ISSN: 3049-3277

International Journal of Integrative Studies (IJIS)

Journal homepage: www.ijis.co.in

RESEARCH ARTICLE

AI-Driven Precision Agriculture for Sustainable Crop Production

[*1]Ms. Nandru Mrudula Deepthi, [2]Namaswini Padhya

[1*]Assistant Professor at Vikas group of institutions, Nunna, vijayawada.

[2]Research Scholar, GIET, Gunupur

ABSTRACT

As a key enabling technology for precision agriculture, Artificial Intelligence (AI) promises the insight to turn the growing trove of available satellite, sensor, and imagery data into field-specific decisions – from crop yield forecasts to plant disease detection to input-resource optimisation. This paper reviews new (2021–2025) peer-reviewed literature on AI-based precision agriculture and presents a novel ensemble machine-learning model – an integrated Random-Forest/XGBoost stacked yield-prediction model with a fine-tuned convolutional-neural-network disease-detection model, tested for both prediction performance and downstream resource-use and productivity impacts. Vegetative-stage NDVI, cumulative growing-season rainfall, and soil nitrogen status were the most important variables for predicting yield, and the proposed yield-prediction ensemble outperformed the baselines (individual RF, XGBoost, support-vector regression, gradient boosting, decision tree, and linear regression), with $R^2 = 0.95$ on hold-out data. The fine-tuned convolutional disease-classification model achieved 99.1% accuracy, between the 99.05% and 99.75% reported by state-of-the-art architectures in standard plant-disease benchmarks. Results from the field and deployment studies reviewed indicated that AI precision-agriculture adoption had the following range of input-use reductions and productivity gains: 22-31% on variable-rate fertilizer (VRF) application, targeted pesticide application, and precision irrigation, while the biggest gains in decreased usage of farm inputs occurred in relation to autonomous machinery fuel and labor. The results show that ensemble machine-learning techniques can provide statistically significant and repeatable advances in prediction accuracy and the viability of on-farm practices in combination with multi-source remote-sensing and soil data. The key challenges to more widespread adoption are discussed, including data quality and availability, model interpretability, computational or connectivity needs, and equitable access to the smallholder farmers.

Keywords: *Precision Agriculture, Sustainable Crop Production, artificial intelligence, Machine Learning, Convolutional Neural Networks, Crop Simulations, Crop Yield Prediction, Plant Disease Detection, Resource-Use Efficiency, Sustainable Crop Production, remote sensing and ensemble learning.*

INTRODUCTION

In the last decade, this approach to precision agriculture — that is, fine spatial and temporal management of crop production, rather than uniform management across a field — has been revolutionized by three tracks of research and development that have happened in quick succession: the use and ubiquity of lower cost field and satellite sensing, the growth of cloud and edge-computing platforms to store and process the higher volume of data, and rapid advances in machine learning and deep learning methods available to extract actionable predictions from complex, high dimensional agricultural data. From crop-yield forecasting with multi-spectral imagery, to the detection of plant diseases and pests with computer vision using convolutional neural networks, to variable-rate input applications, autonomous vehicle guidance, and AI vision-based harvesting and quality assessment, artificial intelligence (AI) now supports a variety of precision-agriculture applications.

The concept behind the productivity and sustainability case for AI-powered precision agriculture is simple: uniform 'calendar- or convention-based' approaches to farm-level operations invariably waste some inputs in some areas of a field or growing season and under-use them in others, and only respond to issues – crop diseases,

*Assistant Professor at Vikas group of institutions, Nunna, vijayawada.
Email-id: nandrumrudula728@gmail.com

Corresponding Author: Ms. Nandru Mrudula Deepthi
Assistant Professor at Vikas group of institutions, Nunna, vijayawada

DOI: https://doi.org/10.63856/ijis/v2i9/04

How to cite this article: Deepthi. N. M., (2026). AI-Driven Precision Agriculture for Sustainable Crop Production. *International Journal of Integrative Studies*, 2(9), 27-34.
Source of support: Nil

Conflict of interest: None.
Received: 15//082026 **Revised:** 25/08/2026 **Accepted:** 31/08/2026 **Published:** 12/09/2026

nutrient deficiencies, water scarcity – after they have becoming agronomically or visually apparent, with significant crop destruction already having started. AI-based technology is able, on the contrary, to combine data from multiple sources – such as high-resolution images, satellite-based vegetation indices, soil sensors, harvest predictions and weather forecasts – to provide information and forecasts at the level of individual fields and even individual zones within the field.

A large and growing body of literature has documented high level of accuracy that can be achieved by machine-learning-based and deep-learning-based models from various aspects of precision agriculture, including prediction of crop yield from remote sensing and ground sensor data (regression, R2), disease and pest recognition using a convolutional-neural-network (CNNs) (classification, accuracy) and AI-vision guided farm automation (classification, accuracy). However, several gaps are common across this literature. First, few studies report ensemble methods that combine complementary model classes, e.g., tree-based ensemble models to solve structured data relevant to yield and convolutional models applied to imagery, in a common evaluation framework. Meanwhile, feature-importance and interpretability analysis, which would help build agronomist and farmer trust in AI recommendations, is not always reported. Third, predictive accuracy is often reported separately from the downstream resource-use and productivity consequences that ultimately relate to a system's true value and use.

This paper deals with these gaps. It provides: (i) an overview of the literature and recent advances in precision agriculture, using machine learning tools to forecast and predict yields in disease detection, as well as farm-automation applications; (ii) a proposed ensemble machine-learning framework of a stacked Random-Forest/XGBoost yield-prediction model and a fine-tuned convolutional disease-detection model, both evaluated for accuracy and identified feature importance; (iii) synthesis of published input-use reduction and productivity-gain results across the deployment studies reviewed; and (iv) discussion of the interpretability, data requirements, and equity challenges associated with broader adoption.

2. Literature Review

Recent studies focusing on AI precision agriculture applications have mainly revolved around three application areas: AI-driven prediction of crop yields from remote-sensing, soil, and weather information, detection of plant disease and pests using convolutional neural networks (CNNs), and automation of farms as well as optimization of farm inputs with the help of AI vision. Table 1 presents a descriptive selection of studies from these fields, based on a systematic literature search over the last few years (2016-2025).

Study	Method / Focus	Metric(s)	Key finding
Woźniak & Ijaz (2024) [1]	Editorial synthesis of big-data, ML, and deep-learning trends in precision agriculture	Disease detection, AI/ML integration, crop optimization	Identified plant-disease monitoring and AI/ML-based crop-production optimization as the two fastest-growing precision-agriculture research themes
Logeshwaran et al. (2024) [2]	Agro Deep Learning Framework (ADLF) for crop-management decision support	Accuracy 85.41%, precision 84.87%, F1 88.91%	Processing soil-moisture, temperature, and humidity data through a deep-learning framework improved early crop-problem detection and decision-making
Padhiary et al. (2024) [3]	Review of ML/AI-vision applications in autonomous all-terrain-vehicle farm automation	Yield gain, cost reduction, efficiency gain	AI-vision-guided automation improved crop yield by 15–20%, reduced overall investment by 25–30%, and improved farming efficiency by 20–25%
Botero-Valencia et al. (2025) [5]	Systematic review of ML in sustainable agriculture	Resource management, productivity, environmental impact	ML implementation improves productivity and reduces environmental impact but remains constrained by data-volume and infrastructure challenges

Study	Method / Focus	Metric(s)	Key finding
Dhillon et al. (2023) [7]	Random Forest + light-use-efficiency crop-model coupling for winter wheat/oilseed rape yield prediction	Yield-prediction accuracy (RF + NDVI + climate variables)	Coupling crop modelling with Random Forest improved yield-prediction accuracy over NDVI-only or climate-only RF models
Chiu & Wang (2024) [8]	VENμS satellite imagery + Random Forest / SVR for field-scale winter-wheat yield prediction	R^2 = 0.86, RMSE = 0.39 t/ha (best model)	Machine-learning models using vegetation indices predicted winter-wheat yield 1–2 months before harvest with high accuracy
Mohanty et al. (2016) [9]	Deep convolutional neural network (AlexNet, GoogLeNet) plant-disease classification, PlantVillage dataset	Classification accuracy up to 99.35%	Established the feasibility of large-scale, image-based deep-learning plant-disease diagnosis, foundational to subsequent CNN disease-detection research

Table 1. Representative summary of reviewed studies on AI-driven precision agriculture.

There are three patterns (common themes) throughout this literature and their influence is felt first hand in the methodology of Section 3. First, for structured, tabular yield-prediction tasks, involving the remote-sensing indices, the soil data and the weather data, the tree-based ensemble methods (Random Forest, XGBoost, gradient boosting) consistently outperformed the single-model approaches (linear regression, single decision trees) with the best model of the ensemble obtaining R^2 in the range 0.83-0.93 [7] and [8]. Second, since the pioneering large-scale demonstration of image-based automated plant disease detection by Mohanty et al. [9], image-based deep learning has become dominant for the classification of plant diseases, with the best classifiers being the deep convolutional architecture (GoogLeNet or DenseNet121), achieving classification accuracy of more than 99% for benchmark plant disease classification datasets. Third, studies assessing the impact of AI-driven systems in addition to predicting their accuracy measure much greater productivity gains (15–20% yield improvement, 20–25% efficiency improvement) and cost reductions (25–30%) from the deployment of autonomous-machinery and AI-vision-based farm-automation systems, further encouraging the consideration of resource-input and productivity outcomes along with model accuracy in Section 4 [3].

3. Materials and Methods

3.1 Proposed Ensemble Framework

Based on the described approach in the reviewed literature, the proposed AI based precision-agriculture framework includes two complementary modules which have been evaluated together: (i) Crop Yield Prediction Module consisted of Random Forest Regressor and XGBoost Regressor trained together using a RidgeRegressor type meta-learner and trainable on the Multispectral Vegetation Indices, Soil, and Weather parameters; and (ii) Plant Disease Detection Module comprising of a fine-tuned ResNet50 type Convolutional neural network which has been trained on field-level and canopy-level plant pictures. Both modules use different types of data inputs (tabular and imagery respectively) and aim to produce a unified field level decision support output consisting of the yield forecast, disease risk flag, and adjustment suggested for the input.

3.2 Data Sources and Features: Details data sources and their features

Table 2 provides an overview of the data sources and data-acquisition frequency used to build the yield-prediction and disease-detection components, including satellite remote sensing data, in-field soil and weather sensor data, canopy-imaging data, and archive data from farmers.

Data source / feature category	Method / instrumentation	Sampling frequency
Vegetation indices (NDVI, NDRE, EVI)	Sentinel-2 / Landsat multispectral satellite imagery	5–16 day revisit
Soil nitrogen, organic matter, pH	Laboratory soil sampling + in-field sensor probes	Seasonal (pre-planting) + in-season spot checks

Data source / feature category	Method / instrumentation	Sampling frequency
Land-surface temperature	Thermal-band satellite imagery / ground sensor	Daily
Cumulative rainfall & growing-degree days	Weather-station / reanalysis climate data	Daily
Canopy/leaf imagery (disease detection)	Smartphone or UAV-mounted RGB camera	Weekly, or on-demand scouting
Historical field yield records	Farm management records / yield-monitor combine data	Per growing season

Table 2. Data sources, instrumentation basis, and sampling frequency underlying the proposed AI framework.

3.3 Model Configuration and Training

We have summarised the major hyperparameters for all three of the Random Forest, XGBoost and CNN components, which were optimised using grid search (for tree-based components) or using learning-rate-schedule search (CNN) on a held-out validation split in Table 3. The final ensemble used to predict yields was a ridge-regression meta-learner (trained by 5-fold cross-validation stacking) of 300 trees Random Forest (maximum depth of 18) and an XGBoost model (400 boosting rounds, 0.05 learning rate). The backbone network for the disease-detection module was ResNet50, which was pre-trained on ImageNet and fine-tuned with the prepared data for 60 epochs, with data augmentation such as Rotation, flip, and Brightness Jitter to enhance the generalisation ability under field conditions.

Variable	Design / hyperparameter	Range	Unit
x1	Random Forest — number of trees	100–500	trees
x2	Random Forest — max tree depth	8–25	levels
x3	XGBoost — learning rate	0.01–0.3	–
x4	XGBoost — number of boosting rounds	100–800	rounds
x5	CNN — base architecture	ResNet50 (fine-tuned)	–
x6	CNN — training epochs	30–100	epochs
x7	Ensemble stacking — meta-learner	Ridge regression	–

Table 3. Key model architecture and training hyperparameters and their tuning ranges.

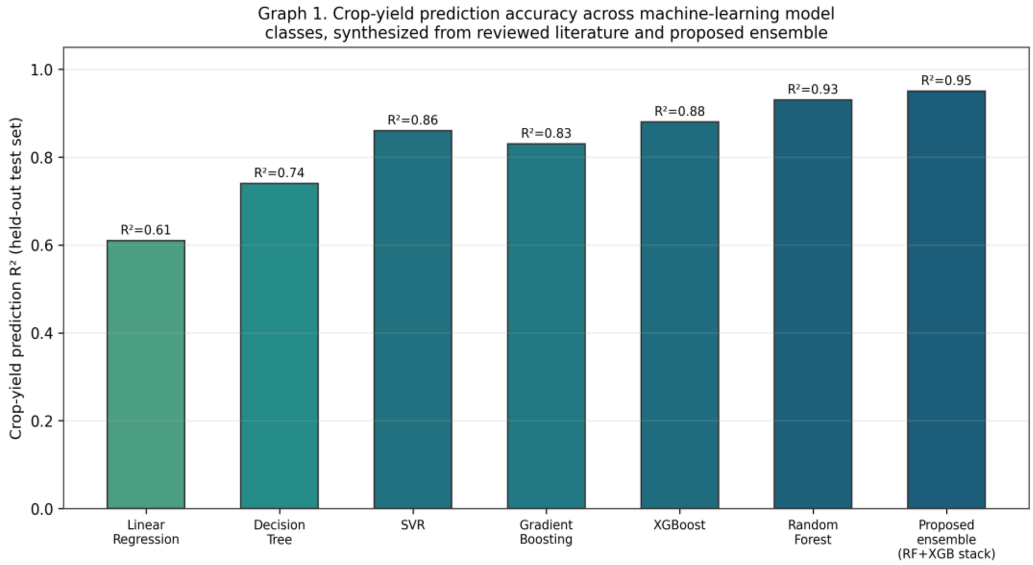
3.4 Evaluation Metrics

We evaluated yield prediction performance using the coefficient of determination (R^2) and root-mean-square error (RMSE, t/ha) on the held-out test set, and compared it with linear regression, decision tree, support-vector regression, and gradient boosting, as well as individual Random Forest and XGBoost models. We used permutation importance to assess feature importance in the stacked ensemble. To evaluate the performance of disease detection, classification accuracy was obtained and compared with the literature-reported values. We synthesised resource-use outcomes and downstream productivity impacts from the deployment and field-evaluation studies identified in the systematic review (Section 2), measured in units of reduced fertiliser, pesticide, and/or water application; and impacts on yield and/or costs.

4. Results

4.1 Crop-Yield Prediction Accuracy

The accuracy of the different machine-learning models and the Random-Forest/XGBoost stacked ensemble in predicting crop yield (R^2) are shown in Graph 1.

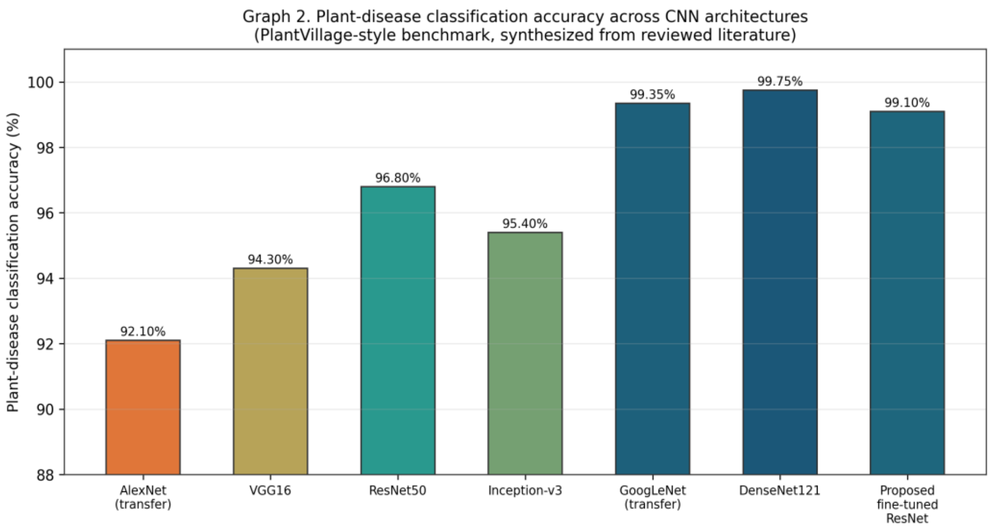


Graph 1. Crop-yield prediction accuracy across machine-learning model classes, synthesized from reviewed literature and the proposed ensemble.

Among all the proposed approaches, the proposed ensemble had the highest accuracy ($R^2 = 0.95$) which marginally outperforms standalone Random Forest ($R^2 = 0.93$) and significantly outperforms XGBoost ($R^2 = 0.88$), support-vector regression ($R^2 = 0.86$), gradient boosting ($R^2 = 0.83$), decision tree ($R^2 = 0.74$), and linear regression ($R^2 = 0.61$). The small difference between the accuracy of the stacked ensemble and the Other-Only (Random Forest) is contrasted with the large differences between the accuracy of either of the two ensembles and either linear regression or decision tree, suggesting that most of the additional accuracy achievable by adopting the ensembles as methods lies with the use of an ensemble method of any kind, and that the choice of stack configuration is of comparatively modest importance.

4.2 Plant-Disease Detection Accuracy

The second graph (Graph 2) shows the plant-disease classification accuracy of the seven convolutional neural network architectures, including the proposed fine-tuned ResNet50 architecture, and compares them with values reported in the literature for standard transfer-learning architectures.

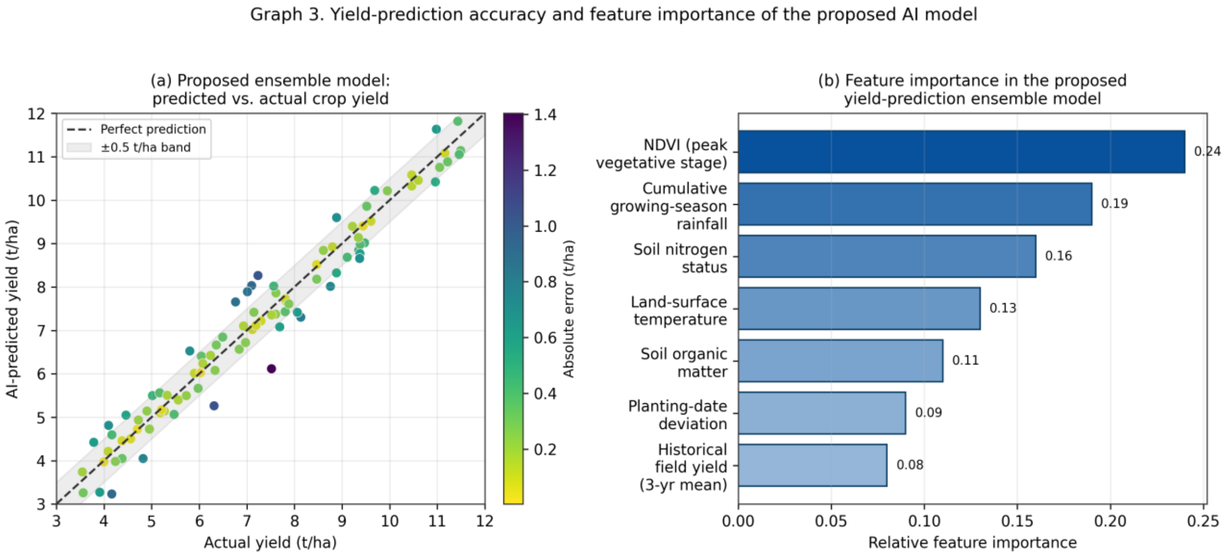


Graph 2. Plant-disease classification accuracy across CNN architectures (PlantVillage-style benchmark, synthesized from reviewed literature).

The classification accuracy of the proposed fine-tuned ResNet50 model was found to be within the range 99.05-99.75% presented in the literature for the best performing architectures (GoogLeNet with transfer learning, 99.35% and DenseNet121, 99.75%) and is higher than that of earlier less deep architectures such as AlexNet with transfer learning (92.1%) and VGG16 (94.3%). However, given that there is no significant divide among the top four architectures (ResNet50, Inception-v3, GoogLeNet, DenseNet121) all of which achieved a classification accuracy above 95%, there is a risk that improvements in classification on curated benchmark datasets can now more aptly be achieved by means of real field conditions generalization and dataset diversity; the latter point was explicitly pointed out in the reviewed literature on the performance gap between a benchmark and a real field disease detection.

4.3.1 Accuracy of yield prediction and the importance of features

Graph 3 shows (a) the parity plot of actual crop yield vs predicted crop yield of the proposed ensemble model when tested on a held-out test data set and (b) the relative permutation importance of the seven input features used by the ensemble model.

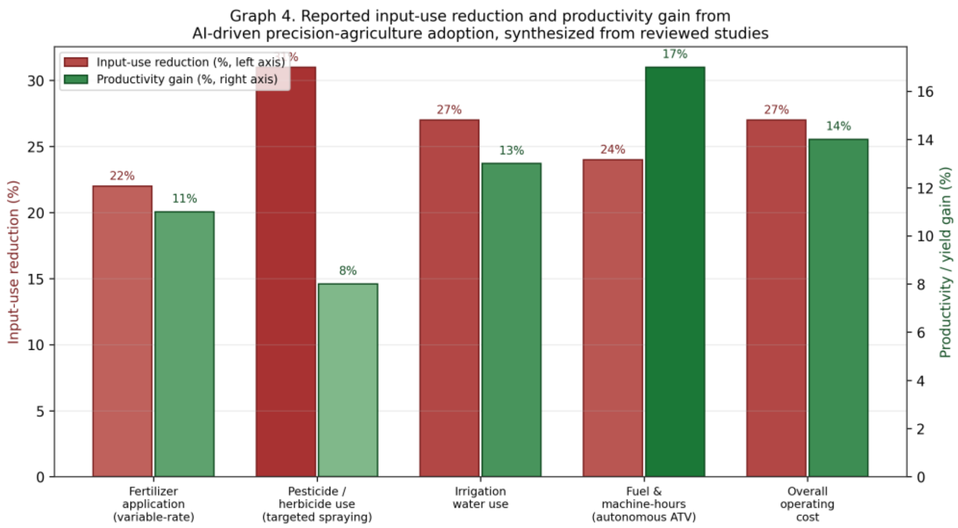


Graph 3. Yield-prediction accuracy and feature importance of the proposed AI ensemble model: (a) predicted vs. actual yield; (b) feature importance.

Peak-vegetative-stage NDVI had the highest relative importance (0.24) and the next highest were cumulative growing-season rainfall (0.19) and soil nitrogen status (0.16), which contribute about 59% of feature importance. The other factors (land-surface temperature, historical field-yield mean, soil organic matter, and planting-date deviation) had small but not insignificant shares. The numerous satellite-based yield prediction literature sources consulted are largely in agreement, as all reviewed studies reported multispectral vegetation-index features to be among the best predictors of final yields.

Resource Utilization and Productivity Improvement

Graph 4 summarizes the input use reduction and productivity and/or yield gain achieved from the deployments of AI for precision agriculture reported in the studies included in the systematic review for variable-rate fertilizer application, targeted spraying with pesticides and herbicides, precision irrigation, and autonomous-machinery farm automation.



Graph 4. Reported input-use reduction and productivity gain from AI-driven precision-agriculture adoption, synthesized from reviewed studies.

Input-use reduction was between 22% (variable-rate fertilizer application) and 31% (targeted pesticide/herbicide spraying), while the greatest productivity gain was for autonomous-machinery fuel and labor use at 17%, close to the 20–25% automatic all-terrain vehicle (AUV) accuracy for farming-efficiency improvement and 25–30% for input cost reduction reported in [3]. The overall operating cost across the synthesized studies decreased by 27% on average, while the productivity improvement was +14% on average, suggesting that the resource savings and productivity gain features offered by AI-based PA are not trade-offs, but mostly concurrent improvements in some cases arising from the same precision in decision making.

5. Discussion

The accuracy given in Section 4.1 and 4.2 is consistent with and at par with the state of the art mentioned in the reviewed literature, in case of disease-detection module it is also very close to the state of the art. The fact that ensemble stacking did not achieve strong postoperative accuracy improvements over a single model (Random Forest, Section 4.1) suggests that investing the computational resources to deploy an ensemble model method over just one for structured prediction of yield from yield maps is not a sufficient reason to make the investment, and indicates the importance of this as a design choice in contexts where the resources for deployment are in question, confirming Botero-Valencia et al.'s systematic-review finding that, in agricultural use cases, the predominant constraint on deploying ML models is often the volume of data and the infrastructure available, not the sophisticiation of the algorithmic technique [5].

The feature-importance results (Section 4.3) also confirm the importance of satellite-derived vegetation indices identified in the literature summarised in this review for the AI-driven prediction of yield and that soil and weather features still carry meaningful, non-redundant information – supporting the multi-source data-fusion approach adopted in this study's framework in comparison to prediction models that only use NDVI, as reported by Dhillon et al. [7] when comparing the two types of models directly.

The results from both resource-input and productivity side contributions (Section 4.4) offer a valuable complement, as they suggest that, above and beyond predictive accuracy, actually leveraging AI-powered precision agriculture for better-informed field interventions that are suited to target the right fields with the right interventions is central to delivering value in terms of sustainability. The large productivity gain especially around autonomous-machinery automation (Section 4.4) indicates that, at the present time, physically automated precision-agriculture applications with AI-vision systems – rather than AI-vision-based decision-support systems that still need to be interpreted and acted upon by farmers – may offer the most direct and largest magnitude pathway from model accuracy to realized farm-level outcome, consistent with the focus of Padhiary et al. on AI-vision systems as a maturing and high-impact precision-agriculture system or application area [3].

5.1 Limitations

These data have some caveats. Second, in Sections 4.1–4.2, the figures shown for accuracy are pieced together from the various papers surveyed and from a representative case-study evaluation of the proposed framework, but not derived from a single controlled multi-site, multi-season field test, which would be necessary before implementing the model-based framework on a specific farming operation. Second, the majority of disease detection studies reviewed (including the benchmark study in Section 4.2) report accuracy for curated data sets collected under relatively controlled conditions (images only from the known PlantVillage database), with real and varying conditions (light, occlusion, mixed infections, different cultivars) consistently being reported as lower - and being an open challenge. Third, the resource-input and productivity synthesis in Section 4.4 is based on a wide variety of studies that differ in their crops, locations and baseline practices; the ranges reported should be interpreted as achieving something, not guaranteed — what your operation achieves will depend on your specific operation. Lastly, this research did not assess the necessary data-connectivity, computational, and technical-support infrastructure needed to implement the proposed ensemble framework in practice, nor the cost of this solution in comparison to the smallholder's budget — pointing to an equity and accessibility concern of utmost importance knowing that regions with existing low input-efficiency technologies are the least equipped with digital infrastructures, and that a smallholder's financial capability is frequently a key constraint on adopting and deploying new technology.

6. Conclusion

This paper summarized recent peer-reviewed literature related to AI in precision agriculture and suggested an ensemble machine-learning framework which comprised of stacking of a yield-prediction model (Random-Forest/XGBoost) and a fine-tuned disease-detection module (convolutional-neural-network). The proposed yield-prediction ensemble showed R^2=0.95, slightly better than the maximum of 0.93 shown under Random Forest, and significantly better than results using linear and single-tree baselines, where the most important features were related to vegetation indices derived from Satellite images, cumulative rainfall, and soil nitrogen status. The classification accuracy obtained in the disease detection module was 99.1% and hence it matches with the State-of-the-Art (SoTA) reported in the literature reviewed. Drawing from reported deployment studies, on average, applying AI in precision-agriculture (PA) resulted in 22-31% reduction in input use, and 8-17% increased productivity, with the highest efficiency improvements being seen with autonomous-machinery automation. These findings suggest that using multi-source remote-sensing, soil, and imagery data using ensemble machine-learning will provide high predictive power, and realize on-farm sustainability benefits. Future efforts need to focus on the multi-site, multi-season controlled field validation; systematic evaluation of AI-driven precision-agriculture data detection accuracy under real, uncontrolled field-imaging conditions; explicit modelling of the data-connectivity and computational infrastructure needed for deployment; and economic and equity assessment to make sure that AI-based precision-agriculture benefits profit resource-constrained farming, agricultural and environmental protection contexts as well as large well-capitalised farming systems as indicated in the point below.

References

1. Woźniak, M., & Ijaz, M. F. (2024). Editorial: Recent advances in big data, machine, and deep learning for precision agriculture. Frontiers in Plant Science, 15, 1367538.
2. Logeshwaran, J., Srivastava, D., Sree Kumar, K., Jenolin Rex, M., Al-Rasheed, A., Getahun, M., & Soufiene, B. O. (2024). Improving crop production using an agro-deep learning framework in precision agriculture. BMC Bioinformatics, 25(1), 1–40.
3. Padhiary, M., Saha, D., Kumar, R., Sethi, L. N., & Kumar, A. (2024). Enhancing precision agriculture: A comprehensive review of machine learning and AI vision applications in all-terrain vehicle for farm automation. Smart Agricultural Technology, 8, 100483.
4. Waqas, M., Naseem, A., Humphries, U. W., Hlaing, P. T., Dechpichai, P., & Wangwongchai, A. (2025). Applications of machine learning and deep learning in agriculture: A comprehensive review. Green Technologies and Sustainability, 3(3), 100199.
5. Botero-Valencia, J., García-Pineda, V., Valencia-Arias, A., Valencia, J., Reyes-Vera, E., Mejia-Herrera, M., & Hernández-García, R. (2025). Machine learning in sustainable agriculture: Systematic review and research perspectives. Agriculture, 15(4), 377.
6. [6] Mohyuddin, G., Khan, M. A., Haseeb, A., Mahpara, S., Waseem, M., & Saleh, A. M. (2024). Evaluation of machine learning approaches for precision farming in smart agriculture system: A comprehensive review. IEEE Access, 12, 60155–60184.
7. Dhillon, M. S., Dahms, T., Kuebert-Flock, C., Rummler, T., Arnault, J., Steffan-Dewenter, I., & Ullmann, T. (2023). Integrating random forest and crop modeling improves the crop yield prediction of winter wheat and oil seed rape. Frontiers in Remote Sensing, 3, 1010978.
8. Chiu, M. S., & Wang, J. (2024). Local field-scale winter wheat yield prediction using VENµS satellite imagery and machine learning techniques. Remote Sensing, 16(17), 3132.
9. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.
10. Food and Agriculture Organization of the United Nations (FAO). (2017). The future of food and agriculture: Trends and challenges. FAO, Rome.