNN, CNN-RNN, CNN-LSTM	Weather (precip., temp, radiation), soil (pH, bulk density, organic matter)	Hybrid CNN-DNN achieved R^2 of 0.87 and RMSE of 0.266	Limited model explainability; hard to reproduce past results due to library mismatches
Scopus	Cheng *et al.*, 2022	LSTM, RF, GBDT, SVR	Seeding rate, farming method, fertilizer application, irrigation, multispectral data	LSTM performed best with RMSE of 0.201 t/ha	SVR unable to map highly non-linear/complex relationships compared to DL
SpringerLink	Cartolano *et al.* (2024)	SHAP, LIME, XGB, MLP, SVM	N, P, K, temperature, humidity, pH, rainfall	XGBoost classifier accuracy of 99.32%	Real-world noise may lead to class misclassification

Table A.38: Summary of the reviewed work (part 38).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
SpringerLink	Stracqualursi (2025)	Exponential-Weighted Ensemble	N, P, K, pH, temp, humidity, rainfall	Attained near-ceiling accuracy of 99.8%	Relies on relatively clean benchmarks; lacks market factors
SpringerLink	Talaat, 2023	DT, RF, ExtraTreeRegressor, GNB, Gradient Boosting, SVR, MLR, CYPA,	Pesticides, Yield, Rainfall, Avg. Temperature, Year, Country,	ExtraTreeRegressor Score: 0.9933; CYPA RMSE: 132.01,	Geographical specificity; relies on highly accurate and timely data collection
SpringerLink	Shams *et al.*, 2024	XAI-CROP, Gradient Boosting, DT, RF, GNB, MNB,	Soil type, weather patterns, historical yields, Season, Area, Production,	XAI-CROP: MSE 0.9412, MAE 0.9874, R^2 0.9415,	Base models lack explainability (partially solved by XAI); limited to Indian cultivation dataset,

Table A.39: Summary of the reviewed work (part 39).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
SpringerLink	El-Kenawy *et al.*, 2024	KNN, Gradient Boosting, XGBoost, MLP, GNNs, GRUs, LSTM,	Weather (temp, precip, humidity), soil properties, yield records, agronomic features,	GNN: MSE 0.02363, R^2 0.51719,	Transferability issues to other plants/locations; KNN noise sensitivity,
SpringerLink	Herrero-Huerta *et al.*, 2020	Random Forest (RF), XGBoost,	UAV optical data (MSI), canopy cover, VIs, plant height,	XGBoost accuracy: 91.36%	Requires specific training for different field conditions/genotypes; high sensor cost
SpringerLink	Lionel *et al.*, 2025	Review: RF, DNN, CNN, RNN, SVM	Meteorological parameters and crop yield	Comparative analysis of MAE, RMSE, and R^2	Focused on a limited number of journals; analysis tools used were not exhaustive

Table A.40: Summary of the reviewed work (part 40).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
SpringerLink	Sawant *et al.*, 2026	Enhanced Gaussian Naive Bayes (EGNB)	Soil and Climate (N, P, K, Temp, Humidity, pH, Precipitation)	99.55% Accuracy	Assumes limited inter-feature dependency in input parameters
SpringerLink	M. Omar and A. Dahir, 2024	Decision Tree (DT), KNN, Random Forest	N, P, K, pH, Temperature, Humidity, Rainfall	99.2% Accuracy (DT); standout performer for interpretability	Dataset primarily sourced from a specific Somali region (Balcad); needs broader geographical expansion
SpringerLink	Mahale *et al.*, 2024	LSTM (Weather Forecast) + Tuned RF (Recommendation)	Historical Weather (Rainfall, Temp, Wind speed, Evaporation)	96.53% training accuracy; 92.73% test accuracy	Analysis omitted soil properties, production costs, and market prices

Table A.41: Summary of the reviewed work (part 41).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
SpringerLink	Stracqualursi, 2025	Lightweight Exponential-Weighted Ensemble (KNN, DT, RF, SVM, NB, LR, XGB)	N, P, K, temperature, humidity, pH, rainfall	99.8% Accuracy	Uses clean Kaggle dataset; real environments may produce noisier data
SpringerLink	Cartolano *et al.*, 2024	XGB, MLP, Linear SVM; XAI: SHAP, LIME	N, P, K, temperature, humidity, pH, rainfall	99.32% Accuracy (XGB); 98.64% (SVM)	LIME lacks stability; SHAP KernelExplainer is slow; potential for adversarial attacks
SpringerLink	Amjad & Ashraf, 2025	RFXG (Ensemble of Random Forest and XGBoost)	Soil (N, P, K, pH) and Weather (Temp, Humidity, Rainfall)	98.10% accuracy; outperformed ETC, MLP, RF, DT, LR, and XGB	Complex architecture increases implementation cost; needs testing on highly imbalanced datasets

Table A.42: Summary of the reviewed work (part 42).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
SpringerLink	Singh & Sharma, 2025	TCRM (Transformative Crop Recommendation Model)	Environmental (NPK, Temperature, pH, Rainfall)	94% Accuracy; 97.67% Five-fold cross-validation score	Lacks evaluation metrics for specific energy and resource savings; requires real-world testing in more resource-limited areas
SpringerLink	N.G.B.A. *et al.*, 2026	FCSNN (Fuzzy C-means Clustering Stochastic Neural Network)	N, P, K, Temperature, Humidity, pH, Rainfall	97.9% Accuracy; 2.1% error rate	Scalability to non-Kaggle derived datasets not fully addressed
SpringerLink	A.K.P. *et al.*, 2026	CNN-based Multi-modal Framework	Multi-modal (Satellite/Drone images, soil sensors, weather API)	High efficiency (89% at 50 inputs); precision ranges from 33% to 92% based on input volume	Lack of unified frameworks that fuse multispectral and IoT data streams in existing literature

Table A.43: Summary of the reviewed work (part 43).

Journal Database	Author & Year	Algorithms Used	Input Parameters	Performance Metrics	Limitations
SpringerLink	SN Kumar *et al.*, 2025	IoT Multi-sensor monitoring (Arduino + ESP-01)	pH, NPK, Moisture, Temperature	Validated with < 2% error margin compared to lab results	Field testing was limited to 10 Kerala farms; connectivity and power issues in rural settings
Wiley	Bouni *et al.*, 2024	Random Forest (RF), SVC, GaussianNB, DT, AdaBoost, Gradient Boosting, SHAP, LIME	N, P, K, Temperature, Humidity, pH, Rainfall	Accuracy (RF 99.3%), Precision, Recall, F1 score, AUC	AdaBoost demonstrated potential for overfitting and poor generalization.
Wiley	Gupta *et al.*, 2022	PSO-SVM, KNN, Random Forest, Relief algorithm	Fertilizer (N,P,K), temp, rainfall, solar radiation, soil pH	PSO-SVM achieved the best accuracy; KNN had superior sensitivity/specificity	Input complexity and unpredictability remain challenges for food sector models