Scientia Review

Website: www.scientiafoundation.in

REVIEW ARTICLE

From Predictive Power to Productive Practice: A Critical Synthesis of Artificial Intelligence Applications in Agricultural Systems

Pranabesh Barman*, Pallavi Deka, Kapil Deb Nath, Himadri Rabha, Pradip Rajbongshi
Krishi Vigyan Kendra, Udalguri
*Email: pranabesh.barman@aau.ac.in

Received: 03 September 2026
Revised: 04 September 2026
Accepted: 05 September 2026
Published online: 05 September 2026

Article ID: SR01157

Citation: Barman, P., Deka, P., Nath, K. D., Rabha, H., & Rajbongshi, P. (2026). From predictive power to productive practice: A critical synthesis of artificial intelligence applications in agricultural systems. *Scientia Review*, 2(9), 26-43.

Abstract

Agriculture faces mounting pressure from population growth, climate volatility, resource scarcity, labour shortages and stagnant farm profitability, and artificial intelligence is widely promoted as a route toward more productive, efficient and resilient food systems. A large and rapidly expanding body of research reports impressive algorithmic performance across crop monitoring, disease detection, yield prediction, irrigation scheduling, weed management, robotics and livestock monitoring. However, the extent to which these algorithmic achievements translate into meaningful farm level and system level outcomes remains contested and, in many domains, poorly evidenced. The evidence indicates that classification and prediction models frequently achieve high accuracy under controlled or benchmark conditions, yet studies that validate these models across seasons, geographies and farm types remain comparatively scarce, and rigorous farm level economic evaluations are even rarer. The literature further reveals persistent disagreement over whether artificial intelligence narrows or widens inequalities between large commercial operations and smallholder or resource constrained farmers. Data governance, algorithmic bias, limited explain ability and weak integration with agricultural extension systems constitute recurring and unresolved concerns. The review concludes that artificial intelligence has produced genuine advances in agronomic prediction and automation, but that the pathway from algorithmic accuracy to agricultural transformation is neither automatic nor uniform, and that farmer centered design, longitudinal field validation and responsible data governance are essential preconditions for translating technical promise into equitable and sustainable agricultural outcomes.

Keywords: Artificial Intelligence, Machine learning, Deep learning, Robotics, Precision

1. Introduction

Agriculture occupies a uniquely constrained position among global production systems. It must simultaneously expand output to meet the food requirements of a growing population, adapt to increasingly erratic climatic conditions, operate within tightening limits on water, soil and biodiversity, absorb persistent shortages of agricultural labour and remain financially viable for farmers whose margins are frequently thin and volatile (Weersink *et al*., 2018; Basso and Antle, 2020). These converging pressures have driven sustained interest in digital agriculture, smart farming and what many scholars term Agriculture 4.0, a paradigm in which sensors, connectivity and computational analysis are integrated into everyday farm management (Klerkx *et al*., 2019; Rose and Chilvers, 2018). Within this broader digital transition, artificial intelligence has attracted particular attention because of its capacity to extract patterns from large and heterogeneous agricultural datasets, including satellite imagery, unmanned aerial vehicle photography, soil and weather sensor streams, farm machinery telemetry and market records (Kamilaris and Prenafeta-Boldú, 2018; Wolfert *et al*., 2017).

The appeal of artificial intelligence in agriculture rests on a straightforward proposition, namely that machine learning, deep learning, computer vision, natural language processing, robotics and predictive analytics can convert raw agricultural data into timely and accurate recommendations that support better farm decisions than unaided human judgement alone (Liakos *et al*., 2018; Benos *et al*., 2021). Convolutional neural networks have achieved remarkably high accuracy in identifying crop diseases from leaf images (Mohanty *et al*., 2016; Ferentinos, 2018), ensemble and deep learning models have been applied extensively to yield forecasting (Chlingaryan *et al*., 2018; van Klompenburg *et al*., 2020), and autonomous machinery, robotic harvesters and unmanned aerial systems are increasingly deployed for field operations that were previously entirely manual (Bechar and Vigneault, 2016; Fountas *et al*., 2020). More recently, generative artificial intelligence and large language models have been proposed as a means of simplifying complex agronomic knowledge and delivering it through conversational interfaces to farmers who previously depended on scarce human extension agents (Tzachor *et al*., 2023).

However, this review resists treating the expansion of artificial intelligence applications as self evidently synonymous with agricultural progress. A technology catalogue that simply enumerates machine learning applications across crop, soil, water, pest and livestock domains risks obscuring a more fundamental and still unresolved question, namely whether reported algorithmic performance actually translates into meaningful agricultural outcomes at the level of the farm and the wider food system. Several existing reviews of artificial intelligence in agriculture have tended toward descriptive cataloguing of techniques and applications rather than critical evaluation of evidentiary strength (Liakos *et al*., 2018; Benos *et al*., 2021), and comparatively few have systematically distinguished the accuracy of a model from the value it creates once deployed under real farm conditions. This gap is consequential because a classification model achieving ninety five percent accuracy in a controlled dataset does not automatically imply that farmers using the corresponding tool will make better decisions, achieve higher yields, reduce costs or improve their incomes (Finger *et al*., 2019; Lowenberg-DeBoer and Erickson, 2019).

This review therefore adopts an explicitly critical and evidence based stance. It synthesises what previous researchers have reported, identifies points of agreement and disagreement across the literature, evaluates the methodological strength of the available evidence, and distinguishes consistently between algorithmic performance, operational performance, farm level outcomes and system level outcomes. The review moves through a logical pathway that begins with artificial intelligence technologies and proceeds through agricultural applications, decision support, farm management, farmer adoption, productivity, profitability, resource efficiency, sustainability, resilience, farmer empowerment and ultimately agricultural transformation, a pathway summarised graphically in Figure 1. In doing so it addresses several review objectives. It sets out to identify the agricultural problems that artificial intelligence currently addresses with strong empirical support and those where evidence remains largely experimental, to examine the reliability of artificial intelligence models under field rather than laboratory conditions, to assess whether improvements in predictive accuracy correspond to improved farm decisions, productivity, cost reduction, income and resource efficiency, to determine which farmers benefit most and least from these technologies, to evaluate the role of agricultural extension in mediating artificial intelligence adoption, to examine emerging ethical and data governance concerns, and to identify the research gaps that must be closed before confident claims of agricultural transformation can be sustained.

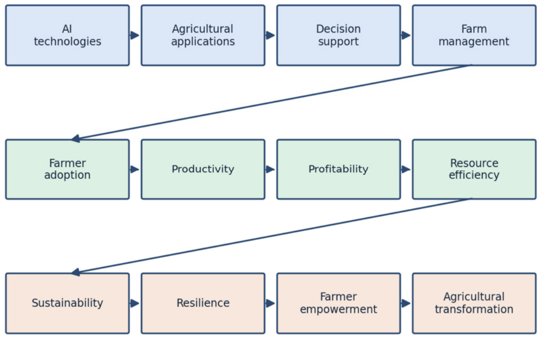

Figure 1. The pathway examined in this review, from artificial intelligence technologies to agricultural transformation. Each stage represents a distinct level of claim commonly made in the literature, and this review evaluates the strength of evidence supporting the transition from one stage to the next rather than assuming the pathway operates automatically.

Review Methodology

This review adopts a systematic and thematically organised methodology consistent with the principles associated with the Preferred Reporting Items for Systematic Reviews and Meta Analyses, commonly known by the acronym PRISMA, while recognising that a review of this scope, spanning computer science, agronomy, engineering, economics and rural development, benefits from a narrative synthesis structure rather than a narrowly quantitative meta analysis. Literature identification drew on the major academic databases relevant to this interdisciplinary field, including Scopus, Web of Science, ScienceDirect, SpringerLink, Wiley Online Library, IEEE Xplore and Google Scholar, since coverage of computer science, engineering and agricultural economics literature varies considerably across these platforms and no single database indexes the full breadth of relevant scholarship. Search terms combined artificial intelligence and its principal subfields, namely machine learning, deep learning, generative artificial intelligence, computer vision, robotics and natural language processing, with agricultural and rural development concepts including precision agriculture, smart farming, digital agriculture, Agriculture 4.0, crop monitoring, crop disease detection, yield prediction, irrigation, soil management, weed detection, pest management, farm robotics, autonomous agriculture, decision support, agricultural extension, smallholder farmers, farmer adoption, agricultural sustainability, farm economics and agricultural productivity. These terms were combined using standard Boolean logic, pairing technology terms with application domain terms to capture studies at the intersection of artificial intelligence and agricultural practice rather than studies addressing either domain in isolation.

The review considered publications primarily from 2016 to 2026, reflecting the period during which deep learning based approaches became widely applied to agricultural problems, while also drawing on earlier seminal contributions where these remain foundational to understanding the development of machine learning, remote sensing, precision agriculture, decision support systems, robotics, computer vision and the Internet of Things within agriculture. Studies were included where artificial intelligence or a closely related computational method constituted a substantive focus of the work and where the application context was directly agricultural, encompassing crop production, livestock, agricultural extension, farm economics or rural development. Studies were excluded where artificial intelligence was mentioned only superficially without methodological substance, where the agricultural relevance was marginal or incidental, where the work constituted a duplicate or near duplicate of an already identified publication, or where the source was a non scholarly website, blog, unverified preprint lacking peer review, or purely commercial material. Priority was given throughout to research published in journals indexed in Scopus or Web of Science and associated with established academic publishers, including Computers and Electronics in Agriculture, Agricultural Systems, Biosystems Engineering, Precision Agriculture, Artificial Intelligence in Agriculture, Smart Agricultural Technology, Sensors, Remote Sensing, Journal of Cleaner Production, Agricultural Economics, the Economic Journal, World Development, Nature, Nature Sustainability, Nature Food and Frontiers journals, among others.

Screening proceeded in stages consistent with the logic of identification, screening, eligibility assessment and final inclusion that PRISMA describes, beginning with an assessment of title and abstract relevance against the inclusion and exclusion criteria described above, followed by full text assessment of studies passing initial screening to confirm methodological substance and direct agricultural relevance, and concluding with citation tracing through the reference lists of key reviews and seminal papers to identify additional studies not captured through database searching alone. Because the precise number of records retrieved at each database varies across platforms and changes over time as indexing is updated, this review does not report a single fixed count of records screened, and instead reports only the verified bibliographic details of the studies ultimately included, consistent with the principle that a review should never present invented or unverifiable numerical figures. The overall screening logic followed in this review is summarised in Figure 2.

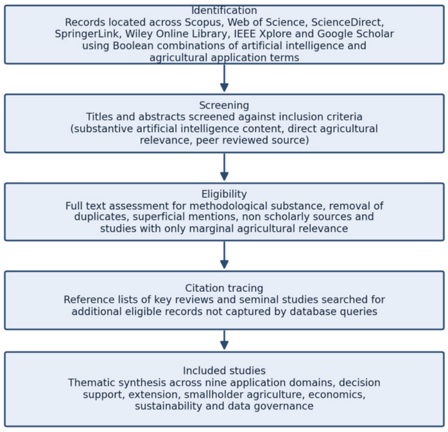

Figure 2. Literature identification and screening process, consistent with PRISMA principles and adapted to a narrative synthesis in which precise record counts at each stage are not fabricated where verified figures are unavailable.

Included studies were then grouped and synthesised thematically rather than chronologically, organised around the major agricultural application domains outlined in the article structure, including crop monitoring and disease detection, yield prediction, soil and nutrient management, irrigation and water management, weed and pest management, weather and climate risk management, agricultural robotics, livestock and poultry production, post harvest management and supply chains, agricultural extension, smallholder agriculture, farm economics, sustainability and data governance. Evidence within each theme was compared explicitly for points of agreement and disagreement, for the methodological rigour of the underlying study designs, particularly the distinction between benchmark or laboratory evaluation and field or farm level validation, and for the practical relevance of reported findings to farmer decision making and measurable agricultural outcomes, which together informed the critical synthesis presented in the sections that follow. Table 1 summarises, at a broad level, the technologies, characteristic methods and representative studies associated with each thematic domain, and provides an orientation point for the more detailed critical discussion developed in sections three through ten.

Table 1. Overview of artificial intelligence technologies and thematic application domains covered in this review

Thematic domain	Principal AI techniques	Typical data sources	Source
Crop monitoring and disease detection	Convolutional neural networks, transfer learning	Leaf images, UAV and smartphone photography	Mohanty *et al.* (2016); Ferentinos (2018); Too *et al.* (2019)

Thematic domain	Principal AI techniques	Typical data sources	Source
Crop yield prediction	Random forests, gradient boosting, deep learning, hybrid models	Remote sensing indices, weather records, soil and historical yield data	Chlingaryan *et al.* (2018); van Klompenburg *et al.* (2020)
Soil and nutrient management	Regression and classification machine learning, remote sensing fusion	Proximal soil sensors, satellite imagery, laboratory soil tests	Chlingaryan *et al.* (2018); Sishodia *et al.* (2020)
Irrigation and water management	Reinforcement learning, sensor fusion, control algorithms	Soil moisture sensors, weather forecasts, evapotranspiration models	Abioye *et al.* (2020); Sishodia *et al.* (2020)
Weed and pest management	Object detection networks, semantic segmentation	Field images, UAV imagery, multispectral sensors	Hasan *et al.* (2021); Rakhmatulin *et al.* (2021)
Agricultural robotics	Computer vision, robotic manipulation, path planning	Onboard cameras, LiDAR, GPS	Bechar and Vigneault (2016); Fountas *et al.* (2020); Zhang *et al.* (2020)
Livestock and poultry farming	Computer vision, sensor based classification, natural language processing	Video and image streams, wearable sensors, herd records	Li *et al.* (2021); Lovarelli *et al.* (2020); Menezes *et al.* (2024)
Agricultural extension and advisory	Recommendation systems, chatbots, large language models	Farmer queries, agronomic knowledge bases	Steinke *et al.* (2021); Tzachor *et al.* (2023); De Clercq *et al.* (2024)

2. Conceptual Foundation of Artificial Intelligence in Agriculture

Artificial intelligence, in the agricultural context, refers to a family of computational methods that enable machines to perform tasks that would ordinarily require human perception, reasoning or judgement, including recognising crop diseases from images, forecasting yields from historical and environmental data, scheduling irrigation, guiding autonomous vehicles and generating agronomic advice in natural language. Machine learning is the subset of artificial intelligence in which algorithms learn statistical patterns from data rather than following explicitly programmed rules, while deep learning refers specifically to machine learning architectures built from multiple layers of artificial neurons, most notably convolutional neural networks for image analysis and recurrent or transformer based architectures for sequential and textual data (Kamilaris and Prenafeta-Boldú, 2018; LeCun *et al*., 2015). Among conventional machine learning algorithms, ensemble methods such as the random forest algorithm described by Breiman (2001) remain widely used across agricultural prediction tasks because they combine strong predictive accuracy with comparative robustness to noisy or limited training data, a combination that Sharma *et al*. (2021) identify as a principal reason for their continued popularity in precision agriculture applications even as deep learning architectures have become more prominent. Computer vision applies these architectures to visual data such as leaf images, aerial photography and livestock video, whereas predictive analytics more broadly encompasses statistical and machine learning techniques used to forecast future agronomic or economic outcomes such as yield, price or disease incidence (Benos *et al*., 2021). Robotics integrates sensing, planning and actuation to allow physical machines to perform field operations autonomously or semi autonomously (Bechar and Vigneault, 2016), while decision support systems combine data, models and user interfaces to help farmers and advisers choose among alternative management options (Zhai *et al*., 2020, as cited in related literature on Agriculture 4.0 decision support). Generative artificial intelligence, exemplified by large language models, differs from these earlier approaches in that it produces novel text, dialogue or synthetic content rather than a single numeric prediction, and has begun to be explored as a natural language interface between complex agronomic knowledge systems and farmers (Tzachor *et al*., 2023; De Clercq *et al*., 2024). Bibliometric analyses of this expanding body of scholarship, such as that conducted by Coulibaly *et al*. (2022), indicate that deep learning publications in agriculture have grown rapidly since the middle of the previous decade, with the majority of this growth concentrated in image based crop applications rather than in economic or social impact evaluation, a pattern that anticipates a central concern developed later in this review.

These systems depend entirely on the data that feed them, and agricultural data are generated through an increasingly diverse set of sources, including satellite remote sensing, unmanned aerial vehicle imagery, soil and weather sensors distributed across fields, telemetry from farm machinery, mobile applications used by farmers, administrative and market records and, increasingly, livestock mounted sensors that record movement, temperature and feeding behaviour (Sishodia *et al*., 2020; Wolfert *et al*., 2017). The relationship between data, algorithms, predictions, decisions and outcomes forms what can usefully be described as the agricultural data pipeline. Data are collected and pre processed, algorithms are trained to detect patterns or make predictions, these predictions are translated into recommendations, farmers or automated systems act on those recommendations, and the resulting management interventions eventually produce measurable agronomic, economic, environmental or social outcomes. Each stage of this pipeline introduces potential sources of error and distortion. Poor sensor calibration, inconsistent labelling, limited geographic and crop diversity in training datasets, and inadequate independent validation can all reduce the reliability of an artificial intelligence system long before it reaches a farmer (Lu and Young, 2020; Weersink *et al*., 2018). Datasets used to train crop disease detection models, for example, are frequently collected under controlled photographic conditions and dominated by a narrow range of crop varieties and geographic regions, which raises legitimate concerns about how well such models generalise to the visual variability, lighting conditions and disease presentations encountered in commercial fields (Mohanty *et al*., 2016; Too *et al*., 2019). This chain of dependency, from data quality through algorithmic design to eventual farm outcome, is the conceptual foundation on which the remainder of this review builds its critical analysis.

3. Artificial Intelligence Applications Across the Agricultural Production System

Bibliometric mapping of this literature, including the analysis conducted by Rejeb *et al*. (2024), indicates that research output on precision agriculture and related artificial intelligence applications has expanded at an accelerating pace over the past decade, with crop monitoring, yield prediction and irrigation management representing the most densely researched clusters, while livestock, post harvest and supply chain applications remain comparatively less developed. The nine thematic areas addressed in this section follow a similar ordering of research intensity, moving from the most extensively studied domains toward those where the evidence base remains thinner.

3.1 Crop Monitoring and Disease Detection

Image based disease detection represents one of the most extensively researched applications of artificial intelligence in agriculture. Convolutional neural network models trained on large repositories of leaf images have repeatedly achieved classification accuracies exceeding ninety percent, with Mohanty *et al*. (2016) reporting accuracy above ninety nine percent on the PlantVillage dataset and Ferentinos (2018) reporting comparably high performance across several deep architectures. Too *et al*. (2019) extended this line of research by comparing multiple fine tuned deep learning architectures for plant disease identification and found that model choice and transfer learning strategy materially affected classification performance. Kamilaris and Prenafeta-Boldú (2018), synthesising forty research efforts applying deep learning to agricultural and food production problems, concluded that deep learning consistently outperformed conventional image processing techniques. Nevertheless, a persistent and largely unresolved limitation runs through this literature. Laboratory performance figures are overwhelmingly derived from images captured under controlled backgrounds, consistent lighting and, in many cases, a narrow set of crop varieties and disease categories, whereas commercial fields present cluttered backgrounds, variable illumination, overlapping leaves, mixed disease symptoms and far greater varietal and geographic diversity. Evidence on how well these models generalise once deployed outside their training distribution remains comparatively limited, and several reviewers have noted that dataset bias toward a small number of well documented crops and disease categories, largely originating from a limited set of geographic regions, constrains the external validity of reported accuracy figures (Kamilaris and Prenafeta-Boldú, 2018; Benos *et al*., 2021). The evidence base for laboratory accuracy is therefore strong, whereas the evidence base for field level diagnostic reliability across diverse commercial growing conditions remains considerably thinner.

3.2 Crop Yield Prediction

Yield prediction has become one of the most active subfields of agricultural artificial intelligence, drawing on remote sensing indices, weather records, soil characteristics and historical yield data. Chlingaryan *et al*. (2018) reviewed machine learning approaches to yield prediction and nitrogen status estimation and concluded that hybrid approaches combining remote sensing with machine learning offered particular promise for decision support, while van Klompenburg *et al*. (2020), conducting a systematic literature review of yield prediction studies, found that random forest and deep learning architectures were the most commonly reported high performing algorithms, but also observed considerable heterogeneity in the features, crops and evaluation metrics used across studies, which complicates direct comparison of reported accuracy. A recurring methodological weakness across this body of work is the tendency to report predictive accuracy, such as coefficient of determination or root mean square error, calculated on data drawn from the same season, region or even the same fields used for model training, rather than through temporally or geographically independent validation. Where studies do report cross regional or cross seasonal testing, predictive performance typically deteriorates relative to within sample results, which illustrates a broader and important distinction that this review develops further in section ten, namely that a model demonstrating strong statistical fit to historical data does not automatically provide farmers with a reliable forecast for a future and potentially atypical season. Furthermore, an accurate yield forecast is not, by itself, a farm management decision. Whether a farmer alters seed rate, fertiliser timing, irrigation scheduling or marketing strategy in response to a predicted yield depends on factors such as trust in the forecast, the cost of altering established practices and the presence of complementary advisory support, none of which are captured by predictive accuracy metrics alone (Finger *et al*., 2019).

3.3 Soil and Nutrient Management

Artificial intelligence has been applied to soil classification, nutrient status estimation, nitrogen management and site specific fertiliser recommendation, often in combination with remote sensing and proximal soil sensing. Chlingaryan *et al*. (2018) specifically highlighted nitrogen status estimation as an area where machine learning methods have shown strong potential to reduce over application of fertiliser while maintaining yield, and Sishodia *et al*. (2020) similarly identified nutrient application among the precision agriculture applications most directly supported by remote sensing derived vegetation indices. The environmental and economic rationale for these applications is compelling in principle, since more precise nutrient application can reduce input costs and limit nitrogen leaching and greenhouse gas emissions associated with over fertilisation (Balafoutis *et al*., 2017). However, the evidence for realised environmental and economic benefits at commercial farm scale is considerably more limited than the evidence for algorithmic prediction accuracy, and much of the supporting literature draws on experimental plots or demonstration farms rather than large samples of ordinary commercial operations, which restricts the generalisability of reported benefits.

3.4 Irrigation and Water Management

Artificial intelligence based irrigation scheduling, soil moisture forecasting and intelligent control systems have been reviewed extensively by Abioye *et al*. (2020), who examined a wide range of monitoring and advanced control strategies for precision irrigation and concluded that the integration of sensor networks with machine learning based decision algorithms could substantially improve irrigation timing relative to fixed schedule approaches. Sishodia *et al*. (2020) similarly identified irrigation management as one of the precision agriculture domains most directly supported by satellite and sensor based remote sensing. The potential water savings reported in individual experimental studies are often substantial, yet the reviewed literature indicates that many of these results derive from small scale trials conducted under specific soil, crop and climatic conditions, and that evidence on water and energy savings achieved when such systems are scaled across heterogeneous commercial farms remains comparatively sparse. This is a meaningful gap given that water scarcity is among the most pressing sustainability concerns motivating investment in artificial intelligence based irrigation in the first place (Finger *et al*., 2019).

3.5 Weed and Pest Management

Computer vision based weed recognition and targeted spraying have advanced rapidly, with Hasan *et al*. (2021) surveying deep learning techniques for weed detection from images and Rakhmatulin *et al*. (2021) reviewing real time deep neural network approaches to distinguishing weeds from crops under field conditions. Both reviews report high classification accuracy for weed species identification in benchmark datasets, and the underlying rationale for these systems, namely that precise weed identification enables targeted rather than blanket herbicide application, aligns closely with sustainability objectives around reduced chemical use. However, Rakhmatulin *et al*. (2021) explicitly note that translating high benchmark accuracy into reliable real time detection under variable field lighting, occlusion, growth stage variation and weed species diversity remains a substantial engineering and agronomic challenge, and Hasan *et al*. (2021) similarly emphasise that most reported systems have not been validated across the wide range of soil, crop and weed conditions found in commercial fields. The reviewed evidence therefore indicates that reductions in herbicide use are technically plausible and have been demonstrated in specific controlled trials, but that robust, generalisable evidence of pesticide reduction achieved at commercial scale across diverse farms remains limited.

3.6 Weather and Climate Risk Management

Artificial intelligence has been applied to short term weather prediction, drought monitoring and broader agricultural early warning systems, often through deep learning architectures applied to meteorological and remote sensing time series. This application area is closely tied to the agenda of climate resilient agriculture articulated by Basso and Antle (2020), who argue that digital agriculture technologies, properly integrated with policy and stakeholder engagement, can help balance the economic, environmental and social dimensions of sustainable food production. While drought and weather prediction models frequently report strong statistical performance in retrospective testing, the practical value of such systems for farm level resilience depends heavily on whether forecasts are delivered in a timely, comprehensible and actionable form to farmers, a condition that intersects directly with the extension and advisory challenges discussed later in this review.

3.7 Agricultural Robotics and Autonomous Systems

Robotic and autonomous systems have been comprehensively reviewed by Bechar and Vigneault (2016), who examined the concepts and components underlying agricultural robots for field operations, and by Fountas *et al*. (2020), who conducted a systematic review of commercial and research robotic systems and found that weeding and harvesting operations were the most extensively developed applications, while disease detection and seeding robots remained comparatively underdeveloped. Zhang *et al*. (2020) reviewed robotic grippers and grasping strategies specifically for agricultural manipulation tasks, noting that reliable grasping of delicate or irregularly shaped produce remains a significant unsolved engineering problem. The labour saving potential of agricultural robotics is intuitively appealing given persistent agricultural labour shortages in many regions, yet Fountas *et al*. (2020) caution that cost, reliability under variable field conditions and the complexity of integrating robotic systems into existing farm workflows continue to constrain commercial adoption, meaning that much of the published literature remains closer to proof of concept demonstration than to documented large scale deployment with measured productivity or cost outcomes.

3.8 Livestock and Poultry Farming

Artificial intelligence applications in livestock production span animal health monitoring, disease detection, behaviour recognition, precision feeding and production forecasting. Li *et al*. (2021) systematically reviewed convolutional neural network based computer vision systems in animal farming and found that classification, detection and tracking tasks had all been successfully demonstrated for cattle, sheep, pigs and poultry, while Mahmud *et al*. (2021) similarly reviewed deep learning applications specifically for precision cattle farming. Lovarelli *et al*. (2020) examined whether precision livestock farming represents an environmentally, economically and socially sustainable compromise for dairy cattle production, concluding that while individual technologies show clear potential for improving efficiency and animal welfare monitoring, comprehensive evidence of simultaneous environmental, economic and social benefit at commercial scale remains incomplete. Menezes *et al*. (2024) further note that while computer vision applications for dairy cattle have matured considerably, the application of large language models and natural language processing to livestock knowledge management remains at an early and largely exploratory stage. Neethirajan (2023) similarly highlights that sensor and artificial intelligence integration in livestock logistics and welfare monitoring is advancing quickly in research settings but that documented commercial scale evidence of welfare and efficiency gains remains limited relative to the pace of technical publication. Bao and Xie (2022), synthesising the broader systematic literature on artificial intelligence in animal farming, reach a similar conclusion, reporting that the volume of published technical solutions has grown considerably faster than the volume of studies evaluating farmer adoption, animal welfare outcomes or economic return under commercial operating conditions.

3.9 Post Harvest Management and Agricultural Supply Chains

Artificial intelligence applications in grading, sorting, quality assessment, storage optimisation and supply chain forecasting have received comparatively less systematic review attention than crop and livestock production applications, despite their direct relevance to reducing post harvest losses and improving farmer market access. Deichmann *et al*. (2016) note that digital technologies can improve agricultural supply chain management and market information flows in developing countries, though they caution that many promising pilot demonstrations of such digital dividends have not scaled to the extent anticipated, primarily because technology alone cannot resolve the institutional, infrastructural and market access barriers that constrain smallholder participation in modern supply chains. This observation anticipates a theme that recurs throughout the remainder of this review, namely that technological capability and realised agricultural benefit are analytically distinct and empirically dissociable.

4. Artificial Intelligence and Agricultural Decision Support

A central conceptual distinction that this review insists upon is the difference between prediction and decision support. A model that predicts crop yield, disease risk or optimal irrigation timing generates information, but a decision support system must additionally translate that information into a recommendation that a farmer can trust, understand and act upon within the constraints of their particular operation. Much of the reviewed literature on predictive modelling, including the yield prediction studies synthesised by van Klompenburg *et al*. (2020) and the disease detection studies synthesised by Kamilaris and Prenafeta-Boldú (2018), reports algorithmic performance in considerable technical detail while devoting comparatively little attention to the interface, communication and trust building elements required to convert a prediction into genuine decision support. Whether artificial intelligence recommendations replace or complement farmer knowledge is a matter of ongoing debate within the literature. Rose and Chilvers (2018) argue explicitly for a model of human artificial intelligence collaboration in which smart farming technologies broaden, rather than narrow, the range of considerations available to farmers, cautioning against approaches that treat farmers as passive recipients of algorithmic instruction. This perspective is consistent with evidence from agricultural extension research indicating that recommendations perceived as ignoring local agronomic knowledge or context specific constraints are frequently disregarded regardless of their underlying statistical accuracy (Steinke *et al*., 2021). The available evidence therefore supports the conclusion that farmer experience, local knowledge, agricultural extension services and artificial intelligence generated recommendations function most effectively as complementary rather than substitutable sources of agricultural knowledge, although rigorous comparative evaluation of purely algorithmic recommendations against human advisory services, or against genuinely integrated combinations of the two, remains surprisingly rare in the peer reviewed literature.

5. Artificial Intelligence and Agricultural Extension

Agricultural extension, the system of advisory services that connects agronomic research to farmer practice, has historically suffered from limited institutional capacity and reach, particularly in smallholder dominated agricultural systems (Aker, 2011). Artificial intelligence enabled advisory platforms, including mobile applications, chatbots, voice based advisory tools and recommendation systems, have been proposed as a means of extending the reach of extension services at comparatively low marginal cost. Steinke *et al*. (2021) reviewed the emerging innovation agenda for digital agricultural extension and concluded that many previous digital advisory initiatives failed not because the underlying technology was inadequate, but because designs were driven by available technology rather than by a rigorous understanding of farmers actual communication needs, literacy levels and trust requirements. Tzachor *et al*. (2023) examined the potential of large language models specifically as a natural language interface between complex agronomic knowledge systems and farmers, using real testing with Nigerian cassava farmers to illustrate both the promise of simplified, personalised agronomic advice and significant shortcomings, including the risk of factually inaccurate or contextually inappropriate recommendations being delivered with unwarranted apparent authority. De Clercq *et al*. (2024) similarly caution that while large language models could meaningfully expand access to agronomic knowledge, their tendency to generate plausible but incorrect information poses a genuine risk in a domain where poor advice can translate directly into crop loss or financial harm. Language diversity, digital literacy, connectivity, affordability and trust all constitute significant barriers to effective artificial intelligence enabled extension, and Steinke *et al*. (2021) explicitly warn that a narrow focus on technological sophistication, without corresponding attention to these barriers, risks widening rather than narrowing the digital divide that already separates well resourced and poorly resourced farming communities. The evidence base on artificial intelligence enabled extension is therefore promising in terms of demonstrated feasibility but still limited in terms of rigorous, independent evaluation of impact on farmer knowledge, practice change or ultimate agronomic outcomes. Oyinbo *et al*. (2020), examining the design preferences of extension agents themselves in Nigeria, add a further and frequently overlooked dimension to this discussion, finding that the professionals expected to mediate between digital decision support tools and farmers have their own distinct preferences regarding interface design, language and the level of technical detail provided, and that digital extension tools designed without consulting these intermediaries risk being underused even where farmers themselves might otherwise value the underlying information.

6. Artificial Intelligence and Smallholder Farmers

Whether artificial intelligence disproportionately benefits large commercial farms relative to smallholder producers is one of the most consequential and contested questions in this literature. Lowenberg-DeBoer and Erickson (2019), reviewing precision agriculture adoption patterns, found that adoption has historically concentrated among larger, better capitalised farms in developed economies, a pattern attributable to the cost of sensors, machinery and subscription services, the technical skills required for effective use and the greater absolute financial return that large scale operations can extract from marginal efficiency gains. Mizik (2023), conducting a systematic review focused specifically on precision farming at small scale, identified small land size, high adoption cost, technology related difficulties, lack of professional support and lack of supporting policy as the principal constraints preventing smallholder farmers from capturing the benefits demonstrated on larger commercial operations, and concluded that while environmental benefits of precision technologies may be broadly scale neutral, demonstrated economic benefits at small scale remain comparatively scarce in the literature. Deichmann *et al*. (2016) similarly argue that digital technologies can, in principle, overcome information asymmetries that disadvantage small scale farmers in output and input markets, but note that realised benefits have frequently fallen short of expectations because technology alone cannot resolve deeper structural barriers related to land tenure, access to finance, market infrastructure and institutional support. This distinction between technological availability and meaningful accessibility runs throughout the literature on artificial intelligence and development, and is directly relevant to research and policy attention on India, South Asia, Sub Saharan Africa and other smallholder dominated agricultural economies, where the majority of the world's farms are located but where the majority of published artificial intelligence agriculture research has not been conducted or validated (Aker, 2011; Steinke *et al*., 2021). Barrett and Rose (2022), examining public perceptions of what is often termed the fourth agricultural revolution, similarly report that farming communities themselves express considerable ambivalence about who is expected to benefit from these technologies, with concerns about corporate control of data and the displacement of smaller operations featuring prominently alongside enthusiasm for genuine productivity gains. Georgopoulos *et al*. (2023), reviewing the determinants of artificial intelligence adoption across agriculture, livestock farming and aquaculture using a PRISMA guided methodology, similarly conclude that perceived cost, perceived usefulness and institutional support consistently emerge as the strongest predictors of adoption across the studies they reviewed, reinforcing the conclusion that structural and institutional factors, rather than the technical sophistication of the artificial intelligence system itself, primarily determine which farmers are able to benefit.

Table 2 draws together the constraints and enabling conditions identified across this literature and contrasts them for smallholder and resource constrained farms relative to large, well capitalised commercial operations. The comparison illustrates that the barriers facing smallholders are rarely purely technical in nature, and instead cluster around cost, complementary infrastructure and institutional support, which suggests that purely technical solutions, however sophisticated, are unlikely by themselves to close the gap in realised benefit between these two groups of farmers.

Table 2. Comparative constraints and enabling conditions for artificial intelligence adoption among smallholder and large commercial farms

Factor	Smallholder and resource constrained farms	Large commercial farms	Source
Cost of adoption	High relative to farm income, often prohibitive	Comparatively small relative to farm turnover	Lowenberg-DeBoer and Erickson (2019); Mizik (2023)
Connectivity and infrastructure	Frequently limited or unreliable	Generally reliable and well supported	Steinke *et al.* (2021); Deichmann *et al.* (2016)
Digital and technical literacy	Variable and often low without targeted training	Generally higher, supported by dedicated staff	Georgopoulos *et al.* (2023); Aker (2011)
Access to extension and advisory support	Often thin or overstretched	Frequently supplemented by private advisory services	Steinke *et al.* (2021); Oyinbo *et al.* (2020)
Land and farm scale	Small and fragmented, limiting economies of scale	Large, enabling efficient use of fixed cost technologies	Lowenberg-DeBoer and Erickson (2019); Finger *et al.* (2019)
Institutional and policy support	Frequently limited or absent	Often embedded in established supply chain relationships	Mizik (2023); Georgopoulos *et al.* (2023)
Trust and perceived relevance	Shaped strongly by local language, culture and past experience of failed initiatives	Shaped by established vendor relationships and demonstrated return on investment	Barrett and Rose (2022); Wiseman *et al.* (2019)

7. Economic Implications of Artificial Intelligence in Agriculture

From an agricultural economics perspective, the central question is not whether artificial intelligence can improve the accuracy of agronomic predictions, but whether it produces measurable improvements in productivity, input efficiency, labour productivity, production cost, farm income, risk reduction and market access at the level of individual farm businesses. Finger *et al*. (2019), reviewing precision farming at the intersection of agricultural production and the environment, note that while precision technologies can reduce private variable costs through more targeted input use, current adoption remains concentrated among large farms in developed countries and that the economic case for smaller operations, particularly in developing countries, requires stronger public and private incentives to be realised. Weersink *et al*. (2018) similarly emphasise that big data and artificial intelligence hold substantial theoretical potential for agricultural and environmental analysis, but caution that data ownership, quality and interoperability issues continue to constrain the translation of analytical capability into demonstrated economic value for farm businesses. Cole and Fernando (2021) provide one of the more rigorous farm level economic evaluations in this literature, examining mobile based agricultural advisory services in India through experimental methods, and find that while such services can improve farmer knowledge and, in some circumstances, agricultural outcomes, sustained behavioural change and lasting income effects depend heavily on complementary factors such as trust, repeated engagement and local relevance of advice, rather than on information delivery alone. Overall, the reviewed literature indicates that long term, rigorously designed farm level economic evaluations of artificial intelligence adoption remain scarce relative to the volume of technical and algorithmic research, and that the distinction between theoretical economic potential, frequently emphasised in technical papers, and demonstrated economic outcomes, rarely measured with comparable rigour, constitutes one of the most significant gaps in the evidence base.

8. Artificial Intelligence and Sustainability

Artificial intelligence is frequently framed as inherently supportive of sustainable agriculture, on the basis that more precise targeting of water, fertiliser and pesticide inputs should reduce environmental externalities while maintaining or improving yield. Balafoutis *et al*. (2017) reviewed precision agriculture technologies with explicit reference to greenhouse gas mitigation potential and concluded that technologies enabling variable rate application of seed, fertiliser, pesticide and irrigation water can, under the conditions documented in the reviewed studies, reduce both greenhouse gas emissions and production costs simultaneously. However, this review resists the assumption that artificial intelligence is inherently or uniformly sustainable. Basso and Antle (2020) explicitly caution that digital agriculture must be deliberately designed around sustainability objectives rather than assumed to produce them as an automatic byproduct of increased data availability, arguing that balancing economic, environmental and social dimensions of sustainability requires active stakeholder engagement rather than technology deployment alone. Furthermore, the computational, sensor and connectivity infrastructure underpinning artificial intelligence systems, including data centres, drones and continuous sensor networks, carries its own energy and material footprint that is rarely quantified alongside the environmental benefits claimed for these systems in the agricultural literature reviewed here. The evidence for positive environmental outcomes is therefore genuine but conditional, resting on specific technologies, crops and management contexts rather than constituting a general property of artificial intelligence as applied to agriculture, and a comprehensive lifecycle accounting of the environmental costs and benefits of digital agricultural infrastructure remains largely absent from the reviewed literature.

9. Data Governance, Ethics, Trust and Responsible Artificial Intelligence

As artificial intelligence systems have become more deeply embedded in agricultural operations, questions of data ownership, farmer data rights, privacy, algorithmic bias, transparency, accountability and platform power have moved from peripheral to central concerns within the literature. Wiseman *et al*. (2019) examined farmers reluctance to share agricultural data and found that legal ambiguity over data ownership, combined with concern about how data might be used by input suppliers, financial institutions or competitors, constitutes a significant and rational barrier to the data sharing on which many artificial intelligence applications depend. Sullivan *et al*. (2024), conducting a systematic review of drivers and barriers to agricultural data sharing, similarly identified socio economic, systemic, technical and legal factors as jointly shaping farmers willingness to participate in data ecosystems, with technical and systemic barriers most frequently cited across the reviewed studies. Ryan (2019) examined the ethics of artificial intelligence and big data use by a large agricultural multinational and highlighted concerns around unequal bargaining power between farmers and large technology providers, while Dara *et al*. (2022) proposed a set of recommendations for ethical and responsible artificial intelligence use in digital agriculture, emphasising transparency, accountability and farmer inclusion in system design. Explainability represents a particularly important and underexplored dimension of this discussion. Adadi and Berrada (2018), in a widely cited general survey of explainable artificial intelligence, argue that the opacity of many high performing machine learning models constitutes a fundamental impediment to trust and adoption across domains, and this concern applies with particular force in agriculture, where farmers bear direct and immediate financial risk from acting on a recommendation whose underlying reasoning cannot be inspected or verified. The reviewed literature indicates that farmers are, understandably, hesitant to adopt systems whose recommendations cannot be understood, questioned or independently verified against their own agronomic knowledge, and that this hesitancy is a rational response to genuine informational asymmetry rather than a matter of simple technological resistance.

10. From Algorithmic Accuracy to Real World Agricultural Outcomes

This section develops the central critical argument of the review, namely that technological performance alone does not establish agricultural impact. It is analytically necessary to distinguish among several conceptually distinct levels of evidence, namely algorithmic performance, which refers to a model's statistical accuracy on a given dataset, operational performance, which refers to how reliably a system functions once deployed in real field conditions with all the noise, variability and edge cases that entails, farm level performance, which refers to whether use of the system changes farmer decisions and management practices, economic outcomes, which refer to measured effects on cost, yield value, income or profitability, environmental outcomes, which refer to measured effects on water use, input use, emissions or biodiversity, social outcomes, which refer to effects on farmer wellbeing, knowledge, autonomy and equity, and system level outcomes, which refer to aggregate effects on food security, rural livelihoods and agricultural sustainability at regional or national scale.

The reviewed literature demonstrates a systematic asymmetry across these levels of evidence. Algorithmic performance is reported extensively and, in domains such as image based disease classification, is often genuinely impressive, with accuracy figures regularly exceeding ninety percent under benchmark conditions (Mohanty *et al*., 2016; Ferentinos, 2018). Operational performance, meaning reliability under the full range of field conditions encountered on commercial farms, is documented far less consistently, and where it is examined, as in Hasan *et al*. (2021) and Rakhmatulin *et al*. (2021) on weed detection or Fountas *et al*. (2020) on field robotics, researchers frequently report meaningful degradation relative to benchmark performance and identify substantial remaining engineering challenges. Farm level performance, meaning whether farmers actually change their decisions and practices in response to artificial intelligence outputs, is addressed even less frequently, and where addressed, as in the extension and advisory literature reviewed by Steinke *et al*. (2021) and Cole and Fernando (2021), the evidence suggests that trust, communication design and local relevance matter as much as, or more than, the statistical accuracy of the underlying model. Economic, environmental and social outcomes, meaning the ultimate justification typically offered for investment in agricultural artificial intelligence, are addressed with the least frequency and the least methodological rigour of all, with Finger *et al*. (2019) and Weersink *et al*. (2018) both explicitly noting the scarcity of long term, farm level economic evaluation relative to the volume of algorithmic and technical publication.

A ninety five percent classification accuracy figure, to take the illustrative example embedded in the review's central argument, describes the proportion of correctly labelled instances in a specific test set, and carries no inherent implication for how much a farmer's yield, cost or income will change upon adopting the corresponding technology, because that translation depends on additional and independent factors including the cost of the intervention prompted by the recommendation, the farmer's capacity to act on it, the baseline practice being replaced and the broader agronomic and market context in which the farm operates. This is not a claim that algorithmic accuracy is unimportant, since a highly inaccurate model is unlikely to produce beneficial outcomes under any circumstances, but rather a claim that accuracy is a necessary and not a sufficient condition for agricultural impact. Field validation, external validation across independent geographies, temporal validation across multiple growing seasons and farmer level trials involving real behavioural response constitute the evidentiary standards that would allow confident inference from algorithmic performance to agricultural outcome, and the reviewed literature indicates that these standards are met inconsistently and, in several thematic areas including robotics, livestock artificial intelligence and generative artificial intelligence for extension, are met only rarely. The evidence therefore points toward a persistent and consequential gap between proof of concept demonstration, which dominates the published literature, and scalable, validated agricultural deployment, which remains comparatively uncommon. This asymmetry in the volume of published evidence across the outcome hierarchy is illustrated in Figure 3.

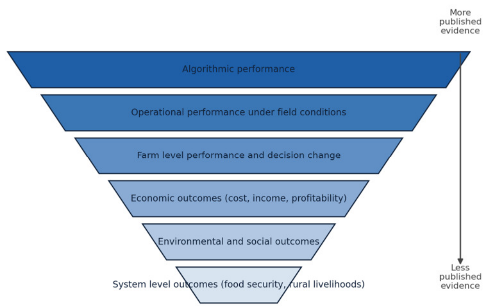

Figure 3. Asymmetry in the volume of published evidence across levels of the outcome hierarchy identified in this review. The funnel narrows sharply below the level of operational performance, indicating that rigorous evidence on farm level, economic, environmental, social and system level outcomes is considerably scarcer than evidence on algorithmic accuracy.

10.1 Cross Domain Appraisal of Evidence Strength

Building on the distinction developed above, Table 3 presents a critical appraisal of the evidence available across the nine application domains reviewed in section three, rating the strength of evidence separately for algorithmic performance, field or operational validation and documented economic or farm level outcome. This appraisal reflects the qualitative judgement of this review, formed through the process of critical synthesis described in the methodology, rather than a formal quantitative meta analytic score, since the underlying studies employ evaluation metrics and study designs that are too heterogeneous to support direct numerical pooling. The purpose of the table is to make explicit, in a single comparative view, an argument that otherwise risks remaining implicit across many pages of narrative discussion, namely that the strength of evidence available to researchers, policymakers and investors declines sharply and unevenly as one moves from algorithmic claims toward claims about agricultural impact.

Table 3. Critical appraisal of evidence strength across artificial intelligence application domains in agriculture

Application domain	Algorithmic performance evidence	Field or operational validation evidence	Farm level economic or outcome evidence	Principal limitation identified in this review
Crop disease detection	Strong	Moderate	Limited	Training data dominated by controlled photographic conditions and narrow crop or disease diversity
Crop yield prediction	Strong	Moderate	Limited	Within sample validation common, cross seasonal and cross regional testing less common
Soil and nutrient management	Moderate	Limited	Limited	Evidence concentrated in experimental plots rather than commercial farms
Irrigation and water management	Moderate	Limited	Limited	Water and cost savings documented mainly in small scale trials
Weed and pest management	Strong	Limited	Limited	High benchmark accuracy but limited testing under variable field lighting and occlusion
Agricultural robotics	Moderate	Limited	Very limited	Cost and reliability constraints restrict most evidence to demonstration rather than commercial deployment
Livestock and poultry farming	Moderate to strong	Moderate	Limited	Computer vision applications more mature than natural language or economic evaluation components
Agricultural extension and generative artificial intelligence	Limited to moderate	Limited	Very limited	Field testing largely exploratory, with risk of confidently delivered but incorrect advice
Smallholder and economic outcomes across domains	Not applicable	Limited	Very limited	Long term, rigorously designed farm level economic evaluation remains rare across the entire literature

10.2 Recurring Methodological Weaknesses Across the Reviewed Literature

Beyond the asymmetry documented above, several recurring methodological weaknesses cut across the thematic domains reviewed in this article and, taken together, help explain why algorithmic accuracy has proved a poor proxy for agricultural impact. First, many of the datasets underpinning image based classification tasks, whether for disease detection or weed identification, are collected under standardised photographic conditions and drawn disproportionately from a limited number of countries, crop varieties and research institutions, which constrains the external validity of reported accuracy figures when models are transferred to unfamiliar agroecological contexts (Kamilaris and Prenafeta-Boldú, 2018; Hasan *et al*., 2021). Second, predictive modelling studies for yield and related agronomic outcomes frequently evaluate performance on data drawn from the same season, region or even the same fields used for model training, a practice that inflates apparent accuracy relative to genuinely independent, out of sample prediction, and the comparatively few studies that do test across seasons or regions typically report weaker performance than within sample evaluation would suggest (van Klompenburg *et al*., 2020). Third, studies evaluating farmer adoption or the impact of digital advisory services frequently rely on samples of farmers who have already chosen to adopt a technology, introducing self selection bias that likely overstates the benefits attributable to the technology itself rather than to the underlying characteristics, such as capital access or risk tolerance, that led certain farmers to adopt in the first place, a concern that is only partially addressed by the more rigorous experimental designs exemplified by Cole and Fernando (2021). Fourth, comparatively few reviewed studies employ control groups, randomisation or other design features that would support causal rather than merely associational claims about the effect of artificial intelligence adoption on farm outcomes, which limits the confidence with which the literature as a whole can attribute observed differences to the technology under study. Fifth, the geographic distribution of published research remains heavily concentrated in a comparatively small number of countries, with Sishodia *et al*. (2020) and related remote sensing literature drawing predominantly on cases from North America, Europe, China and India, a pattern that leaves large parts of Sub Saharan Africa, Latin America and Southeast Asia considerably underrepresented despite hosting a substantial share of the world's smallholder farms (Mizik, 2023; Deichmann *et al*., 2016). Sixth, much of the technical literature reviewed here is authored by computer science or engineering teams with comparatively limited direct engagement with agricultural economics or extension expertise, which may help explain why farm level economic and behavioural outcomes are addressed with less methodological rigour than algorithmic performance, since the disciplinary training and incentive structures of these two research communities differ considerably. Finally, publication practices in fast moving technical fields tend to favour reporting of successful model performance over null or negative results, and while this review cannot quantify the extent of such publication bias within agricultural artificial intelligence specifically, its plausibility is supported by the observation that very few reviewed studies report cases in which artificial intelligence recommendations were tested against, and found inferior to, existing farmer or extension agent practice.

10.3 Points of Agreement and Disagreement Across the Literature

Despite this critical appraisal, meaningful convergence exists across parts of the literature. There is broad agreement that deep learning architectures outperform earlier machine learning and conventional image processing techniques on benchmark classification and detection tasks (Kamilaris and Prenafeta-Boldú, 2018; Liakos *et al*., 2018; Benos *et al*., 2021), and broad agreement that precision agriculture technologies can, under favourable and well documented conditions, reduce input use while maintaining yield (Balafoutis *et al*., 2017; Finger *et al*., 2019). There is also convergence around the observation that adoption remains concentrated among larger, better capitalised farms (Lowenberg-DeBoer and Erickson, 2019; Mizik, 2023; Georgopoulos *et al*., 2023), and around the concern that data governance and trust constitute significant unresolved barriers to wider adoption (Wiseman *et al*., 2019; Sullivan *et al*., 2024; Ryan, 2019). Disagreement is more pronounced, however, regarding the practical significance of these technologies for smallholder farmers. Deichmann *et al*. (2016) adopt a cautiously optimistic position, arguing that digital technologies can meaningfully reduce information asymmetries facing small farmers even where broader structural barriers persist, whereas Mizik (2023) reaches a more guarded conclusion, emphasising that demonstrated economic benefit at small scale remains scarce relative to the volume of claims made about smallholder relevant artificial intelligence. Similarly, while Tzachor *et al*. (2023) present generative artificial intelligence as offering genuine potential to extend agronomic advice to underserved farmers, De Clercq *et al*. (2024) emphasise the risks of confidently delivered but factually incorrect advice with sufficient force to suggest considerably greater caution before such tools are deployed at scale. The reviewed literature remains inconclusive, in other words, not because researchers have failed to study these questions, but because the studies that exist vary widely in context, methodology and underlying assumptions about what constitutes meaningful evidence of benefit, a conclusion that itself constitutes an important finding of this review.

11. Major Research Gaps

Several substantial research gaps emerge from this synthesis. Longitudinal studies that track the same farms or farmers over multiple seasons following artificial intelligence adoption remain rare, which means that the durability of any observed benefits, and the possibility that initial gains erode as novelty effects fade or as farmers revert to prior practices, cannot currently be assessed with confidence. Farmer level impact evaluation using rigorous experimental or quasi experimental designs, of the kind exemplified by Cole and Fernando (2021), remains the exception rather than the norm across the wider artificial intelligence agriculture literature, and this matters because observational or laboratory based evidence cannot reliably establish causal effects on farmer decisions or outcomes. Economic evaluation at farm level is similarly underdeveloped, with Finger *et al*. (2019) and Weersink *et al*. (2018) both highlighting this as a significant constraint on confident policy or investment guidance. Field validation of computer vision and predictive models across diverse geographies, crop varieties and seasons remains limited relative to the volume of algorithmic development, a gap that matters because the reviewed disease detection and yield prediction literature indicates that performance can deteriorate substantially outside training conditions (Kamilaris and Prenafeta-Boldú, 2018;

van Klompenburg *et al*., 2020). Geographic concentration of both datasets and published research toward a comparatively small number of countries and crop systems limits the representativeness of the evidence base, and this concentration compounds the already limited representation of smallholder farming systems in the literature, despite smallholders constituting the majority of farms globally (Mizik, 2023; Deichmann *et al*., 2016). Multilingual and locally adapted artificial intelligence extension systems remain scarce relative to the linguistic diversity of farming populations, a gap that Steinke *et al*. (2021) identify as a direct barrier to equitable access. Explainability of agricultural artificial intelligence systems remains underdeveloped relative to the general explainable artificial intelligence literature (Adadi and Berrada, 2018), which matters because farmer trust and willingness to adopt appear closely tied to the ability to understand and verify recommendations. Interoperability among agricultural data systems, standardisation of datasets and integration of artificial intelligence tools with existing agricultural extension infrastructure all remain limited, constraining the scalability of otherwise promising pilot demonstrations. Evidence on long term sustainability outcomes, gender inclusion, institutional adoption and the policy environment surrounding agricultural artificial intelligence is sparse relative to the volume of purely technical publication, and research specifically addressing generative artificial intelligence in agricultural contexts remains at an early and largely exploratory stage (Tzachor *et al*., 2023; De Clercq *et al*., 2024).

12. Future Research Directions

A forward looking research agenda should prioritise several interconnected directions. Explainable artificial intelligence tailored specifically to agricultural decision contexts is needed to help farmers understand, question and verify algorithmic recommendations rather than accepting or rejecting them without insight into their underlying logic, building on general explainable artificial intelligence research (Adadi and Berrada, 2018) but adapted to the specific risk tolerances and knowledge systems of farming communities. Generative artificial intelligence research should move beyond feasibility demonstrations toward rigorous evaluation of accuracy, safety and contextual appropriateness when deployed as agronomic advisory tools, addressing directly the risks of confidently delivered but incorrect information identified by De Clercq *et al*. (2024) and Tzachor *et al*. (2023). Multimodal artificial intelligence, combining image, sensor, text and market data within unified decision support frameworks, offers a promising direction for overcoming the fragmentation that currently characterises many single purpose agricultural artificial intelligence tools. Edge artificial intelligence and federated learning approaches, which allow models to be trained and run with reduced dependence on continuous connectivity and centralised data pooling, are particularly relevant to the connectivity and data ownership constraints documented in smallholder and data governance research (Wiseman *et al*., 2019; Mizik, 2023). Closer integration of artificial intelligence with the Internet of Things and remote sensing infrastructure, alongside emerging digital twin approaches that simulate whole farm systems, could support more holistic decision making than isolated point predictions currently allow. Human artificial intelligence collaboration, farmer centred design and participatory development processes should be treated as core methodological requirements rather than optional considerations, addressing the concern raised by Rose and Chilvers (2018) that technology led design risks marginalising the very farmers such systems are intended to serve. Multilingual agricultural artificial intelligence and its integration with existing extension services deserve sustained attention given the barriers documented by Steinke *et al*. (2021) and Aker (2011). Responsible and low cost artificial intelligence approaches specifically oriented toward smallholder and resource constrained farming systems represent an urgent priority given the adoption disparities documented by Lowenberg-DeBoer and Erickson (2019) and Mizik (2023). Finally, agricultural economics research applying rigorous causal inference methods to artificial intelligence adoption, of the kind demonstrated by Cole and Fernando (2021), should be expanded substantially relative to the current volume of purely algorithmic and technical research, since without such evidence the field cannot move beyond plausible conjecture about economic impact toward confident, policy relevant conclusions.

13. Proposed Conceptual Framework

Drawing together the evidence and arguments developed across this review, a conceptual framework can be articulated that traces the pathway from agricultural data through artificial intelligence algorithms to prediction, decision support, farmer decision, management intervention, farm outcome, and ultimately economic, environmental and social outcomes. This pathway is not linear or automatic at any stage, and its progression is shaped throughout by a set of moderating factors that the reviewed literature repeatedly identifies as decisive, including farmer capability and digital literacy, trust in the system and its developers, the availability and quality of complementary extension support, physical and connectivity infrastructure, the broader institutional and policy environment, farm size and available capital, the cost of adoption relative to expected benefit, and the quality and representativeness of the underlying data. A model may achieve excellent algorithmic performance yet fail to influence farmer decisions if trust is absent or if the recommendation cannot be integrated into existing farm workflows, exactly as the extension literature demonstrates (Steinke *et al*., 2021). Equally, a farmer may act on a recommendation, yet realise no economic or environmental benefit if the underlying prediction was poorly calibrated to local conditions, exactly as the yield prediction and irrigation literature suggests can occur when models are transferred outside their original geographic and climatic context (van Klompenburg *et al*., 2020; Abioye *et al*., 2020). This framework therefore positions artificial intelligence not as a self executing driver of agricultural transformation but as one input, whose ultimate effect on productivity, profitability, resource efficiency, sustainability, resilience and farmer empowerment is contingent on the surrounding socio technical, economic and institutional system within which it is deployed. Figure 4 presents this framework graphically, showing the central data to outcome chain together with the moderating factors that the reviewed literature repeatedly identifies as decisive.

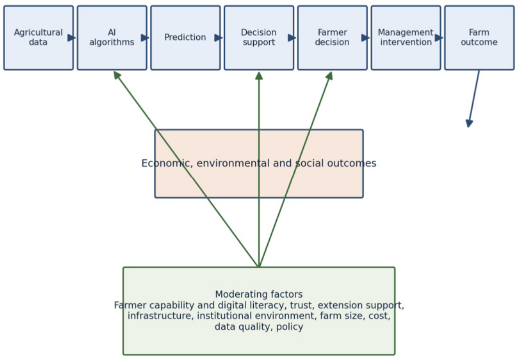

Figure 4. Proposed conceptual framework linking agricultural data, through artificial intelligence algorithms, prediction and decision support, to farmer decisions, management interventions, farm outcomes and ultimately economic, environmental and social outcomes, moderated throughout by farmer capability, trust, extension support, infrastructure, institutional environment, farm size, cost, data quality and policy.

14. Discussion

The preceding sections have synthesised a large and technically sophisticated literature, and this discussion draws together the principal critical threads in order to situate the findings of this review within the broader debate about digital agriculture. Three overarching observations merit particular emphasis. The first is that the artificial intelligence agriculture literature exhibits a pronounced and consistent gradient of evidentiary strength that runs from algorithm to outcome, a pattern documented across every thematic domain examined in section three and formalised in the appraisal presented in Table 3. This gradient is not unique to agriculture, since similar patterns have been observed in other applied domains where technically impressive models struggle to demonstrate commensurate real world impact, but it carries particular weight in agriculture because farmers, unlike many other end users of artificial intelligence, bear direct and often immediate financial and food security consequences when a recommendation proves unreliable. The second overarching observation is that the debate over who benefits from agricultural artificial intelligence cannot be resolved through technical improvement alone. The evidence reviewed in sections six and seven indicates that adoption disparities between large and small farms are driven substantially by cost, infrastructure, institutional support and trust rather than by the technical sophistication of the algorithms themselves, which implies that continued investment in ever more accurate models, absent parallel investment in extension, affordability and locally adapted design, is unlikely by itself to narrow this gap and could plausibly widen it. The third overarching observation is that the field currently lacks a shared standard for what constitutes adequate evidence of agricultural impact. Some studies treat statistical accuracy on a held out test set as sufficient evidence of value, while others insist on farm level economic measurement, and the absence of a widely accepted evidentiary standard, comparable to the hierarchies of evidence used in clinical and public health research, makes it difficult to compare findings across studies or to judge when a given artificial intelligence application has accumulated sufficient evidence to justify wide scale promotion or investment.

These observations carry practical implications for the principal audiences of this review. For researchers, the analysis suggests that the marginal scientific value of yet another crop disease classifier trained and tested on an existing benchmark dataset is now considerably lower than the marginal value of rigorous field validation, farmer level trials or long term economic evaluation of tools that already exist, a rebalancing of research effort that the current volume of technical publication relative to outcome evaluation does not yet reflect. For technology developers and agricultural technology companies, the analysis suggests that investment in explainability, in transparent data governance and in genuine co design with farmers and extension agents is likely to matter as much for eventual adoption and impact as further improvement in predictive accuracy, particularly given the trust related barriers documented in section nine. For policymakers and donors, the analysis suggests that funding decisions should distinguish carefully between technologies that have accumulated genuine field level and economic evidence and those that remain, however promising, at the algorithmic proof of concept stage, and that public investment may be particularly justified in precisely the areas the market is least likely to address on its own, namely field validation across diverse smallholder contexts, multilingual and low cost adaptation and independent evaluation of claimed benefits. For extension services, the analysis suggests that artificial intelligence tools are most likely to add value when integrated into, rather than substituted for, existing advisory relationships, consistent with the human artificial intelligence collaboration model advocated by Rose and Chilvers (2018) and the user centred design principles articulated by Steinke *et al*. (2021).

Table 4 consolidates the research gaps identified throughout this review into a structured agenda, indicating for each gap what is currently known, what remains unknown, and the type of study design that would be needed to close it. This table is intended to function as a practical reference for researchers and funders seeking to identify where additional evidence would carry the greatest value relative to the volume of research already available.

Table 4. Research gap matrix and corresponding evidence needs

Research gap	What is currently known	What remains unknown	Study design needed to close the gap
Longitudinal farm level impact	Short term trials and pilot demonstrations exist for several technologies	Whether benefits persist, grow or erode over multiple seasons	Multi season panel studies following the same farms over time
Farmer level causal impact evaluation	Descriptive and observational adoption studies are common	Whether observed associations reflect causal effects of the technology itself	Randomised controlled trials or well identified quasi experimental designs, as in Cole and Fernando (2021)
Farm level economic evaluation	Theoretical cost savings from targeted input use are well documented	Realised income and profitability effects at commercial farm scale	Farm level cost benefit analysis linked to independent yield and expenditure records
Field validation across geography and season	Benchmark accuracy is well documented within single studies	External validity across independent regions, climates and cropping systems	Cross regional and multi season validation studies with pre registered evaluation protocols
Smallholder representation	Barriers to smallholder adoption are conceptually well described	Documented outcomes for smallholders who have adopted artificial intelligence tools	Field studies conducted in under represented regions including Sub Saharan Africa, Latin America and Southeast Asia
Explainability and trust	General explainable artificial intelligence methods are well developed in computer science	How explainability affects farmer trust and behaviour change in practice	User studies examining farmer comprehension, trust and decision change under varying levels of explanation
Generative artificial intelligence safety in extension	Feasibility of large language model based advisory tools has been demonstrated	Accuracy, safety and contextual appropriateness under real deployment conditions	Controlled field trials comparing generative artificial intelligence advice against expert extension agent advice
Institutional and policy environment	Case level evidence of institutional barriers exists	Systematic, cross country evidence on which policy interventions most effectively support equitable adoption	Comparative policy analysis across countries with differing digital agriculture strategies

15. Conclusion

This review set out to examine whether reported improvements in artificial intelligence performance within agriculture correspond to demonstrated improvements in agricultural outcomes, and the evidence supports a carefully qualified answer. Confident conclusions can be drawn regarding algorithmic performance, since deep learning and machine learning methods have repeatedly demonstrated strong predictive and classification capability across crop disease detection, yield forecasting, weed identification and related tasks under benchmark or controlled conditions. Confidence can also be extended, with somewhat greater caution, to the feasibility of automation and sensing technologies, since robotic, remote sensing and Internet of Things systems have been successfully demonstrated across a wide range of field operations, even where commercial scale reliability remains under active development. What the evidence does not yet support, however, is a confident general claim that these algorithmic and technical achievements have translated into widespread, well documented improvements in farm productivity, profitability, resource efficiency or resilience, nor a confident claim that the benefits of artificial intelligence in agriculture are being shared equitably between large commercial operations and smallholder or resource constrained farmers. Much of the published literature remains concentrated at the algorithmic proof of concept stage, with rigorous field validation, farmer level economic evaluation and attention to data governance, explainability and equity trailing considerably behind the pace of technical innovation. This is not a case for scepticism toward artificial intelligence in agriculture as such, since the underlying technical achievements documented throughout this review are genuine and, in several domains, substantial, but it is a case against treating algorithmic accuracy as a proxy for agricultural progress. Realising the transformative potential that is so frequently attributed to artificial intelligence in agriculture will require sustained investment in farmer centred design, longitudinal and geographically diverse field validation, transparent and responsible data governance, and closer integration between artificial intelligence developers, agricultural extension systems and the farmers whose knowledge, trust and decisions ultimately determine whether any algorithm makes a genuine difference in the field.

Conflict of interest: The authors declare no conflict of interest.

Data availability: No new primary data were generated for this review. All synthesised sources are cited in the reference list.

References

Abioye, E. A., Abidin, M. S. Z., Mahmud, M. S. A., Buyamin, S., Ishak, M. H. I., Rahman, M. K. I. A., Otuoze, A. O., Onotu, P., & Ramli, M. S. A. (2020). A review on monitoring and advanced control strategies for precision irrigation. Computers and Electronics in Agriculture, 173, 105441.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE access, 6, 52138-52160.

Aker, J. C. (2011). Dial "A" for agriculture: A review of information and communication technologies for agricultural extension in developing countries. Agricultural Economics, 42(6), 631-647.

Balafoutis, A., Beck, B., Fountas, S., Vangeyte, J., Van der Wal, T., Soto, I., Gómez-Barbero, M., Barnes, A., & Eory, V. (2017). Precision agriculture technologies positively contributing to greenhouse gas emissions mitigation, farm productivity and economics. Sustainability, 9(8), 1339.

Bao, J., & Xie, Q. (2022). Artificial intelligence in animal farming: A systematic literature review. Journal of Cleaner Production, 331, 129956.

Barrett, H., & Rose, D. C. (2022). Perceptions of the fourth agricultural revolution: what is in, what is out, and what consequences are anticipated. Sociologia Ruralis, 62(2), 162-189.

Basso, B., & Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. Nature sustainability, 3(4), 254-256.

Bechar, A., & Vigneault, C. (2016). Agricultural robots for field operations: Concepts and components. Biosystems engineering, 149, 94-111.

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: a comprehensive updated review. Sensors, 21(11), 3758.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and electronics in agriculture, 151, 61-69.

Cole, S. A., & Fernando, A. N. (2021). 'Mobile'izing agricultural advice: technology adoption, diffusion and sustainability. The Economic Journal, 131(633), 192-219.

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., & Traore, D. (2022). Deep learning for precision agriculture: a bibliometric analysis. Intelligent Systems with Applications, 16, 200102.

Dara, R., Hazrati Fard, S. M., & Kaur, J. (2022). Recommendations for ethical and responsible use of artificial intelligence in digital agriculture. Frontiers in Artificial Intelligence, 5, 884192.

De Clercq, D., Nehring, E., Mayne, H., & Mahdi, A. (2024). Large language models can help boost food production, but be mindful of their risks. Frontiers in Artificial Intelligence, 7, 1326153.

Deichmann, U., Goyal, A., & Mishra, D. (2016). Will digital technologies transform agriculture in developing countries?. Agricultural economics, 47(S1), 21-33.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311-318.

Finger, R., Swinton, S. M., El Benni, N., & Walter, A. (2019). Precision farming at the nexus of agricultural production and the environment. Annual Review of Resource Economics, 11, 313-335.

Fountas, S., Mylonas, N., Malounas, I., Rodias, E., Hellmann-Santos, C., & Pekkeriet, E. (2020). Agricultural robotics for field operations. Sensors, 20(9), 2672.

Georgopoulos, V. P., Gkikas, D. C., & Theodorou, J. A. (2023). Factors influencing the adoption of artificial intelligence technologies in agriculture, livestock farming and aquaculture: a systematic literature review using PRISMA 2020. Sustainability, 15(23), 16385.

Hasan, A. S. M. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. K. (2021). A survey of deep learning techniques for weed detection from images. Computers and Electronics in Agriculture, 184, 106067.

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

Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. NJAS: wageningen journal of life sciences, 90(1), 1-16.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

Li, G., Huang, Y., Chen, Z., Chesser Jr, G. D., Purswell, J. L., Linhoss, J., & Zhao, Y. (2021). Practices and applications of convolutional neural network-based computer vision systems in animal farming: A review. Sensors, 21(4), 1492.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.

Lovarelli, D., Bacenetti, J., & Guarino, M. (2020). A review on dairy cattle farming: Is precision

livestock farming the compromise for an environmental, economic and social sustainable production?. Journal of Cleaner Production, 262, 121409.
Lowenberg-DeBoer, J., & Erickson, B. (2019). Setting the record straight on precision agriculture adoption. Agronomy journal, 111(4), 1552-1569.
Lu, Y., & Young, S. (2020). A survey of public datasets for computer vision tasks in precision agriculture. Computers and Electronics in Agriculture, 178, 105760.
Mahmud, M. S., Zahid, A., Das, A. K., Muzammil, M., & Khan, M. U. (2021). A systematic literature review on deep learning applications for precision cattle farming. Computers and Electronics in Agriculture, 187, 106313.
Menezes, G. L., Mazon, G., Ferreira, R. E., Cabrera, V. E., & Dorea, J. R. (2024). Artificial intelligence for livestock: a narrative review of the applications of computer vision systems and large language models for animal farming. Animal Frontiers, 14(6), 42-53.
Mizik, T. (2023). How can precision farming work on a small scale? A systematic literature review. Precision agriculture, 24(1), 384-406.
Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in plant science, 7, 215232.
Neethirajan, S. (2023). Artificial intelligence and sensor technologies in dairy livestock export: charting a digital transformation. Sensors, 23(16), 7045.
Oyinbo, O., Chamberlin, J., & Maertens, M. (2020). Design of digital agricultural extension tools: perspectives from extension agents in Nigeria. Journal of agricultural economics, 71(3), 798-815.
Rakhmatulin, I., Kamilaris, A., & Andreasen, C. (2021). Deep neural networks to detect weeds from crops in agricultural environments in real-time: A review. Remote Sensing, 13(21), 4486.
Rejeb, A., Rejeb, K., Abdollahi, A., & Hassoun, A. (2024). Precision agriculture: A bibliometric analysis and research agenda. Smart Agricultural Technology, 9, 100684.
Rose, D. C., & Chilvers, J. (2018). Agriculture 4.0: Broadening responsible innovation in an era of smart farming. Frontiers in Sustainable Food Systems, 2, 387545.
Ryan, M. (2019). Ethics of using artificial intelligence and big data in agriculture: the case of a large agriculture multinational. The ORBIT Journal, 2(2), 1-27.
Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2020). Machine learning applications for precision agriculture: A comprehensive review. IEEE access, 9, 4843-4873.
Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. Remote sensing, 12(19), 3136.
Steinke, J., Van Etten, J., Müller, A., Ortiz-Crespo, B., van de Gevel, J., Silvestri, S., & Priebe, J. (2021). Tapping the full potential of the digital revolution for agricultural extension: an emerging innovation agenda. International Journal of Agricultural Sustainability, 19(5-6), 549-565.
Sullivan, C. S., Gemtou, M., Anastasiou, E., & Fountas, S. (2024). Building trust: A systematic review of the drivers and barriers of agricultural data sharing. Smart Agricultural Technology, 8, 100477.
Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 161, 272-279.
Tzachor, A., Devare, M., Richards, C., Pypers, P., Ghosh, A., Koo, J., Johal, S., & King, B. (2023). Large language models and agricultural extension services. Nature Food, 4(11), 941-948.
Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and electronics in agriculture, 177, 105709.
Weersink, A., Fraser, E., Pannell, D., Duncan, E., & Rotz, S. (2018). Opportunities and challenges for big data in agricultural and environmental analysis. Annual Review of Resource Economics, 10(1), 19-37.
Wiseman, L., Sanderson, J., Zhang, A., & Jakku, E. (2019). Farmers and their data: An examination of farmers' reluctance to share their data through the lens of the laws impacting smart farming. NJAS: Wageningen Journal of Life Sciences, 90(1), 1-10.
Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming–a review. Agricultural systems, 153, 69-80.
Zhang, B., Xie, Y., Zhou, J., Wang, K., & Zhang, Z. (2020). State-of-the-art robotic grippers, grasping and control strategies, as well as their applications in agricultural robots: A review. Computers and Electronics in Agriculture, 177, 105694.