Journal of Engineering Science and Technology
Special Issue on INCITEST2025 Vol. 21, No. 2 (2026) 129 - 136

GLOBAL TRENDS, RESEARCH FRONTIERS, AND A PROPOSED AI-BASED PALM OIL SUPPLY CHAIN MANAGEMENT SYSTEM

ZAINAL ARIFIN HASIBUAN[1,*], IRFAN MALIKI[1],
SRI SUPATMI[1], PRAVEENA NAIR SIVASANKARAN[2],
SANTI PHITHAKKITNUKOON[3], OYBEK OLIMOVICH DAMINOV[4]

[1]Universitas Komputer Indonesia, Bandung, Indonesia
[2]Taylor's University, Selangor, Malaysia
[3]Chiang Mai University, Chiang Mai, Thailand
[4]Tashkent State Technical University, Uzbekistan
*Corresponding Author: zhasibua@email.unikom.ac.id

Abstract

This paper maps global artificial intelligence (AI) research trends in the palm oil supply chain for 2020-2025. It also proposes a technical AI-based Palm Oil supply chain management system (AI-Based PO-SCMS). A Scopus query returned 1,345 records, and title-abstract screening produced a curated corpus of 256 peer-reviewed articles. Bibliometric analysis reveals an average annual publication growth of 21.5 percent. Machine learning and remote sensing appear as dominant themes. However, the use of the internet of things (IoT) to collect big datasets and blockchain to maintain traceability of palm fruit from plantation to refinery is almost absent. The proposed AI-Based PO-SCMS integrates vision-based bunch detection, a time-adaptive ensemble for yield prediction, unsupervised anomaly detection, and an explainability module, and specifies an edge-to-cloud deployment model with a staged validation roadmap. Conceptual validation using Sentinel-2 backtesting, augmented unmanned aerial vehicle (UAV)-like images, and synthetic production time series produced a representative baseline: yield mean absolute error near ten percent, object detection mean average precision about 0.72, and anomaly detection area under the curve (AUC) above 0.9. These findings establish a baseline for future multi-site field trials. This work supports trustworthy AI deployment for productivity and sustainability.

Keywords: Artificial intelligence, Machine learning, Palm oil supply chains, Remote sensing, System development.

1. Introduction

Palm oil production supplies a large share of global vegetable oil and supports tropical livelihoods [1]. The sector faces persistent problems in traceability, variable productivity, and environmental sustainability [2]. These problems complicate certification and market access. Rising regulation and consumer demand increase incentives to adopt digital monitoring and verification tools [3].

Recent studies apply machine learning (ML) and remote sensing to mapping, tree detection, ripeness grading, and yield forecasting in oil-palm systems [4]. Several experimental works report high task-level accuracy for tree detection and ripeness classification on curated image sets [5]. Multiple publications evaluate machine-learning regressors and neural networks for block-level yield prediction and report strong explanatory performance under controlled conditions [6].

These technical advances remain fragmented because most studies address isolated tasks, rely on limited or domain-specific datasets, and omit multi-site field validation [7]. Explainability and interpretable methods receive growing attention in agricultural analytics but remain underused in palm oil applications, which constrains stakeholder trust and operational adoption [8]. Sustainability and traceability research highlights systemic challenges that require integrated monitoring and decision support rather than isolated models [9]. Overall, the literature demonstrates advanced algorithms for discrete tasks but little progress toward a validated, end-to-end, explainable AI-based system and curated cross-site datasets.

This paper fills that gap by combining bibliometric mapping with a technical design for an AI-Based system tailored to the palm oil supply chain. A Scopus query returned 1,345 records, and screening produced a curated corpus of 256 articles. A bibliometric analysis quantified publication growth, keyword evolution, collaboration, and citation impact. The results show an average annual publication growth of 21.5 percent, mean citations of 12.22 per document, mean authorship of 5.78 per article, and international collaboration at 35.9 percent.

ML and remote sensing emerged as dominant themes, while explainability, field-scale validation, curated labelled datasets, IoT, and blockchain usage remained scarce. The proposed system specifies vision-based bunch detection, a time-adaptive ensemble for yield prediction, unsupervised anomaly detection, and an explainability module within an edge-to-cloud deployment and staged validation roadmap.

Conceptual validation using Sentinel-2 backtesting, augmented UAV-like images, and synthetic production time series produced baselines: yield mean absolute error (MAE) near 10 %, object detection mean average precision (mAP) near 0.72, and anomaly area under the curve (AUC) above 0.9. Blockchain and IoT aspects are deferred to companion outputs.

This paper makes three concise contributions: (1) a bibliometric mapping (2020-2025) that identifies gaps and research frontiers; (2) a technical, system-level AI-Based Palm Oil Supply Chain Management System (AI-Based PO-SCMS) architecture specifying modular components, edge-to-cloud deployment, and model-management; and (3) a staged validation roadmap with conceptual baselines reported above to accelerate lab proofs-of-concept and multi-site trials.

2. Method

The study performed a systematic Scopus search to identify literature on AI in the palm oil supply chain from 2020 to 2025. Four Boolean queries targeted the title, abstract, and keyword fields and returned 1,345 raw records. The corpus underwent deduplication, year filtering, and title and abstract screening for topical relevance. Only English peer-reviewed articles with explicit AI or ML applications to production, processing, monitoring, or decision support were retained. The final curated corpus comprised 256 articles. Full query strings and the screening flowchart appear in the supplementary repository.

Metadata extraction used standard Scopus fields (title, authors, affiliations, keywords, abstract, source, year, citations). Preprocessing combined scripted routines and manual curation. The team normalized author names and affiliations, harmonized keyword variants, consolidated synonymous labels, and removed residual duplicates and noninformative tokens. All scripts and transformation logs are archived in the project repository to ensure reproducibility.

Bibliometric computations and visualizations used the Bibliometrix R package through the Biblioshiny interface. Descriptive statistics summarized annual counts, citation metrics, authorship, and international collaboration. Thematic structure derived from co-word analysis and thematic mapping on the harmonized keyword set. Network analyses produced keyword co-occurrence and co-authorship graphs.

The proposed AI-Based PO-SCMS synthesizes bibliometric gaps, a targeted literature review, and engineering heuristics. Bibliometric results prioritized candidate modules, and methodological literature guided algorithm selection. The study draws on sampling-based and hybrid path-planning advances, including informed Probabilistic Roadmap (PRM) and hybrid Rapidly-exploring Random Tree (RRT)/optimizer methods, which show fast asymptotic optimality and robust local replanning for mobile platforms in dynamic environments [10-13].

Conceptual validation used retrospective and synthetic experiments. Backtesting paired Sentinel-2 time series with synthetic block-level yields. Vision modules fine-tuned pretrained detectors on augmented agricultural images to simulate UAV and farm-gate views. Anomaly detection trained on synthetic multivariate time series with injected theft and disruption events.

Experiments used ≈60/20/20 train/val/test splits with block-based cross-validation. Evaluation metrics included MAE and RMSE for regression, mAP@IoU0.5 for detection, and ROC AUC, precision, recall, and false-positive rate for anomaly detection. Robustness tests simulated domain shift by varying vegetation index distributions and image quality, and ablation studies assessed the effects of transfer learning and augmentation.

Ensemble models combine tree-based regressors and sequence models to improve robustness under variable weather [6, 7]. Transfer learning and augmentation reduce annotation requirements for vision tasks [4, 5]. Unsupervised encoders and isolation forests detect rare operational events when labelled anomalies are scarce [7]. Explainability modules supply local attributions and saliency maps to address stakeholder trust issues identified in the literature [8].

3. Results and Discussion

3.1. Corpus overview and bibliometric summary

The final curated corpus comprises 256 peer-reviewed articles published between 2020 and 2025, sourced from 191 sources and retrieved and screened from Scopus. The corpus shows an average annual publication growth of 21.5 percent, mean citations of 12.22 per document, mean authorship of 5.78 per article, and international collaboration at 35.9 percent. These aggregate indicators provide a quantitative baseline for evaluating topical concentration and research dynamics in proposed AI-Based PO-SCMS.

3.2. Thematic structure and topic trends

The co-word network results are shown in Fig. 1. The temporal trend analysis shows rapid growth in ML topics from 2020 to 2025, paralleled by an increase in the use of remote sensing and image-based methods. Thematic mapping places ML and palm oil in the motor-themes quadrant (see Fig. 2). These patterns indicate lower maturity and weaker centrality for those technologies and point to a field shifting toward ML-driven prediction and spatial sensing, with digital-ledger technologies remaining peripheral [14].

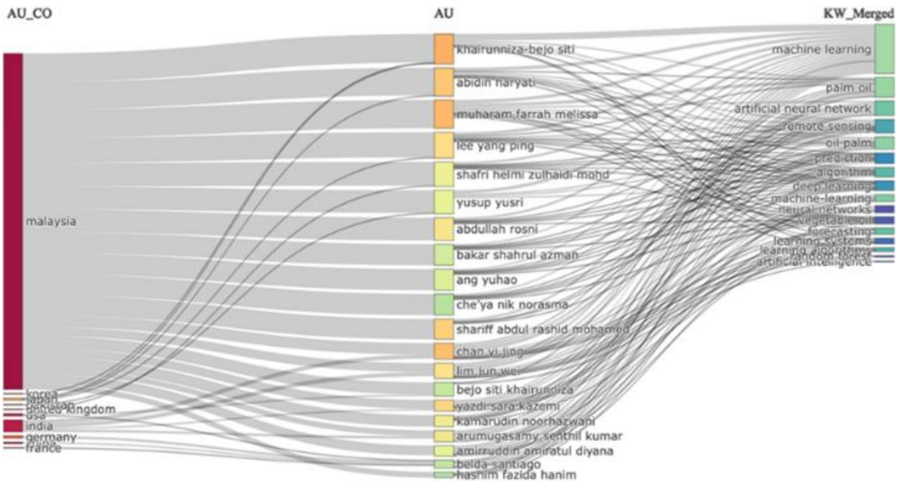

Fig. 1. Three-field plot linking countries, lead authors, and keywords.

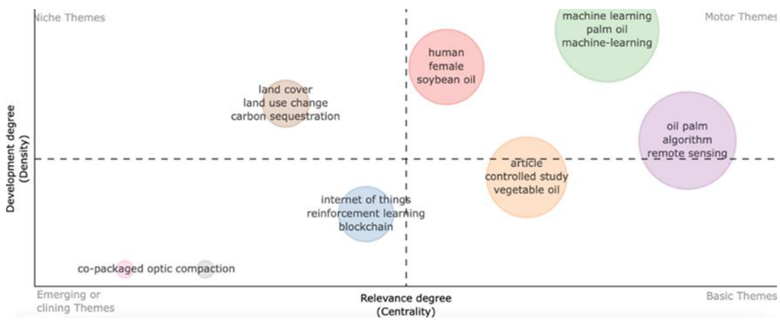

Fig. 2. Thematic classification: motor, basic, niche, and emerging categories.

3.3. Collaboration networks and geographic patterns

Research output is geographically concentrated in major palm oil-producing countries, with notable institutional productivity from Universiti Putra Malaysia, IPB University, and Universiti Teknologi Petronas. Collaboration patterns indicate a predominance of single-country studies and a smaller share of multi-country collaborations (notably regional links between Malaysia, China, and Indonesia). This fragmentation requires curated cross-site datasets and multi-site validation to build generalisable models.

3.4. Proposed AI-based PO-SCMS

The AI-Based PO-SCMS integrates vision, temporal forecasting, anomaly detection, and explainability in an edge-to-cloud architecture. The system ingests satellite time series, drone and camera imagery, production records, meteorological data, and minimal IoT signals. The ingestion pipeline georeferences, cleans, augments, and harmonizes labels. The vision module uses transfer learning with a convolutional backbone and a detection head to produce bounding boxes, counts, and ripeness labels. The forecasting module fuses tabular features and temporal sequences using an ensemble of gradient-boosted trees and a sequence model.

The anomaly module trains unsupervised encoders and isolation forests on multivariate production series to detect rare events. The explainability module provides local feature attributions for tabular outputs and saliency maps for visual outputs. Interpretable multi-fuzzy inference system (FIS)/type-2 fuzzy controllers serve as a human-readable local decision layer [15-17]. The model management layer enforces versioning, scheduled retraining, traceability, drift detection, and latency monitoring via blockchain.

The deployment strategy runs a quantized lightweight detector on edge gateways for near-real-time alerts and reserves heavy ensemble inference for cloud batch processing. Lightweight planners with informed sampling and local metaheuristics enable on-board replanning for drones and ground robots [18]. The interface exposes an API for decision support and a dashboard with explanation panels for human review. Figure 3 shows the architecture and edge-to-cloud data flow, and Table 1 maps bibliometric gaps to modules and design choices. Initial validation targets are relative yield MAE <10%, detection mAP ≈0.72 at IoU 0.5, anomaly ROC AUC >0.9, and edge inference latency <2 seconds per image.

Table 1. Gap-to-Module Mapping

Identified gap	Module / design choice	Rationale
Explainability	Explainability module (local attributions, saliency, fuzzy rules)	Improve transparency and auditability
Field-scale validation	Staged validation roadmap with two estate pilots	Calibrate models and evaluate generalizability
Limited curated datasets	Transfer learning, synthetic augmentation, data-sharing spec	Minimize labelling; standardize labels
IoT/ blockchain absent	Modular edge- cloud architecture with optional ledger/IoT slots	Support incremental adoption

3.5. Preliminary experiments and initial validation

Conceptual validation used synthetic datasets that emulate field conditions. The protocol applied block-based splits and time-aware cross-validation to test spatial and temporal generalization. The vision pipeline fine-tuned a pretrained detector on an augmented image set to approximate UAV and farm-gate views. The forecasting pipeline trained a hybrid ensemble on aggregated satellite indices paired with synthetic historical yields. The anomaly pipeline trained an autoencoder and an isolation forest on normal sequences and evaluated detection on injected events such as theft and missed harvesting. Robustness tests simulated domain shift by varying vegetation index distributions and lowering image quality. Ablation studies measured the impact of transfer learning and targeted augmentation.

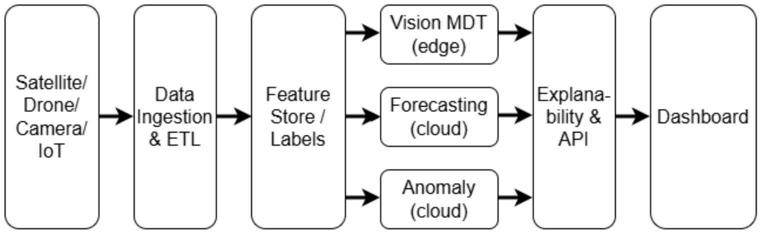

Fig. 3. AI-based PO-SCMS architecture and edge-to-cloud big data flow.

Table 2 reports baseline results from the conceptual validation using synthetic data. Each entry lists the module, dataset description, primary metric, and baseline value. Experiments used explicit versioning, fixed random seeds, and systematic logging to ensure reproducibility. Synthetic datasets included seasonal cycles, measurement noise, and realistic anomalies. Transfer learning and augmentation materially improved detection metrics. Forecasting errors met practical thresholds for block-level planning. Anomaly detectors achieved useful trade-offs between true-positive rates and false alarms.

Table 2. Baseline results from conceptual validation using synthetic data.

Module	Dataset / Methodology	Metric	Value
Yield prediction	Sentinel-2 backtesting with aligned synthetic yields	MAE (relative)	0.098
Vision detection	Augmented UAV-like images (2k annotated, synthetic)	mAP @ IoU 0.5	0.72
Ripeness classification	Same synthetic image set	Accuracy	0.88
Anomaly detection	Synthetic multivariate production time series	ROC AUC	0.91
Anomaly operating point	Same synthetic series	False positive rate (FPR) at true positive rate (TPR) 0.8	0.07

The next stage translates synthetic results into controlled field trials. Performance targets will be updated after empirical calibration. The system will be piloted at two mid-size estates (≈500-1,000 ha) for six months to collect field validation data. Pilot stakeholders include estate managers, agronomists, field technicians, and processor representatives. Required data: labelled imagery, hourly production logs, and minimal IoT streams for model calibration and evaluation. Pilot success criteria: yield MAE <10%, detection mAP ≈0.72, and anomaly AUC >0.9 after three months of continuous operation. The synthetic and augmented experiments provide controlled baselines for module performance and for robustness under designed domain shifts. These experiments do not demonstrate field-level generalizability across estates and sensor setups. We recommend staged field trials to validate operational performance and to calibrate model thresholds.

4. Conclusion

The study maps global research trends and frontiers of AI applications in palm oil supply chains for 2020-2025. Bibliometric analysis of a curated Scopus corpus of 256 articles quantifies rapid publication growth, sustained international collaboration, and a thematic concentration in ML and remote sensing. Thematic analysis identifies gaps in model explainability, field-scale validation, and curated cross-site labelled datasets. A technical design integrates vision-based bunch detection, a time-adaptive ensemble for yield prediction, unsupervised anomaly detection, and an explainability module in an edge-to-cloud deployment for AI-Based PO-SCMS. Laboratory proof-of-concept studies should validate system performance. Multi-site field trials should assess generalizability across production contexts and guide IoT and blockchain integration.

Acknowledgement

The authors would like to express their gratitude to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia for the financial support provided through the Fundamental Research Grant scheme for the 2025 fiscal year, under contract number 008/SP/DP3M/UNIKOM/VI/2025

References

1. Alhaji, A.M.; Almeida, E.S.; Carneiro, C.R.; da Silva, C.A.S.; Monteiro, S.; and Coimbra, J.S.D.R. (2024). Palm oil (Elaeis guineensis): A journey through sustainability, processing, and utilization. *Foods*, 13(17), 2814-2826.
2. Ogahara, Z.; Jespersen, K.; Theilade, I.; and Nielsen, M.R. (2022). Review of smallholder palm oil sustainability reveals limited positive impacts and identifies key implementation and knowledge gaps. *Land use policy*, 120, 106258-106265.
3. Charlebois, S.; Latif, N.; Ilahi, I.; Sarker, B.; Music, J.; and Vezeau, J. (2024). Digital traceability in agri-food supply chains: A comparative analysis of OECD member countries. *Foods*, 13(7), 1075-1087.
4. Akhtar, M.N.; Ansari, E.; Alhady, S.S.N.; and Bakar, E.A. (2023). Leveraging on advanced remote sensing-and artificial intelligence-based technologies to manage palm oil plantation for current global scenario: A review. *Agriculture*, 13(2), 504-515.
5. Junior, F.A.S.; Koeswandy, Y.P.; Nurhayati, P.W.D.; Asrol, M.; and Marimin. (2023). Annotated datasets of oil palm fruit bunch piles for ripeness grading using deep learning. *Scientific Data*, 10(1), 72-87.

6. Khan, N.; Kamaruddin, M.A.; Sheikh, U.U.; Zawawi, M.H.; Yusup, Y.; Bakht, M.P.; and Noor, N.M. (2022). Prediction of oil palm yield using machine learning in the perspective of fluctuating weather and soil moisture conditions: Evaluation of a generic workflow. *Plants*, 11(13), 1697-1708.
7. Jamshidi, E.J.; Yusup, Y.; Hooy, C.W.; Kamaruddin, M.A.; Hassan, H.M.; Muhammad, S.A.; and Tan, C.C. (2024). Predicting oil palm yield using a comprehensive agronomy dataset and 17 machine learning and deep learning models. *Ecological Informatics*, 81, 102595-102608.
8. Ryo, M. (2022). Explainable artificial intelligence and interpretable machine learning for agricultural data analysis. *Artificial Intelligence in Agriculture*, 6, 257-265.
9. Ostfeld, R.; and Reiner, D.M. (2024). Seeing the forest through the palms: developments in environmentally sustainable palm oil production and zero-deforestation efforts. *Frontiers in Sustainable Food Systems*, 8, 1398877-1398886.
10. Pohan, M.A.R.; and Utama, J. (2024). Smart probabilistic road map (Smart-PRM): Fast asymptotically optimal path planning using smart sampling strategies. *Jordanian Journal of Computers and Information Technology*, 10(2), 152-168.
11. Pohan, M.A.R.; Trilaksono, B.R.; Santosa, S.P.; and Rohman, A.S. (2024). Path planning using combined informed rapidly-exploring random tree star and particle swarm optimization algorithms. *IEEE Access*, 12, 56582-56608.
12. Pohan, M.A.R.; and Utama, J. (2023). Efficient autonomous road vehicles local path planning strategy in dynamic urban environment using RRTACS bi-directional rule templates and configuration time-space. *Journal of Engineering Science and Technology*, 18(5), 2388-2397.
13. Malik, A.; and Pohan, M.A.R. (2022). The development of a path planning algorithm combining the rapidly-exploring random tree algorithm and the particle swarm optimization algorithm. *Journal of Engineering Science and Technology*, 17(6), 3742-3754.
14. Albahar, M. (2023). A survey on deep learning and its impact on agriculture: challenges and opportunities. *Agriculture*, 13(3), 540-556.
15. Cao, J.; Zhou, T.; Zhi, S.; Lam, S.; Ren, G.; Zhang, Y.; Wang, Y.; Dong, Y.; and Cai, J. (2024). Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis—A comprehensive review. *Information Sciences*, 662, 120212-120226.
16. Pohan, M.A.R.; Utama, J.; and Herdiana, B. (2024). Novel motion planning strategy with fuzzy logic for improving safety in autonomous vehicles in response to risky road user behaviors. *ASEAN Journal of Science and Engineering*, 4(3), 471-484.
17. Pohan, M.A.R.; and Utama, J. (2025). Controlling traffic for clean air and healthy cities with multi-fuzzy inference systems for a sustainable future. *Asia-Pacific Journal of Information Technology and Multimedia*, 14(1), 324-338.
18. Pohan, M.A.R.; Trilaksono, B.R.; Santosa, S.P.; and Rohman, A.S. (2021). Path planning algorithm using the hybridization of the rapidly-exploring random tree and ant colony systems. *IEEE Access*, 9, 153599-153615.