

REVIEW

Open access

Foundation Models for Smart Farming and Precision Agriculture: A Systematic Review of Opportunities, Risks, Technical Architectures, and Research Directions

Maria Gonzales^{1*}, Jose Santos¹, Anna Cruz², Carlos Reyes³

Abstract

Foundation models are increasingly investigated as reusable architecture-level systems for agricultural sensing, remote-sensing interpretation, multimodal learning, knowledge access, phenotyping, monitoring, and decision-support assistance. However, the evidence remains dispersed across model families, data modalities, agricultural applications, and validation settings. This systematic review synthesises the technical architectures, data requirements, opportunities, risks, deployment challenges, evidence gaps, and research priorities associated with foundation models in smart farming and precision agriculture. Peer-reviewed journal literature was identified through structured database and publisher-platform searches, supplementary reference checking, predefined eligibility criteria, sequential screening, claim-specific evidence extraction, and thematic synthesis. Evidence was organised across language, vision, vision-language, multimodal, geospatial, remote-sensing, time-series, self-supervised, transfer-learning, and domain-adaptation approaches. The reviewed literature indicates that foundation-model approaches may support transferable representation learning, reduced annotation dependence, multimodal data integration, crop and livestock monitoring, remote-sensing analysis, plant phenotyping, agricultural knowledge services, and bounded decision-support workflows. Nevertheless, reported results are commonly task-, dataset-, region-, sensor-, or application-specific. Important limitations include agricultural data scarcity and heterogeneity, geographic and temporal domain shift, uncertain generalisation, bias, weak explainability, inadequate uncertainty communication, privacy and cybersecurity risks, unclear accountability, computing demands, and limited farmer-centred or operational evaluation. Priorities include agricultural benchmark suites, domain-specific pretraining resources, multimodal and spatiotemporal adaptation, low-resource representation, robustness testing, calibrated uncertainty, human-in-the-loop evaluation, privacy-preserving development, sustainable-computing assessment, and lifecycle governance. Foundation models represent a promising research direction, but benchmark performance or pretraining scale alone does not establish agronomic validity, operational safety, farmer trust, or deployment readiness.

Keywords Precision agriculture, Remote sensing, Smart farming, Foundation models, Agricultural AI, Multimodal learning

*Correspondence:

Maria Gonzales

maria.gonzales@up.edu.ph

¹ Department of Foundation Models and Agricultural AI, Faculty of Agriculture, University of the Philippines Los Baños, Laguna, Philippines

² Department of AI Opportunities and Risks in Agriculture, Faculty of Agriculture, University of Santo Tomas, Manila, Philippines

³ Department of Agricultural Research Directions, Faculty of Agriculture, Ateneo de Manila University, Quezon City, Philippines

Introduction

Foundation models are large-scale pretrained artificial intelligence systems designed to learn reusable representations that can be adapted to multiple downstream tasks through fine-tuning, prompting, transfer learning, or other forms of domain adaptation. Within smart farming and precision agriculture, these systems are being investigated for image interpretation, agricultural language processing, multimodal data integration, remote-sensing analysis, plant phenotyping, crop monitoring, knowledge discovery, livestock observation, and decision-support assistance. Their significance lies not merely in model size, but in an architecture-level shift from independently trained task-specific systems toward broadly pretrained models intended to support multiple agricultural applications. Nevertheless, agricultural usefulness depends on the quality and representativeness of pretraining data, the suitability of adaptation procedures, and the strength of task- and domain-specific validation [1].

The emergence of agricultural foundation models builds on two intersecting developments. First, task-specific deep-learning systems demonstrated that learned representations could support classification, detection, segmentation, prediction, and other analytical functions across agricultural imagery and sensor data. Second, smart farming developed as a data-intensive ecosystem involving farm machinery, sensors, platforms, service providers, public institutions, supply-chain actors, and farmers. Foundation models therefore enter agriculture not as isolated algorithms, but as components of heterogeneous technical and institutional systems in which data access, interoperability, ownership, control, and accountability influence how model outputs are produced and used [2, 3].

A persistent constraint across agricultural artificial intelligence is the limited availability of high-quality labels for diverse crops, animal conditions, soils, environments, seasons, and management systems. Agricultural annotation often requires specialist knowledge, repeated field observation, or laboratory confirmation, while collected datasets may remain geographically narrow or operationally inconsistent. These constraints have increased interest in self-supervised learning, weakly supervised learning, transfer learning, and other label-efficient approaches that use large volumes of unlabelled data before adaptation to narrower agricultural tasks. Such approaches may reduce part of the annotation burden, but they do not eliminate the need for reliable reference data, external evaluation, and careful analysis of domain shift [4].

A systematic review is needed because the evidence spans multiple model families and research traditions, including large language models, vision foundation models, vision-language systems, multimodal models, geospatial and remote-sensing foundation models, self-supervised pretraining, transfer learning, and emerging time-series architectures. The literature also differs substantially in agricultural context, data modality, evaluation design, and interpretation of readiness. This review therefore examines foundation models as a heterogeneous class of architecture-level approaches, maps their agricultural data and validation requirements, distinguishes demonstrated research capabilities from proposed opportunities, and integrates technical findings with evidence concerning governance, farmer-centred evaluation, privacy, cybersecurity, uncertainty, sustainability, and deployment risk.

Foundation Models in Agricultural AI

In this review, a foundation model is understood as a model pretrained at substantial scale on broad or heterogeneous data and intended to provide reusable representations or generative capabilities across multiple downstream tasks. Agricultural foundation models may originate from general-purpose language, vision, multimodal, or geospatial systems, or they may be pretrained specifically on crop, livestock, environmental, satellite, sensor, or agricultural knowledge data. This definition distinguishes foundation models from conventional machine-learning systems trained for one predefined outcome, single-task deep-learning networks developed on a bounded dataset, generic decision-support software without reusable pretrained representations, and traditional remote-sensing classifiers that do not employ large-scale transferable pretraining [5].

Large language models represent one foundation-model family relevant to agricultural AI. Their potential functions include information retrieval interfaces, summarisation, question answering, document preparation, knowledge navigation, and conversational assistance. In crop-production contexts, however, their usefulness varies by task and by the extent to which generated outputs depend on current, local, and specialised agronomic information. Linguistic coherence does not

demonstrate factual correctness, and general-purpose pretraining may omit local cultivars, production conditions, regulatory contexts, or recent scientific findings. Agricultural language-model outputs therefore require bounded use, source verification, domain grounding, and expert review where errors could affect consequential decisions [6].

Vision-language and multimodal foundation models extend pretrained representation learning across images, text, labels, sensor streams, or other data sources. Their architectures may combine a visual encoder, a language component, cross-modal projection or fusion layers, and task-specific adaptation mechanisms. In agriculture, these systems may connect crop images with disease descriptions, phenotypic traits, management records, environmental observations, or user questions. Their potential advantage is the ability to integrate visual and semantic evidence, but successful adaptation requires appropriately paired data, domain-relevant terminology, computational resources, and evaluation of whether generated textual interpretations are actually grounded in the supplied agricultural evidence [7].

Agricultural foundation-model development can consequently be viewed as a multistage process involving data collection and preprocessing, large-scale or inherited pretraining, domain adaptation, fine-tuning or prompting, application integration, and task-specific evaluation. Different stages may rely on satellite imagery, close-range images, farm sensors, weather observations, soil information, crop records, livestock data, or textual knowledge. Multimodal systems may create broader analytical possibilities, yet their complexity also increases exposure to data misalignment, missing modalities, inherited bias, high computing requirements, weak interpretability, and uncertain transfer across farms or regions. Foundation models should therefore be treated as adaptable research infrastructures rather than self-validating agricultural intelligence [8].

Review Scope and Research Questions

The review focuses on peer-reviewed evidence concerning foundation models and closely related architecture-level approaches in smart farming, precision agriculture, digital agriculture, agricultural remote sensing, crop science, livestock systems, phenotyping, and farm decision support. Included technologies comprise large language models, vision foundation models, vision-language models, multimodal foundation models, geospatial and remote-sensing foundation models, time-series models where directly supported, self-supervised pretraining, transfer learning, fine-tuning, prompting, retrieval-augmented generation where reported, and domain adaptation. For language-model evidence, the synthesis considers model architecture, agricultural knowledge integration, adaptation strategy, application type, validation design, and documented limitations rather than treating all generative systems as a uniform category [9].

Six review questions guide the analysis. RQ1 examines the model and architecture types studied in agricultural contexts. RQ2 addresses agricultural tasks, data modalities, and system settings. RQ3 evaluates technical architectures, pretraining strategies, adaptation methods, and data requirements. RQ4 synthesises opportunities for monitoring, prediction, knowledge discovery, decision support, automation, and sustainability-oriented applications. RQ5 examines bias, robustness, generalisation, explainability, uncertainty, privacy, cybersecurity, farmer trust, human oversight, governance, and field-validation limitations. RQ6 identifies research directions for responsible development and evaluation. The scope includes digital agri-food evidence when it informs farm-level data, knowledge, decision, or governance questions, but it does not assume that generative capability establishes operational maturity [10].

The conceptual boundary excludes conventional machine-learning studies lacking meaningful pretraining, transfer, self-supervised, multimodal, language-model, geospatial-foundation, or architecture-level relevance. It also excludes conference-only publications, preprints without journal publication, non-peer-reviewed reports, software documentation, websites, repositories, commercial materials, policy documents, and sources without verifiable DOI information. Remote-sensing evidence is included when it addresses large-scale pretraining, transferable geospatial representations, spatial or temporal adaptation, agricultural downstream tasks, or benchmark and validation concerns. This dedicated treatment is necessary because remote-sensing foundation models differ from ordinary image classifiers in their sensor diversity, geographic scale, spatial structure, temporal characteristics, and adaptation requirements [11]. **Table 1** summarises the systematic review scope, research questions, eligibility criteria, and evidence extraction domains.

Table 1. Systematic Review Scope for Foundation Models in Smart Farming and Precision Agriculture: Research Questions, Eligibility Criteria, Databases, Search Concepts, Screening Logic, Evidence Extraction Domains, and Review Boundaries

Review component	Operational definition	Included evidence	Excluded evidence	Purpose in the review	Boundary or limitation
Review domain	Foundation models and architecture-level AI approaches relevant to agricultural systems	Smart farming, precision agriculture, digital agriculture, agricultural remote sensing, crop science, livestock systems, phenotyping and farm decision support	Generic AI without substantive agricultural relevance	Establish the population and application domain	Agricultural relevance must be explicit rather than inferred from generic technical capability
RQ1: model types	Identification of foundation-model and related architecture families	Language, vision, vision-language, multimodal, geospatial, remote-sensing and supported time-series models	Single-purpose conventional models without transferable pretraining	Develop an agricultural foundation-model taxonomy	Terminology varies across studies and may not be used consistently
RQ2: tasks and contexts	Mapping of model use to agricultural tasks, modalities and operational settings	Monitoring, mapping, phenotyping, disease recognition, knowledge services, livestock observation and decision-support assistance	Applications lacking agricultural data or use context	Determine where evidence is concentrated	Presence of an application does not establish field effectiveness
RQ3: architectures and data	Examination of pretraining, adaptation, fusion and data requirements	Self-supervision, transfer learning, fine-tuning, prompting, domain adaptation, multimodal fusion and labelled-data requirements	Performance-only reports without sufficient technical description	Explain how reported systems are constructed and evaluated	Technical details are unevenly reported
RQ4: opportunities	Synthesis of capabilities proposed or demonstrated in reviewed research	Representation reuse, reduced annotation dependence, scalable monitoring, multimodal integration and knowledge access	Speculative commercial or operational claims without peer-reviewed evidence	Identify research and decision-support potential	Opportunities are interpreted as conditional rather than guaranteed outcomes

RQ5: risks and deployment challenges	Analysis of technical and sociotechnical failure conditions	Bias, fairness, robustness, generalisation, explainability, uncertainty, privacy, cybersecurity, farmer trust and governance	Unsupported assumptions that technical performance resolves deployment risk	Define responsible-use constraints	Direct agricultural evidence is limited for several governance dimensions
RQ6: research directions	Identification of evidence gaps requiring further investigation	Benchmarks, domain data, external validation, user evaluation, sustainable computing and lifecycle governance	Universal model, vendor or deployment prescriptions	Develop a research agenda grounded in identified gaps	Directions are proposed priorities, not operational recommendations
Eligible publication type	Peer-reviewed journal article with verifiable DOI	Reviews, technical studies, application studies, methods articles and relevant social-science research	Books, chapters, theses, conferences, preprints, websites, reports and policy documents	Maintain evidence quality and traceability	Journal publication does not by itself establish methodological quality
Eligibility period	Publication within the predefined review interval	Articles published from 1 January 2017 to the search date	Articles outside the interval	Capture contemporary developments underlying agricultural foundation models	Earlier foundational concepts may be underrepresented
Search sources	Academic databases and publisher-platform interfaces	ScienceDirect, SpringerLink, IEEE Xplore, PubMed, MDPI, Wiley and DOI-linked publisher records	Google Scholar as a sole evidence source	Increase disciplinary and publisher coverage	Direct subscription-database export coverage was not identical across platforms
Search concepts	Combined technology, agricultural-domain and evidence-theme terms	Foundation models, LLMs, multimodal learning, remote sensing, self-supervision, agriculture, validation and governance	Narrow searches based only on individual model names	Improve retrieval across inconsistent terminology	Rapidly changing terminology may reduce retrieval completeness
Evidence extraction	Structured extraction of technical, agricultural,	Study aim, context, architecture, modality, task, data need, validation, opportunity, risk and	Unreported information inferred by the reviewers	Support transparent thematic synthesis	"Not clearly reported" is required where evidence is absent

	evaluation and risk information	deployment challenge			
Review boundary	Separation of research capability from readiness	Benchmark findings, architecture descriptions and documented limitations	Claims of autonomous, safe or deployment-ready operation without direct evidence	Prevent overinterpretation	The review does not certify models, recommend vendors or prescribe farm actions

Search Strategy and Study Selection

The review process was structured around transparent identification, screening, eligibility assessment, extraction, and synthesis. Reporting design was informed by systematic-review guidance concerning explicit documentation of study identification and selection and by search-reporting guidance concerning databases, interfaces, search strings, dates, limits, and supplementary retrieval methods [12, 13]. These frameworks were used to organise the review process and reporting record; they were not interpreted as evidence of prospective registration, meta-analysis, formal risk-of-bias scoring, or complete compliance with every optional reporting element.

Searches covered ScienceDirect and Elsevier journal collections, SpringerLink, IEEE Xplore, PubMed, the MDPI journal platform, Wiley Online Library, DOI-linked publisher records, and backward and forward reference checking from eligible articles. Search concepts combined terms for foundation models, large language models, vision and vision-language models, multimodal systems, geospatial and remote-sensing foundation models, time-series models, self-supervised learning, transfer learning, fine-tuning, prompting, retrieval-augmented generation, and domain adaptation with terms for agriculture, smart farming, precision agriculture, crops, plants, livestock, phenotyping, soil, irrigation, weeds, pests, disease, robotics, and decision support. Multiple retrieval systems were used because academic search platforms differ in coverage, query functionality, indexing practices, and reproducibility characteristics [14].

The search covered publications from 1 January 2017 to 13 July 2026. The documented candidate log contained 51 records identified through database and publisher-platform searching and five additional records identified through reference checking. Six duplicate or repeated-version records were removed, leaving 50 records for title and abstract screening. Eight records were excluded at that stage because they lacked agricultural relevance, addressed generic artificial intelligence only, or did not demonstrate sufficient foundation-model, pretraining, transfer-learning, self-supervised, multimodal, or architecture-level relevance. Forty-two articles underwent full-text or detailed publisher-metadata assessment. Eight were subsequently excluded because they were conference or preprint-only publications, lacked an eligible journal or verifiable DOI record, focused on conventional task models without adequate architecture-level relevance, had only indirect agricultural relevance, or could not support a planned systematic-review claim. The resulting evidence set contained 34 journal articles.

For each included article, evidence was extracted concerning study aim, agricultural context, model or architecture type, data modality, agricultural task, technical architecture feature, data requirement, validation or evaluation approach, reported opportunity, reported risk or limitation, deployment challenge, and contribution to the review synthesis. Evidence was recorded as “not clearly reported” when the source did not provide sufficient detail. Claim-specific citation mapping was completed before manuscript drafting, and references were fixed in order of first appearance. The search has limitations: direct institution-level Scopus and Web of Science exports were not available in the evidence-building interface, terminology evolved rapidly during the eligibility period, and some emerging model families may be indexed under task-specific rather than foundation-model terminology. Consequently, the review emphasises reproducible scope, cautious interpretation, and transparent evidence boundaries rather than claiming exhaustive retrieval of every

agricultural foundation-model publication. **Table 2** summarises the study selection and evidence extraction structure used for the systematic review.

Table 2. Study Selection and Evidence Extraction Structure for the Systematic Review: Database Search, Duplicate Removal, Title and Abstract Screening, Full-Text Eligibility, Inclusion Logic, Exclusion Reasons, Evidence Extraction Fields, and Transparency Notes

Review stage	Screening or extraction action	Eligibility or extraction criterion	Output	Transparency note	Potential limitation
Database and publisher-platform search	Execute combined technology, agriculture and evidence-theme searches	Potential relevance to foundation models or related architecture-level agricultural AI	51 candidate records	Search concepts and interfaces were documented	Platform-specific indexing and query syntax may differ
Supplementary reference checking	Examine references and citations of eligible reviews and technical studies	Journal article with DOI and direct relevance to an uncovered review domain	5 additional candidate records	Added before deduplication	Citation checking may favour well-connected publications
Duplicate removal	Compare titles, authors, DOI records and publication versions	Retain one eligible journal version of each article	50 unique records after six removals	Journal versions were preferred over preprints or repeated index records	Metadata differences can complicate duplicate identification
Title and abstract screening	Assess agricultural and architecture-level relevance	Explicit agricultural context and meaningful foundation-model, transfer, self-supervised, multimodal or pretraining relevance	42 articles advanced after eight exclusions	Predefined exclusion categories were used	Abstracts may omit relevant technical or validation details
Full-text eligibility assessment	Evaluate publication type, DOI, technical content, agricultural fit and claim relevance	Peer-reviewed journal article able to support at least one planned review claim	34 included articles after eight exclusions	Exclusion reasons were recorded by category	Full-text access and reporting completeness varied
Inclusion logic	Map each article to review questions and synthesis domains	Direct or carefully bounded support for architecture, data, opportunity, risk, validation or research-gap claims	Final evidence set	No article was retained solely to reach a numerical target	The evidence set remains heterogeneous in design and maturity

Exclusion reasons	Remove ineligible publication types or weakly aligned evidence	Conference or preprint only; no eligible journal or DOI; conventional task model; indirect agriculture; insufficient claim fit	Documented exclusion categories	Categories correspond to predefined criteria	Some borderline studies could reasonably be classified differently
Evidence extraction: study characteristics	Record aim, context, architecture, modality and task	Information explicitly reported by the article	Structured study profile	Missing information was not inferred	Reporting standards differ across technical and review articles
Evidence extraction: technical features	Record pretraining, adaptation, fusion and data requirements	Sufficient description of architecture or learning strategy	Architecture and data map	Technical terminology was normalised cautiously	Articles may use inconsistent definitions of foundation models
Evidence extraction: evaluation	Record benchmark, transfer, validation and generalisation information	Evaluation procedure explicitly described	Validation evidence profile	Internal results were separated from external validation	Many studies used narrow or application-specific datasets
Evidence extraction: opportunities and limitations	Record reported capabilities, risks and implementation constraints	Claim supported directly by study results or discussion	Opportunity–risk synthesis	Proposed opportunities were distinguished from demonstrated outcomes	Reviews and conceptual articles may provide indirect rather than empirical evidence
Evidence extraction: governance and users	Record privacy, security, accountability, trust and oversight information	Directly reported or systematically reviewed evidence	Sociotechnical evidence profile	Governance evidence was integrated without assuming legal sufficiency	Direct farmer-centred evidence remains limited
Citation planning	Assign every included article a specific first-use claim	Citation must match the supported sentence and sequential order	Locked Vancouver citation blueprint	Grouped citations were limited to the preapproved audit	Later manuscript changes could disrupt citation order
Synthesis preparation	Organise findings across architectures, data, applications, risks and research directions	Evidence patterns supported by multiple or directly relevant articles	Narrative and tabular synthesis	No meta-analysis or universal model ranking was attempted	Heterogeneity prevents quantitative pooling

Technical Architectures and Data Requirements

Vision foundation models represent one of the most developed architecture classes in the reviewed agricultural literature. Rather than learning all visual features from a task-specific agricultural dataset, these models begin with broadly pretrained representations and are subsequently adapted to agricultural classification, detection, or segmentation tasks. Agricultural adaptation may modify feature aggregation, intermediate layers, task heads, or fine-tuning procedures to retain transferable visual information while responding to crop-specific structures. Reported application results demonstrate the feasibility of this approach within selected datasets, but they do not establish that visual representations will transfer uniformly across crops, cultivars, imaging systems, growth stages, weather conditions, or field environments [15].

Geospatial and remote-sensing foundation models require additional architectural treatment because agricultural observations are structured across space, time, spectral bands, sensors, and geographic regions. Cropland classification systems may separate spatial and temporal attention so that land-pattern characteristics and seasonal crop development are learned through distinct but coordinated components. This approach is relevant to agricultural landscapes in which crops can have similar appearances at one observation date but diverge across their phenological trajectories. Time-series foundation-model research remains less mature than vision-oriented agricultural work, and temporal architectures require careful evaluation of seasonality, missing observations, changing management practices, sensor differences, and future-period transfer [16].

Self-supervised pretraining provides a complementary route for learning from large agricultural datasets that lack complete labels. Satellite time-series models, for example, can be pretrained on unlabelled observations and then adapted to crop-mapping tasks using smaller labelled datasets. Such designs may learn spectral and temporal relationships before task-specific supervision, potentially reducing dependence on expensive crop labels. Their validity nevertheless depends on the representativeness of the unlabelled pretraining data, the quality of downstream reference labels, and whether evaluation is conducted across genuinely independent regions, seasons, and sensor conditions rather than through random partitions of closely related observations [17].

Vision-language architectures extend these principles by aligning visual evidence with textual descriptions, class labels, prompts, or agricultural knowledge. In small-sample crop-disease recognition, image–text–label integration may provide semantic information that is unavailable to image-only systems and may assist learning when labelled examples are limited. However, the quality of multimodal learning depends on whether visual observations and textual descriptions are accurately paired, whether disease labels reflect reliable diagnoses, and whether generated interpretations remain grounded in the input evidence. Performance within a single crop or disease dataset cannot be assumed to transfer to other crops, disorders, cameras, field backgrounds, or production environments [18].

Prompted vision systems and general-purpose segmentation models provide another adaptation pathway. Crop-field boundary research indicates that a general vision foundation model can be incorporated into an agricultural remote-sensing pipeline through super-resolution preprocessing and task-specific edge correction, although pipeline complexity and geographic transfer remain important limitations [19]. Transferability evidence further shows that agricultural classification performance can change under phenological, temporal, and spatial shifts, making external-domain testing essential [20]. Contrastive self-supervised learning may reduce reliance on exhaustive agricultural annotation by learning representations from unlabelled images, but outcomes remain dependent on data composition, augmentation design, and downstream evaluation [21]. Integrating language and vision models into plant-stress phenotyping may expand systems from trait recognition toward interpretive or decision-support functions, yet hallucination, explainability, uncertainty, and human-oversight requirements remain unresolved [22]. **Table 3** maps technical architectures and data requirements for foundation models in smart farming and precision agriculture.

Table 3. Technical Architectures and Data Requirements for Foundation Models in Smart Farming and Precision Agriculture: Large Language Models, Vision Foundation Models, Vision-Language Models, Multimodal Models, Geospatial Models, Remote Sensing Models, Time-Series Models, Transfer Learning, Fine-Tuning, Domain Adaptation, Multimodal Data Fusion, Benchmarking, Validation, and Deployment Boundaries

Architecture or data component	Core technical question	Required evidence or input	Model function	Agricultural use context	Validation need	Risk of overinterpretation
Large language model	Can a pretrained language model process agricultural terminology, context, and user questions reliably?	Curated agricultural corpora, current knowledge, local terminology, provenance, and expert-reviewed outputs	Text generation, summarisation, question answering, retrieval interface, and reasoning assistance	Advisory support, documentation, knowledge access, and research navigation	Factuality testing, source verification, local-context evaluation, and expert review	Fluency may be mistaken for agronomic correctness
Vision foundation model	Do broadly pretrained visual representations transfer to agricultural imagery?	Diverse crop, livestock, field, lighting, background, sensor, and growth-stage data	Classification, detection, segmentation, and feature extraction	Crop monitoring, disease recognition, weed detection, phenotyping, and robotic perception	Multicrop, multiseason, multisensor, and external-farm validation	Strong internal results may be interpreted as universal transfer
Vision-language model	Does image-text alignment improve agricultural interpretation?	Paired images, descriptions, labels, prompts, and agricultural knowledge	Visual question answering, semantic recognition, explanation generation, and open-vocabulary analysis	Plant disease, phenotyping, livestock observation, and advisory interfaces	Grounding tests, hallucination checks, expert validation, and explanation-fidelity assessment	Generated explanations may appear evidential when they are not
Multimodal foundation model	Can heterogeneous agricultural modalities be aligned without losing temporal or contextual meaning?	Synchronised imagery, sensors, weather, soil, management records, livestock data, and text	Joint representation, multimodal fusion, and cross-modal inference	Integrated crop, livestock, and farm decision-support research	Missing-modality testing, alignment checks, ablation analysis, and external validation	Additional modalities may increase complexity with improving reliability
Geospatial foundation model	Do large-scale spatial representations transfer to agricultural landscapes?	Multi-region, multi-resolution, multi-sensor, and georeferenced data	Mapping, segmentation, spatial inference, and land-pattern representation	Cropland classification, land-use analysis, and field-boundary mapping	Cross-region, cross-resolution, and cross-sensor evaluation	Broad geospatial performance may conceal crop-specific limitations

Remote-sensing foundation model	Can pretrained Earth-observation models represent agricultural spectral, spatial, and seasonal characteristics?	Multispectral, hyperspectral, synthetic-aperture-radar, aerial, and satellite data with metadata	Crop classification, detection, segmentation, and change analysis	Crop mapping, field monitoring, land management, and stress observation	Seasonal, geographic, atmospheric, sensor, and resolution validation	Remote-sensing benchmark performance may be treated as agronomic usefulness
Time-series foundation model	Can the model learn seasonal, longitudinal, and event-dependent agricultural dynamics?	Repeated satellite, sensor, weather, crop, soil, management, or livestock observations	Temporal representation, forecasting, sequence modelling, and anomaly detection	Crop development, yield-support research, irrigation monitoring, and animal behaviour	Future-season testing, missing-data evaluation, and temporal-shift analysis	Historical fit may be misinterpreted as future reliability or causality
Self-supervised learning	Can useful agricultural representations be learned without exhaustive labels?	Large unlabelled datasets and smaller labelled evaluation subsets	Pretraining, contrastive learning, masked prediction, and representation learning	Crop images, satellite time series, sensors, and livestock streams	Label-fraction comparison, independent transfer testing, and data-quality assessment	Reduced annotation need may be described incorrectly as elimination of labels
Transfer learning	How effectively do pretrained features move between tasks, farms, regions, and seasons?	Well-described source and target datasets with documented distribution differences	Reuse and adaptation of pretrained parameters or representations	Crop classification, disease detection, phenotyping, and remote-sensing analysis	Spatial, temporal, phenological, sensor, and management-transfer tests	Successful fine-tuning on one target can conceal poor wider transfer
Fine-tuning	How much agricultural task-specific adaptation is required?	Curated downstream data, labels, task definitions, and validation sets	Full, partial, or parameter-efficient model adaptation	Specialised crop, livestock, remote-sensing, and advisory tasks	Overfitting assessment, calibration, and independent-domain evaluation	Fine-tuned performance may remain narrow dataset-specific
Prompting	Can desired task behaviour be elicited without changing	Instructions, examples, task constraints, visual prompts, and agricultural context	In-context task specification, conversational interaction, and prompted segmentation	Language interfaces, information assistance, and vision segmentation	Prompt-sensitivity, reproducibility, and failure-case testing	Selected successful prompts may be treated as stable model capability

Temporal data	Can models account for seasonality, phenology, management timing, and long-term change?	Timestamps, repeated observations, growth stages, interventions, and multiple seasons	Temporal representation, forecasting, and sequence analysis	Crop growth, yield-related research, weather effects, and animal monitoring	Future-period and cross-season validation	Temporal leakage may inflate performance
Sensor data	Can heterogeneous agricultural devices be harmonised and trusted?	Calibration, sampling frequency, units, maintenance records, drift, and missingness	Environmental, equipment, soil, and physiological-state representation	Irrigation, soil monitoring, machinery, greenhouses, and livestock	Cross-device, degradation, missingness, and real-time evaluation	Laboratory-quality measurement may not represent operational sensor streams
Satellite data	Which spatial, temporal, and spectral characteristics are appropriate for the task?	Multispectral or radar imagery, atmospheric correction, metadata, and reference labels	Large-area observation and geospatial pretraining	Crop classification, land monitoring, field boundaries, and stress assessment	Cross-sensor, cloudy-condition, region, season, and resolution testing	Coarse resolution or processing artefacts may be overlooked
Drone imagery	Can high-resolution aerial data improve fine-scale agricultural analysis?	Flight plans, camera calibration, georeferencing, altitude, weather, and field diversity	Fine-scale detection, segmentation, and phenotyping	Weed mapping, plant counts, crop stress, and structural phenotyping	Cross-flight, field, altitude, season, and camera validation	Curated flight may not represent routine operational conditions
Weather data	Does weather context improve agricultural prediction or interpretation?	Local observations, forecasts, historical records, uncertainty, and spatial resolution	Contextual conditioning and temporal prediction	Disease-risk research, irrigation support, crop monitoring, and yield-related analysis	Forecast-error propagation and climate-regime evaluation	Correlation may be interpreted incorrectly as causal agricultural reasoning
Soil data	Can models represent highly heterogeneous soil characteristics at useful scales?	Physical, chemical, moisture, biological, and spatial soil measurements	Soil-state representation, mapping, and contextual conditioning	Soil monitoring, irrigation, nutrient research, and crop suitability analysis	Ground sampling, cross-field testing, and spatial-uncertainty assessment	Sparse sampling may create false spatial precision

Crop data	Are crop species, cultivar, development stage, treatment, and management adequately documented?	Crop identity, phenology, management operations, environmental context, and outcomes	Crop-specific representation, adaptation, and prediction	Monitoring, disease recognition, phenotyping, and yield support	Multicultivar, multistage, and management-system evaluation	Crop labels also may conceal important production-context differences
Textual agricultural knowledge	Is embedded or retrieved knowledge current, local, traceable, and authoritative?	Peer-reviewed evidence, expert-curated resources, extension material, terminology, and provenance	Language adaptation, retrieval, explanation, and knowledge navigation	Advisory support, question answering, and knowledge discovery	Source-quality, temporal-validity, citation-faithfulness, and local-relevance audits	General web text may contain outdated or contextually unsuitable information
Benchmarking	Do evaluations permit meaningful and reproducible comparison?	Shared datasets, predefined splits, suitable baselines, metrics, compute reporting, and uncertainty measures	Comparative assessment	All model families and agricultural tasks	Hidden external tests, subgroup reporting, and replication	Leaderboard ranking may reward narrow optimisation
External validation	Does model performance persist beyond development conditions?	Independent farms, regions, seasons, devices, sensors, and user groups	Generalisation assessment	All applications with potential operational relevance	Prespecified independent evaluation and prospective testing	Internal testing may be misrepresented field readiness
No deployment-readiness boundary	What technical, human, organisational, and governance evidence is still missing?	Robustness, safety, field, integration, monitoring, accountability, and user evidence	Interpretation constraint rather than a model function	All proposed agricultural uses	Staged field evaluation, fallback testing, monitoring, and responsibility assessment	Technical feasibility may be labelled prematurely as readiness

Opportunities, Risks, and Deployment Challenges

The reviewed evidence identifies potential opportunities not only for replacing narrow analytical components, but also for connecting information across tasks and modalities. Large language models may assist with documentation, communication, information synthesis, and task-level knowledge access, while their impact on agricultural work may include both complementary support and changes in occupational responsibilities. Such effects cannot be evaluated only

through model accuracy because they also concern skill requirements, role redesign, unequal access, and the distribution of responsibility between workers and automated systems [23].

Conversational and generative interfaces may make agricultural analytical tools more accessible to users who cannot interact directly with programming languages, complex dashboards, or specialist modelling systems. They may support natural-language questioning, explanation, translation, or navigation of analytical outputs. At the same time, farmer-facing interfaces can reproduce hallucinations, weak local knowledge, digital-literacy barriers, unreliable connectivity, language exclusions, and misplaced confidence in plausible answers. Their usefulness therefore depends on local context, understandable uncertainty, source transparency, escalation to qualified expertise, and evaluation with intended users rather than assumptions about user needs [24].

Farmer trust and technology acceptance are also deployment conditions rather than automatic consequences of technical capability. Agricultural AI adoption can be influenced by perceived usefulness, behavioural intention, compatibility with established work, confidence in the technology, social expectations, available support, and perceptions of control. These factors may differ among farm types, regions, production systems, and levels of digital experience. A model that performs adequately in a technical study may still be rejected, misunderstood, overtrusted, or used in unintended ways if its purpose, limitations, required actions, and accountability arrangements are unclear [25].

Remote-sensing and geospatial foundation models create opportunities for scalable agricultural observation, but responsible use requires attention to dataset provenance, geographic representation, bias, uncertainty, explainability, human rights, environmental implications, and the downstream consequences of spatial decisions. These concerns are particularly important when maps or classifications influence resource allocation, monitoring, credit, insurance, compliance, or access to services. Responsible Earth-observation AI therefore requires lifecycle evaluation that extends from data collection and model design to interpretation, communication, institutional use, and mechanisms for contesting or correcting outputs [26].

Privacy and cybersecurity become increasingly important as foundation-model pipelines combine farm-management data, machinery records, sensor streams, imagery, communications, livestock information, and external platforms. Federated learning may support collaborative model development without centralising every raw record, but it does not by itself prevent parameter leakage, poisoning, re-identification, non-independent data effects, or unequal control over the resulting models [27]. Connected smart-farming systems also remain vulnerable to attacks affecting devices, communications, identities, data integrity, service availability, and automated control functions [28]. These risks reinforce two claim boundaries: the reviewed evidence does not establish field readiness, and it does not support autonomous agricultural decisions without meaningful human review, override, monitoring, and accountable fallback procedures.

Evidence Synthesis

The integrated evidence shows that agricultural foundation-model research develops through an interdependent chain rather than a simple progression from larger models to better farming outcomes. Architecture types determine what representations can be learned; agricultural data quality and modality determine what information those representations contain; adaptation and validation determine whether capabilities transfer to a crop, animal, region, season, sensor, or user group; and governance conditions determine whether outputs can be used responsibly within real agricultural institutions. Opportunities in monitoring, mapping, phenotyping, knowledge access, and decision-support assistance therefore coexist with risks involving domain shift, data scarcity, bias, uncertainty, privacy, cybersecurity, accountability, and environmental cost. A systemic interpretation is necessary because narrow performance metrics can omit externalities occurring beyond the model-development setting [29]. **Figure 1** synthesises how foundation-model architectures, agricultural data requirements, application opportunities, risks, deployment challenges, and research directions are connected in smart farming and precision agriculture.

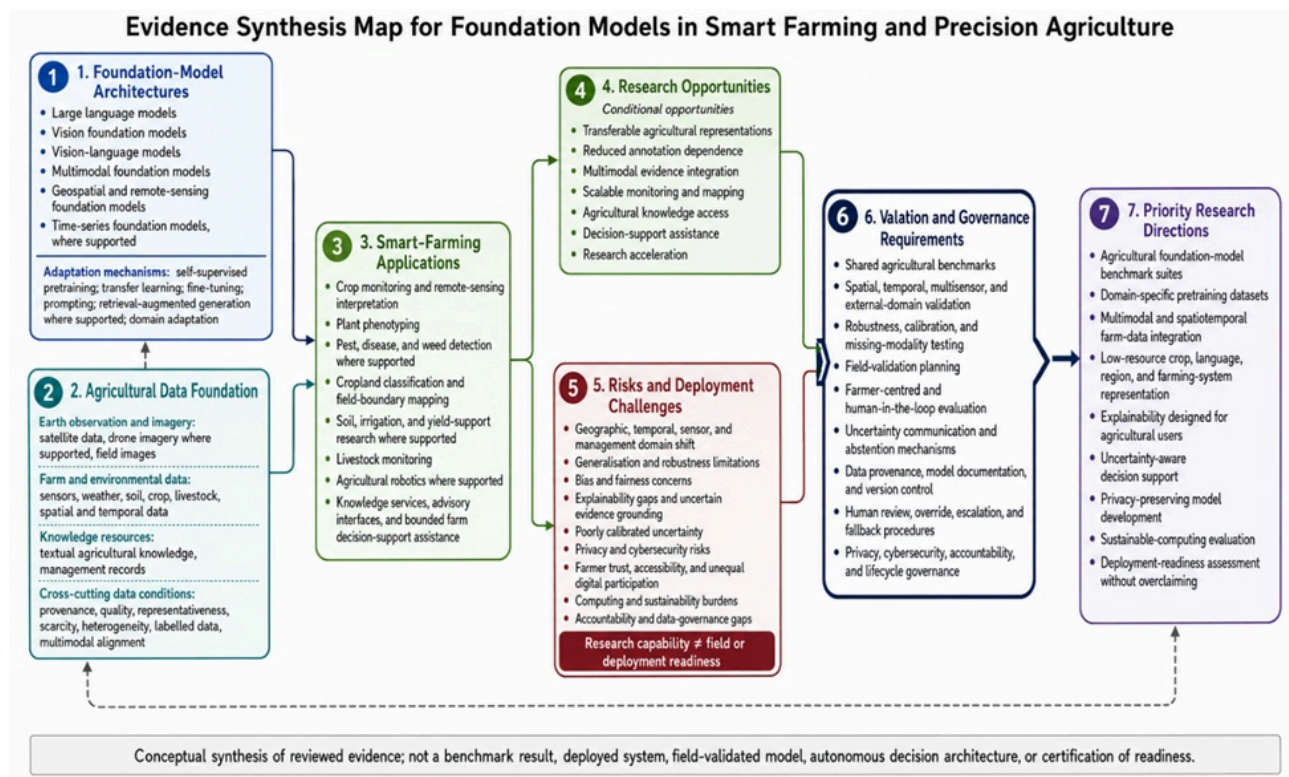


Figure 1. Evidence Synthesis Map for Foundation Models in Smart Farming and Precision Agriculture. The figure connects foundation-model architecture types, agricultural data requirements, application opportunities, risks, deployment challenges, validation needs, governance requirements, and future research directions identified in the systematic review.

The figure is conceptual and does not represent a benchmark result, deployed system, field-validated model, or autonomous agricultural decision-support architecture.

Alt-text description

A conceptual evidence synthesis map showing foundation-model types connected to agricultural data modalities, smart farming applications, technical opportunities, risks, deployment challenges, validation needs, governance requirements, and research directions.

Responsibility within this evidence chain is distributed. Model developers influence pretraining data, architectures, optimisation, documentation, and known limitations; data providers influence coverage, consent, quality, and provenance; agricultural organisations influence integration, training, monitoring, and escalation; and users influence interpretation and contextual judgment. Trustworthy practice therefore cannot be assigned to a single “responsible user” at the final stage. Responsibility gaps can emerge when developers assume deployers will validate models, deployers assume vendors have addressed bias and safety, and users receive outputs without sufficient information or authority to question them [30].

Cybersecurity evidence forms a related cluster that is sometimes separated from model evaluation but is directly relevant to foundation-model deployment. Agricultural AI systems may depend on sensors, edge devices, gateways, cloud services, APIs, farm-management platforms, mobile interfaces, and automated equipment. A technically accurate model can still become unreliable if the data stream is spoofed, corrupted, interrupted, or manipulated. The reviewed cybersecurity literature covers threat detection, prevention frameworks, technical countermeasures, organisational safeguards, and training, but evidence from long-term real-farm evaluation remains limited [31].

The social-science evidence introduces a further corrective to technology-centred interpretation. Digital agriculture changes relations among farmers, platforms, advisers, technology providers, supply-chain organisations, and public institutions. Access to infrastructure, data ownership, digital skill, labour effects, platform power, service dependency, and inclusion can influence who benefits from foundation-model systems and who bears their errors or costs. These dimensions are not secondary implementation issues; they shape which problems are defined, which data are collected, whose knowledge is represented, and how model outputs enter agricultural decisions [32].

Across the evidence base, the strongest clusters concern visual representation learning, remote-sensing analysis, crop classification, multimodal recognition, language-based knowledge access, and label-efficient pretraining. Evidence is thinner for independent field validation, causal decision impact, uncertainty calibration, farmer-centred longitudinal evaluation, low-resource regions, autonomous robotics, soil and irrigation applications, sustainable-computing assessment, and governance effectiveness. The literature therefore contains an important tension: foundation models are designed for broad transfer and reuse, but agricultural evidence remains predominantly narrow, context-dependent, and experimentally bounded. This contradiction does not invalidate the research direction; it defines the validation and governance agenda required before wider claims can be made.

Research Directions

Agricultural foundation-model benchmarks should move beyond single-dataset accuracy and include annotation efficiency, external-domain transfer, calibration, robustness, subgroup performance, computational demand, and reproducibility. Controlled evidence from crop classification indicates that self-supervised pretraining can retain useful performance with substantially smaller labelled subsets, making label-efficiency evaluation a meaningful benchmark dimension [33]. Future benchmark suites should nevertheless include independent crops, farms, seasons, regions, sensors, and management conditions so that reduced annotation dependence is not confused with general agricultural validity.

Precision livestock farming requires a parallel research programme because livestock systems generate different combinations of video, imagery, audio, physiological sensors, environmental measurements, movement records, feeding data, production records, and textual information. Foundation models may support animal identification, health observation, behaviour interpretation, feeding analysis, and decision-support research, but progress depends on synchronised longitudinal datasets, animal-level ground truth, robust multimodal fusion, welfare-sensitive validation, and safe integration with human and automated workflows [34]. Livestock benchmarks should therefore distinguish perception, interpretation, prediction, and intervention rather than treating them as one capability.

Domain-specific pretraining resources should be developed with documented provenance, geographic and production-system coverage, licensing, crop or animal representation, sensor metadata, and known limitations. Multimodal datasets should align satellite, drone, field-image, sensor, soil, weather, management, livestock, and textual data through consistent identifiers and timestamps. Research should give particular attention to low-resource crops, languages, farming systems, and regions that are poorly represented in general-purpose or high-income-country datasets. Geospatial and temporal adaptation methods should be evaluated under future seasons, climate variability, sensor changes, and management shifts, while bias and fairness assessments should report disaggregated performance rather than only aggregate averages.

Responsible evaluation also requires explanation methods designed for agricultural users, calibrated uncertainty, abstention when evidence is insufficient, robustness testing under field variability, and realistic human-in-the-loop studies. Farmer-centred validation should examine comprehension, workload, reliance, contestability, accessibility, and decision consequences rather than measuring acceptance alone. Privacy-preserving methods should be tested against leakage, poisoning, and non-independent farm data; sustainable-computing studies should report energy, hardware, and lifecycle effects; and governance standards should document datasets, model versions, limitations, intended uses, monitoring,

incidents, and accountable parties. **Table 4** presents research directions for responsible foundation-model development and evaluation in smart farming and precision agriculture.

Table 4. Research Directions for Responsible Foundation Models in Smart Farming and Precision Agriculture: Agricultural Benchmarks, Domain-Specific Pretraining, Multimodal Farm Data, Geospatial and Temporal Adaptation, Low-Resource Representation, Explainability, Uncertainty, Robustness, Bias, Human-in-the-Loop Evaluation, Privacy, Sustainable Computing, Governance, and Deployment-Readiness Assessment

Research direction	Problem addressed	Required methodological advance	Required data or infrastructure	Validation need	Governance or user need	Expected contribution
Agricultural foundation-model benchmarks	Fragmented tasks, metrics, datasets, and evaluation procedures	Multitask protocols incorporating transfer, calibration, robustness, annotation efficiency, uncertainty, and computing requirements	Curated crop, livestock, sensor, text, geospatial, and multimodal benchmark resources	Hidden external test sets and independent replication	Transparent benchmark governance and version control	More comparable and reproducible evidence
Domain-specific pretraining datasets	General-purpose corpora may not represent agricultural knowledge or environments adequately	Documented sampling, curation, quality control, and balanced domain coverage	Large, licensed, geographically diverse agricultural image, text, sensor, and remote-sensing corpora	Downstream transfer, contamination, bias, and subgroup evaluation	Data rights, provenance, consent, and benefit-sharing arrangements	Better representation of agricultural concepts and variations
Multimodal farm-data integration	Agricultural data are heterogeneous, asynchronous, incomplete, and institutionally fragmented	Fusion methods robust to missing, delayed, noisy, and conflicting modalities	Interoperable farm-data infrastructure, metadata, identifiers, time alignment, and quality indicators	Modality ablation, missing-data simulation, and external farm testing	Data-access, ownership, and sharing agreements	More context-aware agricultural representation
Geospatial and temporal model adaptation	Models may fail across regions, seasons, sensors, and climate conditions	Spatial–temporal adaptation, shift detection, continual evaluation, and uncertainty-aware transfer	Multiyear, multiregion, multi-resolution, and multi-sensor datasets	Future-season, unseen-region, and cross-sensor evaluation	Regional stakeholder participation and transparent limits	More defensible transfer across agricultural contexts
Low-resource crop and	Dominant crops, languages, and	Few-shot, self-supervised,	Data partnerships for neglected	Disaggregated evaluation	Local governance,	Broadened scientific

region representation	regions receive disproportionate research attention	active-learning, and transfer methods combined with representation audits	crops, smallholder systems, local languages, and underrepresented environments	under low-resource conditions	compensation, participation, and benefit sharing	geograph relevance
Explainability for agricultural users	Technical explanations may be unfaithful, incomprehensible, or irrelevant	Task-specific faithful explanation methods and explanation-quality evaluation	User-interface prototypes, expert rationales, evidence links, and counterfactual cases	Fidelity testing plus farmer and agronomist comprehension studies	Participatory design, correction mechanisms, and contestability	Better cap to review challenge, contextual outputs
Uncertainty-aware decision support	Foundation models can be confidently wrong or fail under distribution shift	Calibration, abstention, out-of-distribution detection, prediction intervals, and evidence-linked confidence	Representative validation streams and continuous monitoring infrastructure	Prospective calibration, error-severity, and decision-impact evaluation	Clear escalation, fallback, and override procedures	Safer bound decision support assistance
Robustness under field variability	Environmental noise, missing data, equipment differences, and biological variation reduce reliability	Agricultural stress tests, corruption suites, adversarial evaluation, and shift monitoring	Diverse real-field datasets, controlled test environments, and simulation resources	Cross-farm, cross-device, extreme-weather, and degraded-input testing	Defined safe failure states and incident response	Better characteris of failure condition
Bias and fairness assessment	Errors and benefits may be distributed unevenly across farms, crops, regions, and user groups	Agricultural subgroup definitions, distributional metrics, causal audits, and impact assessment	Geographic, crop, farm-structure, language, and access metadata	Disaggregated evaluation and repeated monitoring	Stakeholder-defined fairness objectives and redress	Visibility unequal performance and bene
Human-in-the-loop evaluation	Human oversight is frequently assumed without empirical workflow evidence	Controlled and field-based studies of review, correction, escalation, and override	Decision-support prototypes, realistic tasks, audit trails, and qualified participants	Prospective evaluation of reliance, correction quality, workload, and outcomes	Clear authority, responsibility, and access to supporting evidence	Evidence a when hur oversight improves fails to imp decision
Farmer-centred validation	Technical benchmarks omit usability, local	Participatory, mixed-method,	Diverse farmer panels, local-language	Comprehension, accessibility, trust, reliance,	Compensation, inclusion, autonomy,	Better alignme between

	relevance, autonomy, and long-term effects	and longitudinal evaluation	interfaces, and representative farm settings	contestability, and decision-impact assessment	recourse, and transparent purpose	systems and farmer ne
Privacy-preserving model development	Farm-data collaboration can expose commercially, personally, or operationally sensitive information	Federated learning, secure aggregation, differential privacy, access controls, and privacy accounting	Distributed infrastructure, secure identity management, and governance agreements	Leakage, poisoning, re-identification, and non-independent-data testing	Enforceable data rights, withdrawal, auditability, and responsibility allocation	Safer collabora agricultu learnin
Sustainable-computing evaluation	Training and operating large models may create energy, water, hardware, and rebound costs	Lifecycle, marginal-impact, and task-normalised efficiency assessment	Compute, energy, carbon, water, equipment, and utilisation records	Reproducible comparisons across architectures, scales, and deployment modes	Reporting standards and net-benefit criteria	More cred environme interpreta
Governance and documentation standards	Data histories, model limitations, roles, and incidents are inconsistently documented	Agriculture-specific dataset, model, evaluation, and deployment documentation frameworks	Versioned registries, provenance records, audit logs, monitoring systems, and incident databases	Documentation-quality and governance-process audits	Accountability, contestability, stakeholder access, and clear ownership of decisions	More trace model lifec and responsibi
Deployment-readiness assessment without overclaiming	Prototype or benchmark performance is often conflated with operational readiness	Staged assessment covering technical validity, human factors, infrastructure, safety, governance, and monitoring	Independent field-test facilities, fallback systems, monitoring tools, and incident-response processes	Prospective, external, multiseason, and workflow-level evaluation	Named accountable parties, escalation routes, recourse, and post-deployment review	Cleare evidenc thresholds future rese and evaluat

Conclusion

This systematic review integrates evidence on foundation models in smart farming and precision agriculture across model architectures, agricultural data requirements, application opportunities, technical risks, deployment challenges, governance concerns, and research gaps. The literature indicates that pretrained language, vision, multimodal, geospatial, remote-sensing, and self-supervised systems may support agricultural representation learning, monitoring, mapping, phenotyping, knowledge access, and bounded decision-support research. However, the evidence remains uneven across crops, regions, sensors, production systems, user groups, and real operating conditions. Responsible progress requires representative domain-specific data, transparent pretraining and adaptation procedures, multimodal

alignment, external spatial and temporal validation, robustness testing, calibrated uncertainty, faithful explanations, farmer-centred evaluation, meaningful human oversight, privacy and cybersecurity safeguards, sustainable-computing assessment, and accountable lifecycle governance. Foundation-model scale, generative fluency, or benchmark performance should therefore be interpreted as evidence of research capability rather than proof of agronomic validity, autonomous intelligence, operational safety, farmer trust, or deployment readiness.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 26 Aug 2025

Revised: 23 Sep 2025

Accepted: 06 Nov 2025

Published online: 25 May 2026

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Li J, Xu M, Xiang L, Chen D, Zhuang W, Yin X, et al. Foundation models in smart agriculture: Basics, opportunities, and challenges. *Comput Electron Agric.* 2024;222:109032.
<https://doi.org/10.1016/j.compag.2024.109032>.
- Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: A survey. *Comput Electron Agric.* 2018;147:70-90.
<https://doi.org/10.1016/j.compag.2018.02.016>.
- Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big data in smart farming—A review. *Agric Syst.* 2017;153:69-80.
<https://doi.org/10.1016/j.agsy.2017.01.023>.
- Li J, Chen D, Qi X, Li Z, Huang Y, Morris D, et al. Label-efficient learning in agriculture: A comprehensive review. *Comput Electron Agric.* 2023;215:108412.

<https://doi.org/10.1016/j.compag.2023.108412>.

Yin S, Xi Y, Zhang X, Sun C, Mao Q. Foundation models in agriculture: A comprehensive review. *Agriculture (Basel)*. 2025;15(8):847.

<https://doi.org/10.3390/agriculture15080847>.

Kuska MT, Wahabzada M, Paulus S. AI for crop production—where can large language models provide substantial value? *Comput Electron Agric*. 2024;221:108924.

<https://doi.org/10.1016/j.compag.2024.108924>.

Haghighat M, Saleh A, Rahimi Azghadi M. Multimodal language models in agriculture: A tutorial and survey. *Inf Fusion*. 2026;129:104042.

<https://doi.org/10.1016/j.inffus.2025.104042>.

Liu J, Wang X. Multimodal large models for smart agriculture: Frameworks and applications. *Expert Syst Appl*. 2026;300:130394.

<https://doi.org/10.1016/j.eswa.2025.130394>.

Li H, Wu H, Li Q, Zhao C. A review on enhancing agricultural intelligence with large language models. *Artif Intell Agric*. 2025;15(4):671-85.

<https://doi.org/10.1016/j.aiia.2025.05.006>.

Shahriar S, Corradini MG, Sharif S, Moussa M, Dara R. The role of generative artificial intelligence in digital agri-food. *J Agric Food Res*. 2025;20:101787.

<https://doi.org/10.1016/j.jafr.2025.101787>.

Huo C, Chen K, Zhang S, Wang Z, Yan H, Shen J, et al. When remote sensing meets foundation model: A survey and beyond. *Remote Sens (Basel)*. 2025;17(2):179.

<https://doi.org/10.3390/rs17020179>.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;372.

<https://doi.org/10.1136/bmj.n71>.

Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *J Med Libr Assoc*. 2021;109(2):174-200.

<https://doi.org/10.5195/jmla.2021.962>.

Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Res Synth Methods*. 2020;11(2):181-217.

<https://doi.org/10.1002/jrsm.1378>.

Ye J, Yu Z, Lin J, Li H, Lin L. Vision foundation model for agricultural applications with efficient layer aggregation network. *Expert Syst Appl*. 2024;257:124972.

<https://doi.org/10.1016/j.eswa.2024.124972>.

Yan J, Gu X, Chen Y. CropSTS: A remote sensing foundation model for cropland classification with decoupled spatiotemporal attention. *Remote Sens (Basel)*. 2025;17(14):2481.

<https://doi.org/10.3390/rs17142481>.

Xu Y, Ma Y, Zhang Z. Self-supervised pre-training for large-scale crop mapping using Sentinel-2 time series. *ISPRS J Photogramm Remote Sens*. 2024;207:312-25.

<https://doi.org/10.1016/j.isprsjprs.2023.12.005>.

Cao Y, Chen L, Yuan Y, Sun G. Cucumber disease recognition with small samples using image-text-label-based multi-modal language model. *Comput Electron Agric*. 2023;211:107993.

<https://doi.org/10.1016/j.compag.2023.107993>.

Sun H, Wei Z, Yu W, Yang G, She J, Zheng H, et al. SIDEST: A sample-free framework for crop field boundary delineation by integrating super-resolution image reconstruction and dual edge-corrected Segment Anything model. *Comput Electron Agric.* 2025;230:109897.

<https://doi.org/10.1016/j.compag.2025.109897>.

Hoppe H, Dietrich P, Marzahn P, Weiß T, Nitzsche C, von Lukas UF, et al. Transferability of machine learning models for crop classification in remote sensing imagery using a new test methodology: A study on phenological, temporal, and spatial influences. *Remote Sens (Basel).* 2024;16(9):1493.

<https://doi.org/10.3390/rs16091493>.

Güldenring R, Nalpantidis L. Self-supervised contrastive learning on agricultural images. *Comput Electron Agric.* 2021;191:106510.

<https://doi.org/10.1016/j.compag.2021.106510>.

Musazade E, Mrisho II, Gao J, Feng X. Integration of LLMs and VLMs in plant stress phenotyping: From trait recognition to decision support. *Plant Phenomics.* 2026;8(1):100161.

<https://doi.org/10.1016/j.plaphe.2025.100161>.

Marinoudi V, Benos L, Camacho Villa C, Kateris D, Berruto R, Pearson S, et al. Large language models impact on agricultural workforce dynamics: Opportunity or risk?. *Smart Agric Technol.* 2024;9:100677.

<https://doi.org/10.1016/j.atech.2024.100677>.

Zhang W, Yadav A. Generative AI as an interface between farmers and AI algorithms: Opportunities and challenges. *J Indian Inst Sci.* 2025;105:305-21.

<https://doi.org/10.1007/s41745-025-00490-8>.

Mohr S, Kühl R. Acceptance of artificial intelligence in German agriculture: An application of the technology acceptance model and theory of planned behavior. *Precis Agric.* 2021;22(6):1816-44.

<https://doi.org/10.1007/s11119-021-09814-x>.

Ghamisi P, Yu W, Marinoni A, Gevaert CM, Persello C, Selvakumaran S, et al. Responsible artificial intelligence for Earth observation: Achievable and realistic paths to serve the collective good. *IEEE Geosci Remote Sens Mag.* 2025;13(3):72-96.

<https://doi.org/10.1109/MGRS.2025.3529726>.

Dembani R, Karvelas I, Akbar NA, Rizou S, Tegolo D, Fountas S. Agricultural data privacy and federated learning: A review of challenges and opportunities. *Comput Electron Agric.* 2025;232:110048.

<https://doi.org/10.1016/j.compag.2025.110048>.

Yazdinejad A, Zolfaghari B, Azmoodeh A, Dehghantanha A, Karimipour H, Fraser E, et al. A review on security of smart farming and precision agriculture: Security aspects, attacks, threats and countermeasures. *Appl Sci (Basel).* 2021;11(16):7518.

<https://doi.org/10.3390/app11167518>.

Tzachor A, Devare M, King B, Avin S, Ó hÉigeartaigh S. Responsible artificial intelligence in agriculture requires systemic understanding of risks and externalities. *Nat Mach Intell.* 2022;4(2):104-9.

<https://doi.org/10.1038/s42256-022-00440-4>.

Alexander CS, Yarborough M, Smith A. Who is responsible for 'responsible AI'? Navigating challenges to build trust in AI agriculture and food system technology. *Precis Agric.* 2024;25(1):146-85.

<https://doi.org/10.1007/s11119-023-10063-3>.

Campoverde-Molina M, Luján-Mora S. Cybersecurity in smart agriculture: A systematic literature review. *Comput Secur.* 2025;150:104284.

<https://doi.org/10.1016/j.cose.2024.104284>.

Klerkx L, Jakku E, Labarthe P. A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. *NJAS Wageningen J Life Sci.* 2019;90-91:100315.

<https://doi.org/10.1016/j.njas.2019.100315>.

Safonova A, Stiller S, Yordanov M, Ryo M. Self-supervised learning outperforms supervised learning for crop classification by annotating only 5% of images. *Precis Agric.* 2026;27:4.

<https://doi.org/10.1007/s11119-025-10302-9>.

Tian M, Guo H, Long C, Zhu J, Zhao W, Li H, et al. How large-scale foundation models benefit precision livestock farming: A survey. *Artif Intell Agric.* 2026;16(2):1109-32.

<https://doi.org/10.1016/j.aiia.2026.04.013>.