

Review

AI-Assisted Selective Harvesting and Smart Forest Management: A State-of-the-Art Review of Multimodal Sensing and Decision-Support Approaches

Janis Peksa ^{1,2}

¹ Information Technology Faculty, Turība University, Graudu Street 68, LV-1058 Riga, Latvia; janis.peksa@turiba.lv

² Institute of Information Technology, Riga Technical University, Kalku Street 1, LV-1658 Riga, Latvia

Abstract

Selective harvesting requires tree-level decisions that balance operational productivity, stand development, and sustainable forest management, yet current AI and sensing studies often remain fragmented across remote sensing, machine perception, optimization, and forestry domains. This review synthesizes research on AI-assisted selective harvesting and smart forest management, with emphasis on multimodal sensing, decision support, and harvester-oriented deployment. A PRISMA-informed scoping review was conducted using Scopus, Web of Science Core Collection, IEEE Xplore, ScienceDirect, and SpringerLink, resulting in 82 studies retained for qualitative synthesis. The reviewed literature was organized into eight thematic groups covering smart forestry, harvesting operations, RGB-based perception, LiDAR and point-cloud processing, multimodal fusion, edge deployment, decision-support systems, and forecasting-oriented digital forestry. The analysis shows that RGB imaging provides semantic tree recognition, LiDAR enables spatial localization and structural assessment, and decision-support methods can translate tree-level observations into transparent cut/keep recommendations. However, integrated harvester-mounted systems remain underdeveloped, particularly regarding real-time RGB–LiDAR fusion, operator-facing recommendations, and forecasting integration. This review proposes a reference architecture for human-in-the-loop AI-assisted selective harvesting and identifies future research priorities for field validation and smart forest-management integration.

Keywords: harvester-mounted perception; LiDAR; RGB imaging; point clouds; edge computing; operator interface

Academic Editors: Rodolfo Picchio,
Francesco Latterini, Rachele Venanzi

Received: 4 June 2026

Revised: 13 July 2026

Accepted: 14 July 2026

Published: 16 July 2026

Copyright: © 2026 by the author.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Modern forestry is undergoing a transition, moving away from traditional, purely operator-dependent practices toward data-rich digitally assisted management. This transition is driven by the need to improve operational efficiency, support sustainable management, respond to climate-related uncertainty, and integrate recent advances in sensing, artificial intelligence, robotics, and forest-machine assistance [1–8]. Yet, if we are to successfully integrate these technologies in the field, they must remain both scientifically rigorous and practically transparent. Driven by this need, our review employs a PRISMA-

informed methodology to map the current literature and evaluate how emerging research can genuinely enable AI-assisted selective harvesting [9].

Recent works on precision forestry, “Forest 4.0”, Internet of Things (IoT)-based monitoring, human-centered artificial intelligence and digital forest-management systems show that forestry digitalization has grown far beyond traditional geographic information systems and remote inventory [1–6]. Future forestry systems are likely to rely not on isolated technological novelties, but on deeply integrated workflows that bind sensor data and machine learning to silvicultural models and operator expertise. Indeed, the recent literature covering LiDAR fusion, point-cloud analytics, and forest-product supply chains illustrates the technical breadth of this transition [7,8,10–12].

Selective harvesting provides a demanding test case for these technologies. Unlike broad, stand-level inventory tasks, selective logging demands highly localized, tree-by-tree decision-making. Choosing which trees to fell and which to retain directly impacts residual stand density, timber quality, machine efficiency, and long-term ecological health. Recent studies on mobile LiDAR-based thinning-density assistance, onboard LiDAR in cut-to-length harvesting, harvester-mounted mobile laser scanning, LiDAR-supported thinning-tree selection, and automated thinning frameworks demonstrate that forest operations research is moving toward machine-mounted and operator-facing assistance [13–17]. Concurrently, optimization studies show that these individual-tree decisions can be mathematically structured using spatial, economic, and silvicultural criteria [18–20]. Analyses of European harvesting regimes and forest resilience also show that this decision matrix is not merely technical; it is also embedded in ecological and policy realities [21–24].

AI-based visual perception has become an important research direction for tree detection and characterization in complex forest scenes. Convolutional neural networks, one-stage object detectors, instance-segmentation methods, and related deep learning models have been used for crown delineation, tree detection, species classification, live/dead tree identification, and trunk recognition [25–38]. These studies show substantial progress, especially in aerial, unmanned aerial vehicle, and remote-sensing contexts, where tree crowns and canopy structures can often be represented as image objects. However, the transfer of such models to harvester-mounted perception is not straightforward. Aerial and unmanned aerial vehicle data usually represent trees from above, whereas forest harvesters observe stems, lower crowns, understory vegetation, branches, shadows, and partial tree structures from an oblique and moving ground-level perspective. This cross-perspective domain shift changes object appearance, visible geometry, scale, occlusion patterns, illumination, and even the relevant labeling target, since harvesting support often requires stem- or tree-candidate recognition rather than crown-only detection.

This domain shift helps explain why image-based models that perform well in aerial or offline datasets may degrade under practical harvesting conditions. Ground-level forest scenes include partial visibility, overlapping stems, strong background clutter, motion and vibration, changing weather and illumination, and obstruction by neighboring vegetation or machine components. In addition, RGB imagery does not directly provide metric distance, local tree coordinates or structural attributes. Thus, lightweight detection models for logging scenarios cannot rely only on feature representations learned from aerial imagery. They require adaptation toward stem and lower-crown features, multi-scale detection, occlusion-aware training data, field-specific augmentation, and, where available, depth or LiDAR-guided spatial constraints [33,34]. In this context, RGB cameras are best understood as a semantic perception channel rather than as a complete standalone basis for selective-harvesting decisions.

LiDAR and point-cloud technologies address many of the geometric limitations of image-based perception. Terrestrial, mobile, UAV-based, and harvester-mounted laser

scanning can support tree localization, stem detection, individual-tree segmentation, structural parameter extraction, local mapping, and machine positioning [10–12,39–54]. For selective harvesting, this geometric information is critical; an intelligent retention decision cannot be made without understanding a tree's precise spatial relationship to its neighbors and the machine's mechanical reach. However, LiDAR brings its own inherent limitations, including high sensor costs, complex calibration, massive computational overhead, and degraded point density in thick brush. These limitations suggest that the most realistic near-term direction is not a single-sensor solution but a complementary sensing strategy.

Multimodal sensing is, therefore, a central theme in future harvester-oriented systems. Studies combining LiDAR with RGB, hyperspectral, multispectral, UAV, or other data sources show that fusion can improve tree segmentation, species classification, and structural parameter extraction compared with single-modality workflows [10,55–65]. The main value of fusion for selective harvesting is not only improved classification accuracy but also the creation of actionable tree-level entities that combine semantic identity, spatial position, structural descriptors, local neighborhood context, and confidence information. Nevertheless, most multimodal forestry studies are still oriented toward inventory, species mapping, or offline analysis rather than real-time operation on a moving forest machine. The unresolved problem is, therefore, how to make RGB–LiDAR alignment, calibration, synchronization, and decision-level association robust enough for field deployment.

Deployment conditions further distinguish selective harvesting from many remote-sensing or laboratory studies. Forest harvesters work under vibration, dust, weather variation, changing illumination, limited network connectivity, and continuous operator workload. Recent work on edge artificial intelligence, forest robotics, mobile mapping, and operator assistance indicates that onboard inference and local processing are feasible for tasks such as trunk detection, tree mapping, localization, and real-time assistance [63–67]. However, these studies expose a persistent tension between academic benchmarks and field realities. While large models may perform well in offline benchmarks, practical harvesting demands leaner, highly robust architectures. We must design systems that deliberately decouple real-time edge inference from the heavier tasks of offline model training and validation.

However, data collection and data processing are just prerequisites for the core task, which is the decision support. As the literature on multi-criteria decision-making illustrates, forest management requires a delicate balancing act between timber yield, carbon retention, biodiversity, and strict silvicultural rules [68–74]. In the context of a harvester, tree labeling alone is insufficient for operational decision support. It must help the operator evaluate that specific tree against a web of constraints, from minimum diameter thresholds to habitat preservation and spatial distribution. Because of this, algorithmic transparency is paramount. This requires transparent and configurable cut/keep recommendations rather than opaque automatic commands. In this respect, AI-assisted harvesting should be aligned with human-centered Forestry 5.0 principles, where digital tools support expert judgment, operator awareness, and decision consistency instead of replacing professional responsibility [4,13,64].

A similarly, pragmatic approach is needed for Digital Twins and long-term forecasting. Digital Twins in agriculture and forestry, forest-specific Digital Twin architectures, and AI-supported ecosystem-management frameworks show promising directions for linking sensing, simulation, and interactive decision environments [75–77]. Growth-model emulation, deep learning-based forest growth retrieval, LiDAR-optical monitoring, biomass estimation, harvesting-practice identification, and smart forest modeling further suggest that sensing data can support longer-term forecasting and scenario analysis [78–

84]. Recent complementary reviews also emphasize the importance of simulation in forest engineering, remote-sensing support for harvesting-operation planning and monitoring, and the emerging landscape of AI applications in forest operations engineering [85–87]. However, mature operational Digital Twins for real-time forest-harvester decision support remain limited. A more realistic near-term approach is, therefore, to treat forecasting as a digital-forestry-oriented extension: tree-level observations and harvesting actions should be recorded and structured so that they can later support stand-state updates, growth projections, and five- to ten-year harvest estimates.

The necessity of this review follows from the fragmentation of the existing literature. Important progress has been made in tree detection, point-cloud processing, multimodal sensing, harvester-mounted LiDAR, edge deployment, optimization, and forest forecasting, but these components are rarely analyzed together as parts of one selective-harvesting decision-support chain. The purpose of this review is, therefore, not only to summarize the state of the art across these domains, but also to evaluate their readiness for an integrated, harvester-oriented system. The innovative aspect of this review lies in its system-level synthesis: it connects cross-perspective tree perception, local RGB–LiDAR mapping, edge-compatible model deployment, operator-interface requirements, transparent cut/keep recommendations, and forecasting-ready data structures into a coherent reference architecture. Its significance lies in connecting forestry objectives with technical implementation requirements: this review identifies where current methods are mature, where cross-perspective and field-deployment gaps remain, and how sensing, perception, spatial mapping, operator interaction, decision support, and forecasting can be organized into a coherent framework for future research and development.

Taken together, the reviewed literature indicates that AI-assisted selective harvesting should not be framed primarily as a path toward fully autonomous logging. The more realistic and scientifically defensible near-term pathway is a human-in-the-loop decision-support architecture that combines RGB–LiDAR perception, local tree mapping, transparent cut/keep recommendations, offline model improvement, and forecasting-ready data structures. Thus, the main contribution of this review is to translate fragmented advances in sensing, AI perception, forest-machine assistance, optimization, and forecasting into an implementation-oriented roadmap for future selective-harvesting systems. Such an approach can help bridge the gap between general forest-management objectives and the complex field conditions under which selective harvesting is actually performed.

2. PRISMA-Informed Review Methodology

2.1. Review Design and Reporting Framework

This study was designed as a systematic scoping review informed by the PRISMA-ScR framework. The aim was to synthesize the current research landscape surrounding AI-assisted selective harvesting and smart forest management. Because the topic spans precision forestry, multimodal sensing, artificial intelligence, robotics and digital forecasting, the literature presents a high degree of methodological heterogeneity. A scoping review was, therefore, more appropriate than a rigid meta-analysis, as it provides a better structure for mapping diverse approaches into a cohesive overview.

The workflow of this research was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statement alongside its extension for Scoping Reviews (PRISMA-ScR) [9]. Specifically, PRISMA 2020 was used to handle the mechanics of literature identification and screening, and PRISMA-ScR to shape data-charting and thematic synthesis. This hybrid approach aligns well with current MDPI guidelines for systematic mapping and scoping [9].

2.2. Review Objective and Research Questions

This review aims to aggregate and classify research that can drive the development of intelligent, harvester-mounted decision-support systems for smart forestry. Rather than isolating a single technology, the goal was to understand how sensing, AI-based perception, spatial localization, and operational forecasting can interlock within a real-world forestry framework.

To meet the aim, the review was framed around seven key research questions:

- RQ1. Which sensing modalities are currently used to support tree-level recognition, and individual-tree localization as well as structural assessment in forest environments?
- RQ2. Which AI-based approaches are most relevant for individual-tree detection, stem recognition, tree-species identification, and tree-level characterization under forestry conditions?
- RQ3. How are RGB imaging, LiDAR, point clouds, multimodal fusion and similar approaches used to improve tree-level perception and spatial mapping?
- RQ4. What forms of decision support have been proposed for selective harvesting and thinning operations aiming for operator assistance?
- RQ5. How are edge computing, real-time deployment, human-in-the-loop interaction and other assistive technologies addressed in forestry-oriented AI systems?
- RQ6. How do forecasting models and Digital Twin-oriented concepts extend tree-level sensing toward smart forest management?
- RQ7. What research gaps currently limit the transition from isolated sensing or AI methods to integrated, operational, harvester-oriented decision-support architectures?

Answering these questions allows us to build a thematic state-of-the-art synthesis, which later was used to outline a reference architecture for AI-assisted selective harvesting.

2.3. Information Sources and Database Selection

To ensure balanced coverage across engineering, computer science, and forestry disciplines, the records were retrieved from five scholarly databases:

- Scopus;
- Web of Science Core Collection;
- IEEE Xplore;
- ScienceDirect;
- SpringerLink.

Scopus and Web of Science served as primary multidisciplinary databases. IEEE Xplore was integrated specifically to capture engineering-oriented literature focused on computer vision, edge AI, mobile robotics, LiDAR processing and other engineering solutions in the field. ScienceDirect and SpringerLink were targeted to capture the domain-specific literature in forest operations, growth modeling, silvicultural decision-support studies.

The final query was executed on 1 May 2026. The main search window was limited to publications between 2015 and 2026 to capture the growth of deep learning and smart-forestry research during this decade. A few foundational papers published prior to 2015 were included selectively if they established critical baseline concepts for forest optimization or growth modeling.

2.4. Search Strategy

Because no single search string could adequately cover such a cross-disciplinary topic, the search strategy was broken into three complementary “query families.” Each

family was tailored to a distinct dimension of the review: (A) perception, (B) operations, and (C) forecasting.

The syntax rules were adapted to fit each database (e.g., using title, abstract, keyword fields in Scopus versus topic fields in Web of Science) but strictly maintained the underlying Boolean logic across all platforms.

2.4.1. Search Family A: Multimodal Sensing and AI-Based Tree Perception

This family focused on visual tree detection, point-cloud processing, and multimodal sensor fusion:

(forest* OR forestry OR woodland* OR "forest stand*" OR silvicultur*) AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "computer vision" OR YOLO OR transformer* OR "object detection" OR "instance segmentation" OR "semantic segmentation") AND (LiDAR OR "laser scanning" OR "point cloud*" OR RGB OR camera* OR "multimodal sensing" OR "sensor fusion") AND (tree* OR stem* OR crown* OR "individual tree" OR "tree detection" OR "tree localization" OR "tree species")

2.4.2. Search Family B: Selective Harvesting and Decision-Support Systems

This query linked sensing technologies to operational realities like thinning, machine assistance, and cut-to-length harvesting:

("selective harvesting" OR thinning OR "tree selection" OR "forest harvester*" OR "cut-to-length" OR "forest operation*" OR "harvesting operation*") AND ("decision support" OR optimization OR ranking OR "operator assistance" OR "operator support" OR "advanced driver assistance" OR "tree marking" OR "harvesting recommendation*") AND (LiDAR OR sensing OR "machine learning" OR AI OR "digital forestry" OR "precision forestry")

2.4.3. Search Family C: Forecasting, Growth Modeling, and Digital-Forestry Extensions

This family targeted the long-term management aspects, such as stand development and simulation:

("forest management" OR forestry OR "forest planning" OR "smart forestry" OR "digital forestry" OR "forest stand*") AND (forecast* OR "growth model*" OR "yield model*" OR "Digital Twin*" OR simulation OR "scenario analysis" OR "future harvest*" OR "stand development") AND ("decision support" OR LiDAR OR AI OR "machine learning" OR "remote sensing" OR multimodal OR sensing)

2.5. Eligibility Criteria

Eligibility criteria were established as a strict set of inclusion and exclusion parameters designed to exclude papers that lacked a clear operational or smart-forestry connection.

2.5.1. Inclusion Criteria

Studies were retained if they met the following baseline criteria:

- Focused on forestry, forest operations, inventory, silviculture, or precision forestry;
- Addressed at least one of the core technical areas: AI-based tree detection/segmentation; LiDAR/point-cloud localization; multimodal sensing; machine-mounted/mobile perception; DSS for harvesting; or Digital Twin-oriented forecasting;
- Presented peer-reviewed journal findings or substantial conference proceedings;
- Published in English;

- Published between 2015 and 2026 (excluding foundational papers added manually).

2.5.2. Exclusion Criteria

Conversely, studies were excluded if they:

- Focused exclusively on non-forest agricultural settings (e.g., orchards);
- Addressed only broad, landscape-scale remote sensing without actionable relevance to tree-level assessment or operational harvesting;
- Discussed AI or automation in vague, generic terms without substantive application to forestry;
- Lacked sufficient methodological depth to be classified accurately;
- Were editorials, promotional materials, theses, or non-peer-reviewed reports;
- Duplicated earlier versions of research already covered by a more comprehensive; updated publication.

2.6. Study Identification and Screening Procedure

The literature was processed through a standard four-step workflow: identification, deduplication, title/abstract screening, and full-text assessment.

At the first stage, records were retrieved independently from the five targeted databases using proposed search families, supplementing them with forward and backward citation snowballing from highly relevant cornerstone papers. The raw exports were then consolidated into a centralized screening matrix and duplicates were stripped out via DOI matching, exact-title checks, and manual verification to prevent double-counting of conference-to-journal extensions.

During the first screening phase, titles and abstracts were reviewed against predefined eligibility criteria, quickly eliminating research that had no bearing on tree-level sensing, smart forestry, or operational decision support. Full texts of the remaining candidate papers were then retrieved. Studies that failed to offer usable methodological data or lacked a concrete connection to selective harvesting and smart management were dropped at this stage, with the specific reason for exclusion logged in the screening matrix.

2.7. Study Selection Results

The study-selection process is summarized in Figure 1. The initial database search retrieved 1963 records from Scopus, Web of Science Core Collection, IEEE Xplore, ScienceDirect, and SpringerLink. An additional 28 records were identified through backward and forward citation snowballing, resulting in 1991 records before duplicate removal. After removing 718 duplicate records, 1273 records remained for title and abstract screening.

The title and abstract screening stage excluded 1084 records that did not meet the thematic or methodological scope of the review. Most exclusions at this stage concerned studies outside forestry, papers without tree-level relevance, or studies addressing broad remote sensing without a direct connection to selective harvesting, machine-mounted perception, or smart-management decision support. The remaining 189 full-text articles were assessed for eligibility. At the full-text stage, 107 articles were excluded because they had insufficient forestry relevance, lacked a tree-level or selective-harvesting connection, did not provide a meaningful decision-support or smart-management contribution, focused only on regional or landscape-scale monitoring without operational relevance, or lacked sufficient methodological detail. In total, 82 studies were retained for the final qualitative synthesis.

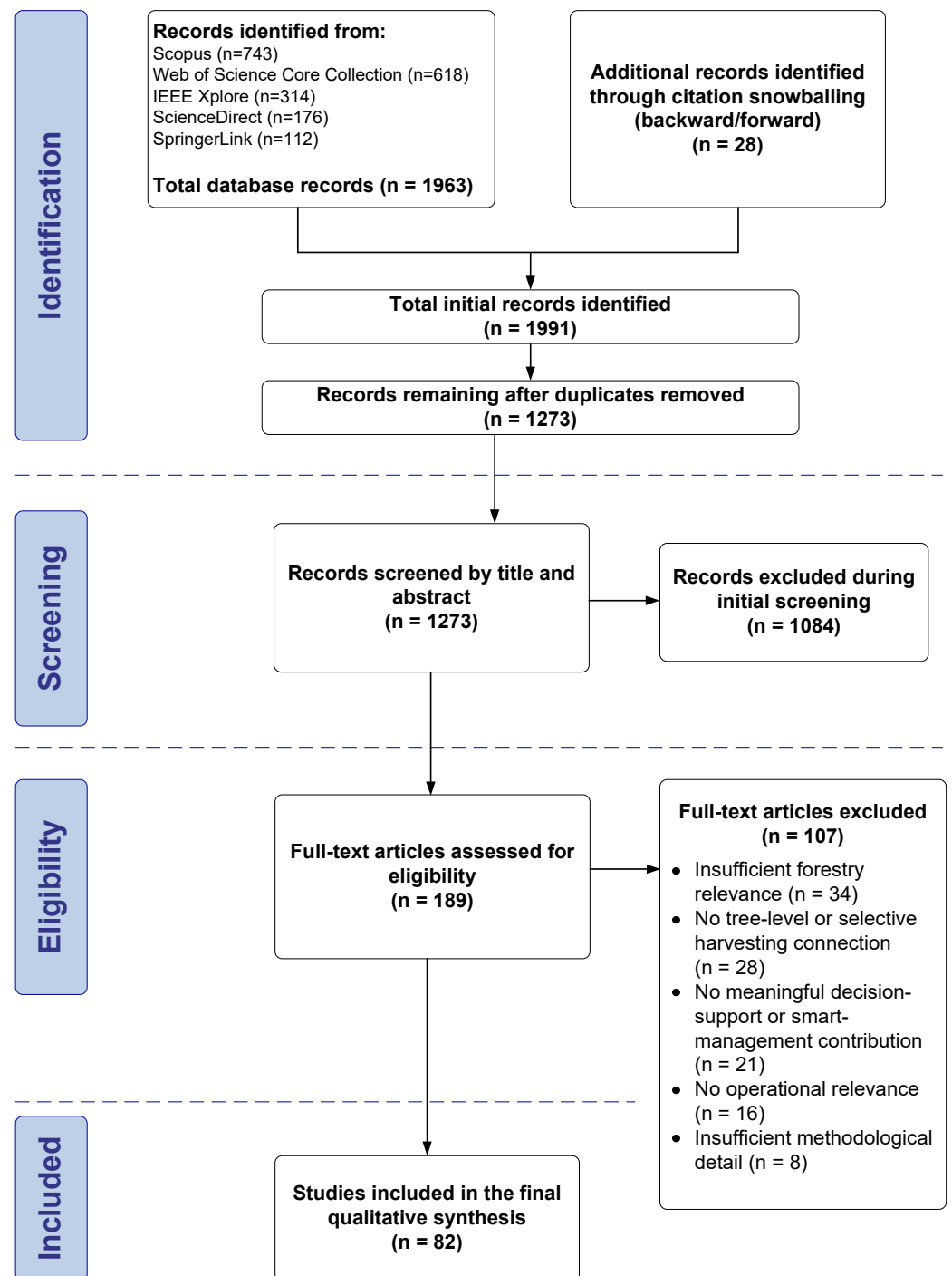


Figure 1. PRISMA-style flow diagram of the literature identification, screening, eligibility assessment, and inclusion process.

2.8. Data Extraction and Coding Framework

A standardized data-extraction matrix was developed to ensure consistent classification of the retained studies. Each of the 82 retained studies was tagged with specific bibliographic, technical, operational, and analytical attributes (Table 1).

Table 1. Main variables extracted from the reviewed studies.

Extraction Dimension	Main Coded Variables
Bibliographic info	Authors, year, venue, DOI
Forestry context	Inventory, thinning, harvesting, forest management
Sensing modality	RGB, LiDAR, point cloud, multimodal
Observation perspective	Aerial, terrestrial static, mobile, harvester-mounted
Analytical method	Detection, segmentation, optimization, DSS, forecasting
System output	Tree detection, localization, recommendation, forecast
Deployment	Offline, edge, real-time, operator assistance
Evaluation metrics	Accuracy, mAP, F1, RMSE, latency, operational outcome
Reported limitations	Occlusion, calibration, data scarcity, cost, generalization

2.9. Thematic Synthesis Strategy

Because the retained studies used diverse methods and evaluation settings, a qualitative thematic synthesis was used instead of statistical meta-analysis. Based on the data extraction, the 82 retained studies were classified into eight distinct analytical categories:

1. Smart forestry, precision forestry, and digital transformation (10 records);
2. Selective harvesting, thinning, and forest-harvester operations (12 records);
3. RGB and computer-vision methods for tree detection (12 records);
4. LiDAR, point-cloud processing, and tree localization (15 records);
5. Multimodal sensing and RGB–LiDAR fusion (10 records);
6. Edge AI, robotics, and operator-facing deployment (6 records);
7. Decision-support systems, optimization, and tree selection (8 records);
8. Forecasting, growth modeling, and Digital Twin-oriented forestry (9 records).

Each study was assigned to one primary thematic category for counting purposes, although some studies contributed conceptually to more than one section of the synthesis. This framework was then used to structure the state-of-the-art synthesis in the following section, which serves as the blueprint for the proposed reference architecture.

2.10. Methodological Scope and Limitations

While this review was structured to be as comprehensive as possible across several disciplines, several inherent limitations also should be acknowledged. First of all, restricting the search to English-language, peer-reviewed literature means that valuable regional research published in other languages may have been missed. Also, the primary 2015–2026 window inherently prefers modern deep learning developments, meaning some older, conceptually foundational papers may be underrepresented unless revealed via citation snowballing. Another point is the lack of standardized terminology across forestry, robotics, and computer science means that, despite running three distinct search families, some studies, which might use highly niche or unconventional phrasing, may have been omitted.

At last, it should be mentioned that technologies such as multimodal sensor fusion and Digital Twins are evolving at a high speed. As a result, this review should be viewed as a snapshot of the literature up to 1 May 2026, rather than an exhaustive endpoint for the field.

3. Results: State-of-the-Art Landscape of AI-Assisted Selective Harvesting and Smart Forest Management

3.1. Literature Landscape and Thematic Taxonomy

The 82 retained studies show that research relevant to AI-assisted selective harvesting is not concentrated in a single disciplinary stream. In contrast, the field is distributed across smart forestry, precision forestry, remote sensing, computer vision, LiDAR-based forest inventory, forest-machine assistance, optimization, and digital-forestry forecasting. Smart-forestry and Forest 4.0 studies provide the broader conceptual frame, emphasizing that forestry digitalization depends on coordinated sensing, data quality, interoperability, and human-centered decision support rather than on isolated digital tools [1–8].

Within this landscape, selective harvesting occupies a particularly demanding position. It requires tree-level observations, local spatial awareness, operator-facing interpretation, and silvicultural decision logic. Recent studies on mobile LiDAR thinning-density assistance, onboard LiDAR in cut-to-length harvesting, harvester-mounted mobile laser scanning, and automated thinning-tree selection demonstrate that the operational forestry community is beginning to move from post facto analysis toward real-time or near-real-time assistance for machine operators [13–17].

The retained studies were grouped into eight thematic groups (Table 2). The first group concerns the digital transformation of forestry and precision forestry concepts. The second group focuses on selective harvesting, thinning, and forest-harvester operations. The third and fourth groups address AI-based visual perception and LiDAR-based tree localization. The fifth group covers multimodal fusion, especially RGB-LiDAR and multi-source remote-sensing integration. The sixth group concerns edge AI, robotics, localization and operator-facing deployment. The seventh group addresses decision-support systems and optimization. The last, the eighth group, covers forecasting, growth modeling and Digital Twin-oriented forest management. This taxonomy provides the structure for the following synthesis.

Table 2. Main thematic streams identified in the retained literature and their relevance to AI-assisted selective harvesting.

Thematic Stream	Typical Focus	Contribution to Selective Harvesting	Main Unresolved Limitation	Representative References
Smart forestry and Forest 4.0	Digital transformation, IoT, precision forestry, human-centered AI	Provides the conceptual basis for intelligent and interoperable forest operations	Often remains at conceptual, infrastructure, or stand-management level	[1–8]
Selective harvesting and forest operations	Thinning, cut-to-length harvesting, harvester assistance, operational constraints	Defines the practical decision environment for operator support	Limited integration with real-time perception and forecasting	[13–24]
RGB and computer vision	Tree detection, crown delineation, trunk recognition, tree-species classification	Provides semantic tree-level perception	Sensitive to occlusion, lighting, viewpoint changes, and lack of metric depth	[25–38]
LiDAR and point clouds	Stem mapping, localization, DBH/height proxies, structural attributes	Provides geometry, local tree position, and structural context	Computational burden, calibration, and field robustness remain challenging	[10–12,39–54]

Multimodal fusion	RGB-LiDAR, LiDAR-hyperspectral, multi-platform mapping	Combines semantic and geometric evidence into tree-level entities	Requires calibration, synchronization, and more complex data handling	[10,55–62,65]
Edge AI and deployment	Onboard inference, robotics, localization, operator interfaces	Supports real-time operation under field constraints	Few complete machine-mounted systems have been validated in forestry	[13–15,47,48,53,63–67]
Decision support and optimization	MCDM, tree-level optimization, thinning rules, harvesting-system selection	Transforms sensing outputs into actionable management recommendations	Often disconnected from real-time sensing streams	[16–20,68–74]
Forecasting and digital forestry	Growth modeling, Digital Twin concepts, scenario analysis, future harvest assessment	Links current interventions to future stand condition and harvest potential	Operational Digital Twin implementations remain immature	[75–84]

3.2. Multimodal Sensing for Tree-Level Assessment

3.2.1. Image-Based Sensing

Red–green–blue (RGB) imagery remains the most accessible sensing modality for tree-level visual interpretation. It supports tree-crown detection, trunk recognition, and tree-species classification, as well as operator-oriented visualization. Reviews of CNN-based individual-tree crown detection and drone-based tree-species identification confirm that image data continue to prevail in many deep learning forestry workflows because of their relatively low cost, direct interpretability and compatibility with already well-established detection and segmentation models [25,26].

Nevertheless, RGB data alone are insufficient for many selective-harvesting tasks. In general, forest scenes present strong occlusion, irregular lighting, repeated stem-like structures, and background clutter. Even though aerial RGB studies have achieved strong results in crown detection and delineation [29–32,38], the ground-level and machine-mounted perspective, which is more difficult because trees are partially visible, overlap strongly, and are observed from a moving platform, still remains understudied. For this reason, RGB imaging is best understood as a semantic perception channel rather than as a complete standalone basis for selective-harvesting decisions.

3.2.2. LiDAR and Point-Cloud Sensing

LiDAR complements RGB imaging by providing metric geometry. Terrestrial and mobile point-cloud studies demonstrate that LiDAR can support individual-tree segmentation, stem detection, structural parameter extraction, local forest mapping and spatial structure assessment [10–12,39–46]. This is central for selective harvesting, where a detected tree must be linked at least to its spatial position and neighbor relationships. It is also useful to evaluate machine accessibility and measurable structural attributes.

The operational relevance of LiDAR is particularly evident in recent harvester-oriented studies. Sagar et al. investigated onboard LiDAR applications for cut-to-length harvesting [14], while Faitli et al. integrated a mobile laser scanning system with a forest harvester to support localization and tree-stem measurements [15]. Pohjala et al. evaluated a mobile LiDAR-based thinning-density assistant and demonstrated that the technology can directly support harvester operators during thinning [13]. These and other studies in the field indicate that mobile and machine-mounted LiDAR are moving from inventory-focused research toward operational decision assistance.

3.2.3. Complementary Sensing for Operational Forestry

The reviewed literature supports a complementary sensing strategy rather than a single-sensor design. RGB imaging is better suited for visual recognition and operator interpretation, while LiDAR provides distance, geometry, and local spatial structure. Multi-modal studies combining LiDAR, RGB, hyperspectral, multispectral, or multi-platform data show that fusion can improve tree segmentation, species classification, structural estimation and the extraction of semantic attributes [8,55–62,65]. For selective harvesting, the key value of fusion is not simply higher classification accuracy, but the creation of actionable tree-level objects that may combine several valuable parameters, including identity, location, structural attributes and local context. Table 3 summarizes the relevance of major sensing modalities for AI-assisted selective harvesting.

Table 3. Comparative relevance of major sensing modalities for AI-assisted selective harvesting.

Sensing Modality	Typical Focus	Main Limitation	Contribution to Selective Harvesting	Representative References
RGB camera	Visual semantics, appearance, tree/trunk detection, operator display	No direct metric depth; sensitive to illumination and occlusion	Essential baseline for semantic recognition	[25–38,63]
UAV or aerial imagery	Stand-scale context, crown mapping, species and biomass assessment	Limited correspondence with machine-level ground perspective	Useful for planning and model pretraining, less sufficient for real-time harvester guidance	[25–32,38,55–62]
Terrestrial laser scanning	Dense 3D forest structure and high-quality reference data	Mostly static or survey-oriented; limited direct operational use	Important for reference datasets and algorithm validation	[10–12,40–46,54]
Mobile or harvester-mounted LiDAR	Tree position, local mapping, stem measurements, machine localization	Requires robust calibration, processing, and field integration	Highly relevant to real-time selective-harvesting support	[13–15,39,47–54]
Multimodal sensing	Combines semantics, geometry, and structural attributes	Increased calibration, synchronization, and data-management complexity	Most promising pathway toward actionable tree-level decision support	[10,55–62,65]

3.3. AI Methods for Individual-Tree Detection and Characterization

3.3.1. Deep Learning-Based Tree Detection

Deep learning has become the dominant paradigm for tree detection and characterization. Systematic reviews show that convolutional neural network-based methods have substantially improved individual-tree detection and crown delineation compared with classical image-processing methods [25,26]. One-stage detectors, including YOLO-type models, are particularly attractive for forestry applications because they provide a favorable balance between inference speed and detection performance, as illustrated by recent studies using YOLO-based models for tree detection and trunk recognition [27,28,33,34].

However, the literature also shows that high performance in aerial or satellite imagery does not automatically translate into robust harvester-mounted perception. Ground-level scenes introduce partial visibility, moving viewpoints, shadows, understory clutter, and occlusion by neighboring stems or branches. These factors can reduce detection reliability and make direct transfer from aerial or satellite benchmarks unreliable. Therefore, one-stage detectors are best interpreted as strong baseline models for future harvester systems, while their performance must be validated under operational field conditions rather than assumed from remote-sensing benchmarks alone.

3.3.2. Cross-Perspective Domain Shift in Harvester-Mounted Perception

A key reason for degraded transfer from aerial or unmanned aerial vehicle datasets to harvester-mounted perception is cross-perspective domain shift. Aerial and UAV imagery often represents trees as crowns or canopy objects observed from above. In contrast, a harvester-mounted camera observes stems, lower crowns, branches, understory vegetation, machine parts, shadows, and partially visible tree structures from a moving ground-level or oblique perspective. This changes the visible object geometry, the spatial scale of candidate trees, the occlusion pattern, and the relationship between the labeled object and the operational decision [25–38,63].

The target definition also changes. In aerial tree-detection studies, the label is often an individual crown or crown boundary. In selective harvesting, the relevant object may be a stem, a tree candidate, a retained tree, or a tree marked for possible removal. This creates a label-space mismatch: a detector trained to recognize crown-like image regions may not reliably identify partially visible trunks or lower-crown structures needed for harvester assistance. Ground-level perception must also tolerate motion blur, vibration, rapidly changing distance, strong illumination gradients under the canopy, seasonal variation, and obstruction by neighboring trees or understory vegetation [33,34,63–67].

For this reason, performance indicators reported for aerial or offline datasets should not be treated as direct evidence of field readiness for selective harvesting. The more relevant evaluation setting is a machine-perspective dataset collected from ground-level or harvester-mounted sensors, preferably synchronized with LiDAR and annotated not only for tree presence but also for operational relevance. This does not reduce the value of aerial datasets; they remain useful for pretraining, stand-level context, and broader forestry mapping. However, harvester-mounted decision support requires additional domain adaptation and field validation.

3.3.3. Feature Adaptation Strategies for Lightweight Logging-Scenario Models

Lightweight detection models for logging scenarios should be optimized for the features that remain stable under ground-level, machine-mounted forest observation. In contrast to aerial crown detection, selective-harvesting assistance requires stronger attention to stems, lower-crown geometry, visible bark and branching patterns, local tree spacing, and partial tree evidence. Therefore, feature adaptation should prioritize trunk- and stem-oriented representations, multi-scale features for trees located at different distances, and robustness to occlusion and background clutter [33,34,63].

Based on the reviewed literature, several practical directions for adaptation arise. The most important thing is that training datasets must shift away from a strict reliance on aerial data to incorporate ground-level or harvester-perspective imagery. To ensure real-world robustness, data augmentation strategies should actively simulate field-specific degradations. This means accurately reproducing conditions such as shadows, sun patches, motion blur, airborne dust, partial occlusions, dense undergrowth, seasonal variations, and fluctuating camera distances. Also, model architectures must be designed to maintain low latency while simultaneously improving their sensitivity to challenging features, particularly thin vertical structures and partially obscured objects. For the systems, when LiDAR is available, leveraging geometric constraints offers significant advantages; these constraints can be used to filter candidate regions, facilitate distance-aware interpretation, and accurately map image-based detections to localized point-cloud clusters [13–15,39–54,63].

In operational terms, the most important metric may not be only mean average precision on an offline test set. A harvester-oriented detector should be evaluated by missed candidate trees, false candidate recommendations, temporal stability across frames, inference latency, and the reliability of subsequent association with LiDAR-derived tree ob-

jects. This means that lightweight models should be optimized for the full perception-to-decision chain rather than for image detection alone.

3.3.4. Instance Segmentation and Tree-Level Delineation

Instance segmentation methods, especially Mask R-CNN-based workflows, are important when the objective is to separate individual trees or crowns rather than simply detect tree-containing regions. Several studies demonstrate the usefulness of Mask R-CNN for individual-tree crown delineation, live/dead tree classification, biomass-related assessment, and multi-temporal crown mapping [29–32]. These methods are particularly relevant to dense or overlapping stands, where bounding boxes may be insufficient to separate adjacent trees.

For selective harvesting, instance segmentation can support more reliable RGB-LiDAR association by providing more precise object boundaries. Its main drawbacks are the higher annotation burden and greater computational cost. Consequently, segmentation appears most suitable as an advanced component for difficult scenes, dataset preparation, or validation subsets, while lightweight detection models may remain preferable for real-time onboard operation.

3.3.5. Ground-Level and Deployment-Oriented Visual Perception

A smaller but particularly relevant part of the literature addresses ground-level trunk detection and deployment-oriented visual perception. Lightweight trunk-recognition models and forest robotics studies show that real-time or near-real-time perception is technically feasible, but accuracy must be balanced against latency, hardware limitations, and environmental variability [33,34,63]. This research direction is important for harvester-oriented systems because operator assistance requires timely and interpretable outputs, not only high offline benchmark scores.

Table 4 summarizes the main AI perception families and their expected roles in selective harvesting systems.

Table 4. Main AI perception families and their expected roles in selective-harvesting systems.

AI/Model Family	Typical Output	Main Strength	Main Limitation	Recommended Role	Representative References
One-stage object detectors	Bounding boxes and confidence scores	Fast inference and mature implementation ecosystem	Less precise separation in dense or overlapping scenes	Operational baseline for onboard tree/trunk detection	[27,28,33,34]
Convolutional segmentation and Mask R-CNN-based workflows	Per-object masks and bounding boxes	Better delineation of individual crowns or trees	Higher annotation and computational cost	Advanced analysis for dense stands and validation datasets	[29–32]
Multisource classification models	Species or condition labels	Combines visual and structural cues	Often trained in aerial or survey contexts	Useful for extending pine-only detection toward broader species recognition	[35–38,55–62,65]
Lightweight ground-level detectors	Stem/trunk candidates	Good fit for edge deployment and machine perception	Narrower representation of full tree condition	Practical component for real-time harvester assistance	[33,34,63]
Harvester-perspective-adapted models	Tree or stem candidates with field-oriented confidence scores	Better alignment with machine-mounted perception and operational constraints	Requires dedicated ground-level datasets and validation	Preferred direction for future selective-harvesting prototypes	[13–15,33,34,47,53,63–67]

3.4. LiDAR-Based Localization and Tree Attribute Extraction

LiDAR-based localization and attribute extraction form one of the most important technical foundations for AI-assisted selective harvesting. Mobile LiDAR studies show that point clouds can be used in many forestry applications, such as forest inventory, stem-attribute estimation, operational mapping and related tasks [39,49,52]. Ground-based and terrestrial point-cloud studies further demonstrate the potential of deep learning and geometric methods for tasks such as individual-tree segmentation, trunk detection and structural parameters extraction [40–46,54].

The relevance of these methods increases when they are connected to forest-machine operation. Harvester-mounted mobile laser scanning has been used for tree-stem measurements and localization [15], while recent work on LiDAR-enabled tree-map matching addresses real-time and drift-free harvester positioning [53]. Such approaches are important because a selective-harvesting assistant must identify not only whether a tree is visible, but also where the tree is located relative to the machine and to other retained or candidate stems.

At the same time, LiDAR-based processing remains challenging. Dense point clouds can be computationally demanding, and forest scenes contain leaning stems, occlusion, understory vegetation, variable terrain, and irregular point density. Reviews of terrestrial point clouds and software tools note that reproducibility, benchmark availability, and robust processing workflows remain important barriers [10–12]. Therefore, the most practical near-term strategy is to combine interpretable point-cloud preprocessing and clustering with AI-based visual perception, rather than relying immediately on a fully end-to-end learned 3D forest-understanding pipeline.

A direct universal quantitative comparison between geometric trunk extraction and deep learning instance segmentation cannot be reported from the reviewed literature because the available studies differ in scanner type, point density, species composition, platform motion, hardware, annotation protocol, and evaluation metrics [40–48,54]. Nevertheless, the reviewed evidence supports a clear operational trade-off. Geometric preprocessing and clustering are usually more interpretable and better suited to low-latency onboard filtering, but they can be brittle under occlusion, leaning stems, and irregular point density. Deep learning-based 3D segmentation has stronger potential for complex forest structures when trained on suitable datasets, but it requires more annotated data, higher computational resources, and careful validation before field deployment. For selective harvesting, a hybrid workflow is, therefore, the most defensible near-term approach: use geometric rules and clustering to extract candidate spatial objects, and use image-based or multimodal AI to improve semantic interpretation and decision relevance.

The trade-offs between geometric point-cloud extraction, deep learning segmentation and hybrid approaches for selective-harvesting support derived from this review are summarized in Table 5.

Table 5. Review-based trade-offs between geometric point-cloud extraction, deep learning segmentation and hybrid approaches for selective-harvesting support.

Approach	Expected Accuracy Profile	Expected Latency Profile	Field Robustness	Interpretability	Main Limitation	Representative References
Geometric rule-based stem extraction	Moderate to high when stems are clear and approximately cylindrical	Low to moderate; suitable for preprocessing	Sensitive to leaning stems, occlusion, undergrowth, and irregular point density	High	May fail in complex or cluttered stands	[40,51,54]
Clustering with model fitting	Moderate; useful for separating	Moderate, depending on	Moderate; parameter-sensitive un-	High to moderate	Requires careful tuning and may	[39,40,49,52]

	nearby spatial objects	point density and clustering parameters	der dense under-story		merge or split stems incorrectly	
Deep learning-based 3D instance segmentation	Potentially high on suitable annotated datasets	Moderate to high computational load; hardware-dependent	Potentially strong in complex structure, but dataset-dependent	Lower than geometric methods	Requires labeled point-cloud datasets and field validation	[41–45]
Mobile or harvester-mounted mapping pipelines	Moderate to high for localization and stem mapping when calibrated	Depends on sensor rate, localization pipeline, and onboard hardware	Field-relevant but sensitive to calibration, vibration, and platform motion	Moderate	Requires machine integration and robust synchronization	[13–15,47,48,53]
Hybrid geometric preprocessing plus AI visual perception	Balanced; combines spatial filtering with semantic interpretation	Moderate; more deployable than full end-to-end 3D learning	More robust than either modality alone if calibration is stable	Moderate to high	Requires sensor synchronization and association logic	[10–15,33,34,39–54,63–67]

This table should not be interpreted as a universal numerical benchmark. Instead, it summarizes the trade-offs identified in this review and should guide future quantitative experiments. A controlled harvester-oriented benchmark should report tree-detection precision and recall, stem-localization error, instance association accuracy, end-to-end latency, dropped-frame rate, and robustness under occlusion, vibration, illumination change, and different stand densities.

3.5. Multimodal Fusion of Image and Laser Scanning Data

Multimodal fusion is repeatedly identified as a key direction in smart forestry and forest-resource monitoring [8,60]. Fusion is especially relevant for selective harvesting because the system must combine semantic recognition with geometric localization. RGB imagery can identify likely tree or species candidates, while LiDAR can provide position, distance, structural context and other spatial parameters. Studies using UAV LiDAR, RGB, hyperspectral, multispectral, and other sources demonstrate that fused data can improve many typical tasks for smart forestry, such as tree segmentation, species classification, and structural parameter extraction [55–62,65].

The reviewed studies suggest three general fusion strategies: early fusion, feature-level fusion, and late fusion. Early and feature-level fusion may extract richer joint representations, but they require well-aligned multimodal datasets, careful calibration, and more complex training. Late fusion, in which independently produced detections and LiDAR objects are associated after separate processing, is more modular and interpretable. For an initial harvester-oriented implementation, late RGB-LiDAR fusion is, therefore, the most realistic starting point. It allows visual detections to be linked with LiDAR-derived clusters, local tree coordinates, and structural descriptors, while preserving transparency for operator-facing decision support.

The main unresolved issue is that most multimodal forestry studies are still oriented toward inventory, species mapping, or aerial/survey workflows rather than real-time forest-machine operation. The next research step is, therefore, not only improving fusion accuracy, but also validating whether fusion outputs can reliably support thinning, cut/keep recommendations, and operator decisions under moving-machine conditions.

3.6. Edge Computing and Real-Time Forestry Deployment

Real-time deployment is a defining requirement for AI-assisted selective harvesting. Forest operations often occur in areas with unreliable connectivity, variable lighting, vibration, dust, changing machine viewpoints, variable weather and seasonal variation. Therefore, many critical functions must be executed locally on the machine rather than relying on cloud processing. Edge AI studies and forestry robotics research show that onboard inference is feasible for trunk detection and forest-monitoring tasks, although model selection must consider latency, power use, memory, and robustness [63,67].

The edge-AI study by da Silva et al. is particularly relevant because it evaluates tree-trunk detection across deployment-oriented hardware categories [63]. Real-time assistance using augmented reality (AR) or operator displays has also been explored for forest harvesters, suggesting that human-machine interaction is becoming a practical research topic rather than a purely conceptual addition [64]. These works support the view that the near-term role of AI in harvesting should be human-in-the-loop assistance, not full autonomy.

The human-centered forestry AI literature reinforces this point. Forestry 5.0 emphasizes the role of expert users, explainability, and operator acceptance [5]. In selective harvesting, this means that AI should help operators maintain thinning targets, identify candidate trees, detect structural conditions, and reduce cognitive burden. Final responsibility, however, still remains with the human operator.

3.7. Decision-Support Approaches for Selective Harvesting

The reviewed literature shows that currently, decision-support systems in forestry address ecosystem services, planting, harvesting-system selection, multi-criteria analysis, and management planning in typical scenarios [22,68,69]. However, many DSS approaches operate at stand or planning scales and are not directly connected to real-time machine perception. This creates a gap between sensing outputs and operational recommendations.

Selective harvesting requires decisions about individual trees under several constraints, the most important of which are silvicultural, ecological, operational and economic. Recent optimization studies show that tree-level decisions can be formalized through mathematical programming, spatial connectivity constraints, and multi-objective models, as well as optimization of thinning or harvesting operations [18–20,70–74]. These studies demonstrate that algorithmic tree selection is an established and growing research direction.

The most directly relevant examples link sensing-derived structural information with thinning decisions. Lee et al. developed and evaluated a thinning tree-selection system using optimization techniques and multi-platform LiDAR data [16], while Yang et al. proposed an automated forest-thinning framework for selective harvesting at the individual-tree level [17]. These studies implement a decision-support structure in which candidate trees are first detected and localized, then filtered by retention rules, ranked by utility or competition indicators, and finally presented to the operator as cut/keep recommendations. For a forest-harvester assistant, the decision layer should therefore be constraint-aware, interpretable and configurable according to the current silvicultural objective.

3.8. Forecasting and Digital-Forestry Extensions

The reviewed literature also indicates that smart harvesting systems should eventually connect immediate tree-level decisions with longer-term forest development. Digital Twin concepts have been reviewed for agriculture and forestry [75]. Forest-specific Digital Twin frameworks generally propose the integration of spatiotemporal data, simulation and interactive management environments [76,77]. These studies are important because

they suggest that future forest machines may become data-generating nodes within broader digital forestry systems.

However, Digital Twin claims should be treated carefully. Operational Digital Twins for forest-harvester decision support are not yet mature. More realistic near-term approaches include forecasting-oriented smart management, growth-model emulation, and scenario analysis using updated tree or stand descriptors [78–84]. Deep learning approaches to forest growth retrieval, LiDAR-optical structure monitoring, biomass estimation and harvest-practice identification demonstrate how sensing data may support longer-term planning beyond immediate harvesting operations [79–84].

Accordingly, this review frames forecasting as a digital-forestry-ready extension rather than as a fully implemented Digital Twin layer. The most realistic pathway is to aggregate tree-level perception outputs into stand-state descriptors, update those descriptors after thinning actions, and use growth or scenario models to estimate future harvest potential over a 5–10-year horizon. This approach connects selective harvesting with smart forest management while avoiding overstatement of the current maturity of Digital Twin implementations.

3.9. *Synthesis of the State of the Art*

The reviewed evidence highlights the following integration-relevant findings.

Tree-level perception has advanced substantially, but most computer-vision studies remain aerial, static, or offline rather than harvester-mounted and operational [25–38]. As a result, cross-perspective domain shift remains a key barrier for transferring models from aerial or UAV images to moving ground-level harvester perception.

Next, LiDAR and mobile point-cloud processing provide the strongest basis for spatial localization, stem mapping, and structural-attribute extraction, especially from mobile or machine-mounted platforms [13–15,39–54]. However, the reviewed literature does not yet provide a single controlled benchmark that quantifies the accuracy–latency–robustness trade-off across geometric trunk extraction, deep learning segmentation, and hybrid workflows under identical harvesting conditions.

Multimodal fusion is well justified because selective harvesting requires both semantic interpretation and metric geometry [10,55–62,65]. The most realistic near-term direction is interpretable late fusion, in which image detections are associated with LiDAR-derived tree objects, while deeper learned fusion remains a future research path.

Decision-support and optimization methods are mature enough to support transparent tree-selection logic, but they remain insufficiently connected to real-time sensing streams [16–20,68–74].

Finally, forecasting and Digital Twin-oriented approaches provide an important long-term extension, but the current state of the field supports a staged, forecasting-oriented interpretation rather than claims of fully mature forest-harvester Digital Twins [75–84].

Together, these findings inform the next section of the manuscript, which synthesizes the evidence into a review-derived reference architecture for AI-assisted selective harvesting. The architecture should therefore be understood not as a fully validated final system, but as an integrative framework that organizes the most promising sensing, perception, fusion, decision-support, and forecasting approaches into a coherent direction for future smart forest-management systems.

4. Integrative Synthesis: Reference Architecture for AI-Assisted Selective Harvesting

The previous section showed that the relevant literature has not yet converged into one fully validated harvester-mounted system. Instead, it provides several mature or emerging building blocks: smart-forestry sensing, RGB and LiDAR-based tree perception, ground-level and machine-mounted mapping, decision-support methods for thinning and tree selection, and forecasting-oriented digital-forestry concepts. The purpose of this section is, therefore, not to claim a completed technological solution, but to synthesize these building blocks into a review-derived reference architecture for future AI-assisted selective harvesting systems. This synthesis follows the evidence base identified in the PRISMA-informed review and is intended as a conceptual bridge between current research and future experimental implementation [1–8,13–18,63–76].

The proposed reference architecture is framed as an operator-support and forest-management system rather than as a fully autonomous harvesting controller. Recent work on Forestry 5.0 and human-centered artificial intelligence emphasizes that digital systems should support expert decision-making, transparency, and human oversight in complex forest environments [5]. Similarly, operational studies on thinning assistance and onboard LiDAR indicate similar trends: the most realistic near-term role of AI is to improve situational awareness, reduce decision uncertainty, and support the consistent application of silvicultural prescriptions [13–16].

4.1. Design Principles Derived from the Review

The reference architecture is based on six design principles derived from the reviewed literature. The first design requirement is to operate at the tree level. Selective harvesting and thinning decisions are not made only at stand scale; they require identifying individual trees, estimating their local context, evaluating them against retention or removal criteria and supporting future stand-development estimates. This requirement is supported by recent studies, including individual-tree optimization, the use of LiDAR-derived thinning-tree selection, and implementations of automated thinning frameworks [16–20,70,71]. For the proposed architecture, tree-level operation also provides the link between immediate cut/keep recommendations and future stand-development estimates.

Next, the system should combine semantic and geometric perception. RGB imaging and computer vision provide visual and species-related cues, while LiDAR and point-cloud processing provide distance, position, and structure. Reviews of smart forestry and LiDAR data fusion show that robust digital-forestry workflows increasingly depend on multimodal sensing rather than a single sensor channel [2,8,9,55–62,65]. For this reason, RGB–LiDAR fusion is treated as a core architectural requirement, not as an optional enhancement.

A further requirement is that the architecture should remain modular. The reviewed studies differ in sensing modality, acquisition perspective, AI model family, and deployment maturity. A modular architecture allows major components to be improved independently, such as RGB detection, LiDAR clustering, fusion, decision support and forecasting. At the same time, all mentioned components are still contributing to a common operational workflow. This is especially important because AI perception models evolve quickly, whereas forestry decision rules and operational constraints may remain more stable [4,25–33,63].

A further design principle is that the system should preserve human-in-the-loop control. Forest operations involve multiple different considerations, including economic, ecological, safety and regulatory, that cannot be fully reduced to a detection score. Therefore, AI outputs should be presented as recommendations, confidence cues, or decision aids for operators and managers, but not as unqualified automatic commands. This is consistent with human-centered Forestry 5.0 concepts and recent work on harvester operator assistance [5,13,64].

Another design principle is deployability under field constraints. Machine-mounted systems must tolerate field conditions typical of forestry operations: machinery vibrations, changing illumination during the day or caused by different forestry density, partial occlusion, dust, variable weather and seasonal conditions, and limited connectivity. The reviewed literature related to edge AI and mobile forest robotics indicates that real-time or near-real-time processing must be handled locally whenever possible, with offline training and model updates performed separately [14,15,47,48,53,63,67].

Finally, the architecture should connect current actions to future stand development. Selective harvesting affects important forestry indicators and features such as future growth, structural diversity and harvest potential, which are important for smart planning. Digital Twin and forecasting studies suggest that tree-level observations and harvesting actions can gradually feed into digital-forestry-ready models, although full Digital Twin implementation remains an emerging direction rather than a mature standard [6,75–83].

4.2. Review-Derived Reference Architecture

Based on these principles, the proposed architecture can be organized into the following six functional layers: sensing, perception, localization and fusion, decision support, forecasting and deployment. These layers are not independent research topics, but they are connected components of a sensing-to-decision pipeline. The architecture begins with multimodal data acquisition from RGB cameras and LiDAR sensors, proceeds through AI-based tree detection and point-cloud processing, combines these outputs into a local tree map and then converts the tree-level representation into operator-facing cut/keep recommendations and longer-term management information.

In this architecture (Figure 2), the forest harvester acts as a mobile observation and decision-support node. It collects local information from the operational environment, constructs a machine-local representation of visible and measurable trees, and eventually assists the operator in applying thinning or selective-harvesting objectives. This interpretation is supported by studies on harvester-mounted mobile laser scanning, onboard LiDAR applications in cut-to-length harvesting, and thinning-density assistance systems [13–15]. At the same time, the architecture remains connected to broader smart-forestry and forest-management goals by allowing data export, model updating, and stand-level forecasting [2,7,75,76].

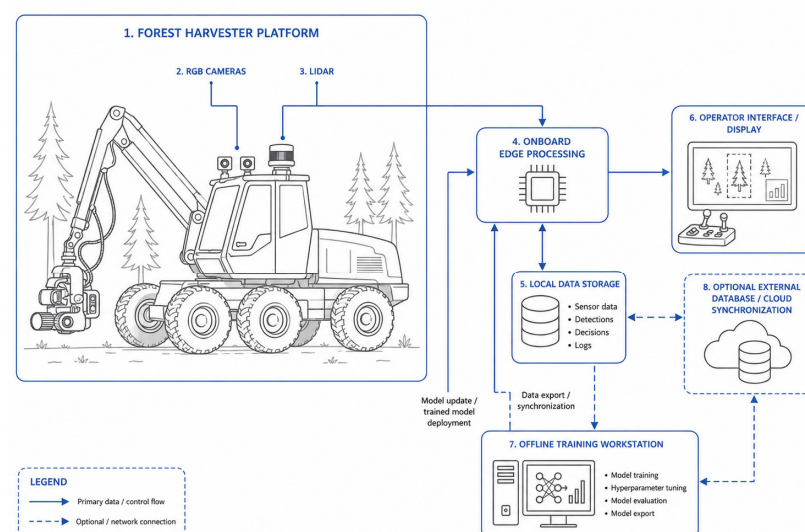


Figure 2. Overall reference architecture of the AI-assisted selective harvesting system, showing the forest-harvester platform, RGB and LiDAR sensing, onboard edge processing, local storage, operator interface, offline training workstation, and optional external database or cloud synchronization.

The proposed architecture layers were derived from the review results and are detailed in Table 6.

Table 6. Review-derived architecture layers and their evidence base.

Architecture Layer	Main Function	Representative Evidence Base	Implication for the Proposed System
Sensing layer	Acquire visual and geometric information from the local forest environment.	Smart forestry and sensor integration studies; LiDAR fusion and point-cloud reviews [1–8,10–12].	Use RGB imaging and 3D LiDAR as complementary sources rather than competing alternatives.
AI perception layer	Detect and characterize visible tree instances from image data.	Deep learning detection, segmentation, and lightweight trunk-detection studies [25–38,63].	Begin with real-time detection, with segmentation reserved for dense or ambiguous scenes.
Localization and fusion layer	Associate detected trees with LiDAR-derived spatial objects and construct a local tree map.	Harvester-mounted MLS, machine localization, RGB–LiDAR fusion, and point-cloud studies [13–15,39–62,65].	Prefer interpretable late fusion as the first deployable strategy, with deeper fusion as a future upgrade.
Decision-support layer	Convert tree-level observations into cut/keep recommendations under silvicultural and operational constraints.	Thinning-tree selection, individual-tree optimization, DSS, and multi-criteria decision studies [16–20,68–74].	Use transparent constraint-aware ranking or optimization rather than opaque automatic decisions.
Forecasting layer	Aggregate current tree and stand states into future harvest or stand-development estimates.	Growth, biomass, harvesting-regime, and Digital Twin-oriented studies [6,75–83].	Treat forecasting as a staged digital-forestry extension, not as a fully mature harvester Digital Twin.
Deployment layer	Separate field inference from offline training, storage, and model updating.	Edge AI, mobile robotics, harvester assistance, and localization studies [13–15,47,48,53,63,64,67].	Run operational inference onboard; perform training, benchmarking, and major updates offline.

The architecture presented in Table 6 reflects the synthesis of the reviewed literature rather than a fixed implementation. Its purpose is to identify the minimum functional chain required to move from sensing to selective-harvesting support. Depending on budgets, site conditions, and management objectives, some components may be simplified. For example, a reduced-cost prototype could use fewer cameras or a simpler LiDAR configuration, whereas a research-grade system should prioritize 3D LiDAR, synchronized multimodal data, and a structured data-management loop.

4.3. Suggested Multisensor Configuration

The sensing configuration should reflect the operational geometry of a forest harvester. The reviewed harvester-mounted and mobile mapping studies indicate that sensors must be positioned to observe trees before the harvesting head reaches them, while minimizing occlusion from the boom, cabin, and nearby understory [13–15,47,53]. For this reason, a cabin-roof configuration is a reasonable reference option: RGB cameras can provide forward and lateral visual coverage, while a 3D LiDAR can capture the surrounding local structure and support machine-local mapping.

The review supports 3D LiDAR as the preferred geometric sensing option for a full research system. Two-dimensional LiDAR may remain useful for low-cost obstacle awareness or simplified stem detection, but it does not preserve enough vertical structure for

robust tree separation, height-related descriptors or local stand reconstruction. By contrast, terrestrial and mobile point-cloud studies show that 3D LiDAR can support tree segmentation, stem extraction, structural parameter estimation, machine localization and provide a wider range of information for smart-forestry systems [9,39–54].

RGB cameras remain essential because LiDAR geometry alone does not provide the same semantic information as visual imagery. RGB and multispectral image-based studies demonstrate the usefulness of visual data for tree detection, crown delineation, trunk recognition, and species-related classification [25–38]. For the proposed architecture, RGB cameras serve as the main semantic channel, and LiDAR provides the spatial channel. This division also makes the first fusion layer easier to validate: visual detections can be associated with LiDAR-derived clusters or local tree objects rather than relying immediately on a fully learned multimodal network.

The proposed harvester-oriented sensor configuration is illustrated in Figure 3.

The 25 m observation zone should be treated as an illustrative design target rather than a universal parameter. Its practical value depends on sensor range, forest density, machine speed, terrain, and the degree to which recommendations must be generated before the operator reaches a candidate tree. In dense stands, useful perception may be limited by occlusion before sensor range is reached. In more open stands, a larger local map may be feasible. The architecture therefore defines the observation radius as a configurable operational parameter rather than a fixed biological or silvicultural threshold.

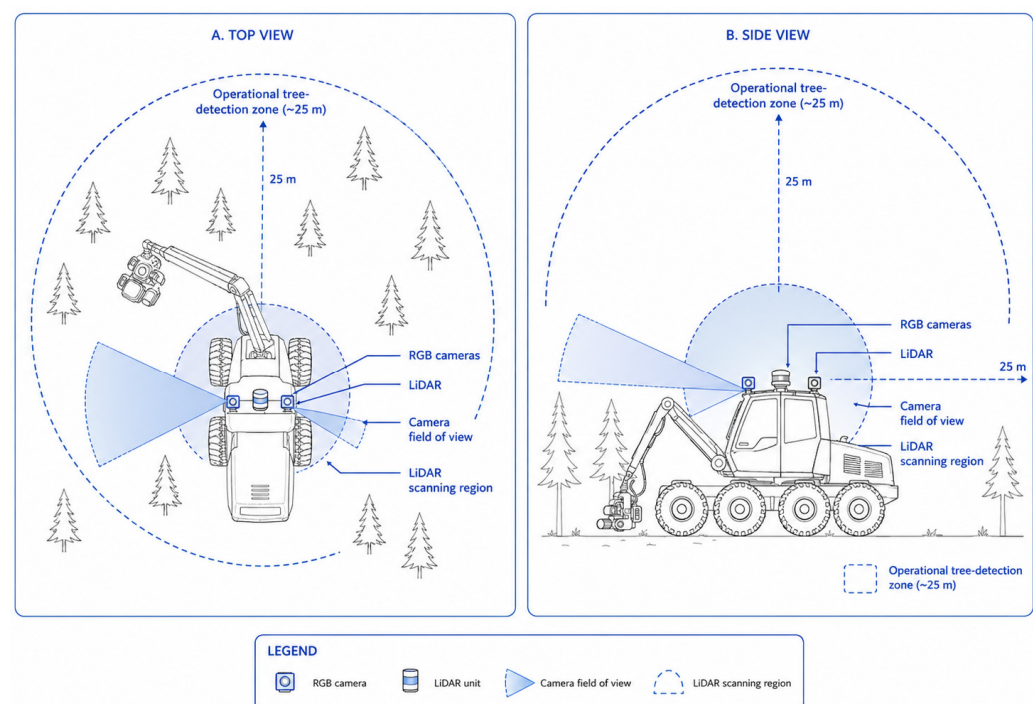


Figure 3. Illustrative harvester-oriented sensor configuration showing top and side views of the harvester, cabin-roof RGB cameras and LiDAR, camera fields of view, LiDAR scanning region, and an operational tree-detection zone of up to approximately 25 m.

4.4. Decision-Support Workflow and Operator-Facing Recommendations

The decision-support workflow is the central forest-management component of the proposed architecture. It converts sensor-derived tree entities into recommendations that correspond to a silvicultural task, including selective thinning, removal of competing stems, retention of young or protected trees, and maintenance of a target residual density. The reviewed DSS and optimization literature shows that such decisions should be repre-

sented as multi-criteria and constraint-aware rather than as a simple detection/classification problem [16–20,68–74].

In the proposed architecture, decision support has two closely connected functions. The first function is analytical: detected and localized trees must be filtered, ranked, and evaluated against the current thinning prescription or management objective. The second function is operational: the result of this analysis must be presented to the operator in a form that supports field decisions without replacing professional judgment. For this reason, the workflow is divided into the decision-support logic itself and the operator-facing interface requirements.

4.4.1. Decision-Support Logic for Cut/Keep Recommendations

The workflow begins with a detected and localized tree (Figure 4). From this tree, the system estimates or records relevant features such as position, distance from the harvester, structural descriptors, species or class label, confidence score, and neighborhood context. Depending on available sensors and models, these features may include either direct LiDAR measurements or image-based descriptors, or may include derived proxies. The tree is then evaluated against preservation constraints, with minimum diameter, minimum height, minimum age proxy, habitat-tree designation, seed-tree status, protection buffers, and machine-accessibility limitations being the most important constraints.

After exclusion rules have been applied, the remaining eligible trees can be ranked according to a utility or priority score. The score may reflect different parameters, among which can be distinguished thinning objectives, competition reduction, economic value, spatial distribution, accessibility, and broader management goals. Finally, the ranking is constrained by the target removal level, for example, a prescribed thinning percentage. This structure is consistent with recent tree-level optimization approaches and LiDAR-informed thinning-tree selection systems [16–20,70,71].

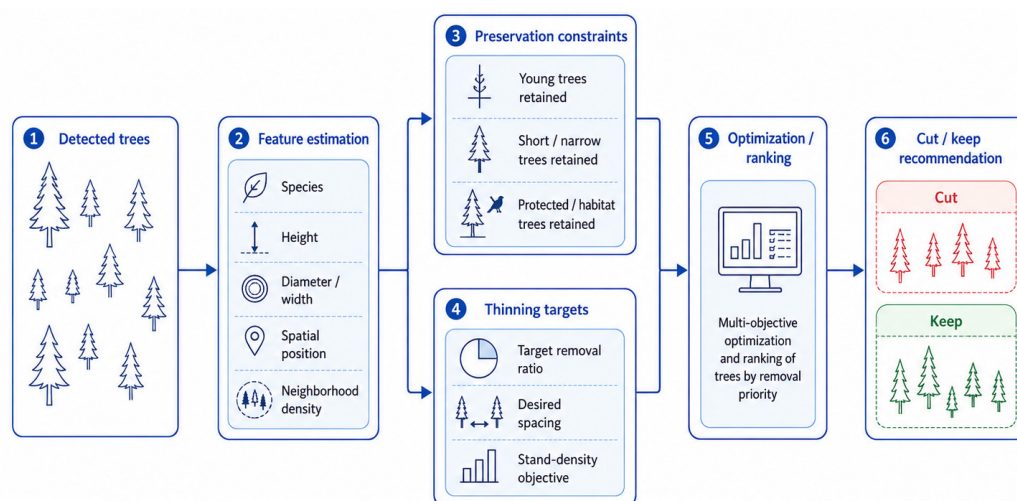


Figure 4. Decision-support workflow for selective harvesting. The workflow links detected trees, feature estimation, preservation constraints, thinning targets, optimization or ranking, and final cut/keep recommendations.

A key advantage of this workflow is interpretability. The operator can see why a tree was excluded, why another tree was ranked as a candidate for removal, and how the recommendation relates to the thinning objective. This is important because forestry decisions involve not only data but also professional judgment and site-specific priorities. In this sense, the proposed system follows the logic of human-centered Forestry 5.0: it assists,

structures, and documents decisions, while at the same time keeping the final responsibility with the human operator or forest manager [5,13,64].

4.4.2. Operator Interface and Human-in-the-Loop Decision Cues

For the proposed architecture, the operator interface should be designed as a decision-support layer rather than as an automated command interface. This distinction is important because selective harvesting requires simultaneous consideration of tree attributes, machine accessibility, residual stand structure, safety, and silvicultural rules. Human-centered Forestry 5.0 concepts and recent harvester-assistance studies support the view that AI should improve operator awareness and decision consistency while preserving expert control [5,13,64].

To reduce cognitive load, the interface should avoid displaying every low-confidence detection or every intermediate model output. Instead, it should prioritize a small number of operationally relevant tree candidates and present them with clear decision cues. These cues should include the recommendation type, confidence level, and the reason for the recommendation. For example, the system may present a tree as “candidate for removal: high local competition”, “retain: below diameter threshold”, “retain: protected or habitat tree”, or “uncertain: operator confirmation required”. This design allows the operator to understand the logic behind the recommendation rather than responding to an unexplained binary output.

The interface should also help prevent overreliance on automation. Recommendations should therefore be expressed as advisory categories rather than automatic commands. Uncertain cases should be visually distinguished from high-confidence recommendations, and the operator should be able to confirm, reject, or override a recommendation. Such overrides should be stored, together with the local tree map and decision context, because they can later support model evaluation, dataset improvement, and refinement of decision-support rules.

Table 7 summarizes the proposed operator-interface requirements derived from the reviewed evidence and the human-in-the-loop architecture.

Table 7. Operator-interface requirements for human-in-the-loop selective-harvesting support.

Interface Requirement	Practical Decision Cue	Purpose in Operator Support	Evidence Base
Prioritized candidate display	Show only operationally relevant tree candidates rather than all detections.	Reduces visual clutter and limits unnecessary cognitive load.	[5,13,64]
Explainable cut/keep reason	Provide short reason labels such as “retain: below threshold” or “candidate: high competition”.	Helps the operator understand why a recommendation was produced.	[5,16–20,68–74]
Confidence and uncertainty indication	Mark uncertain detections or weak RGB–LiDAR associations.	Prevents low-confidence outputs from being interpreted as commands.	[13–15,63,64]
Spatial anchoring	Link recommendations to the visible tree position or local tree-map object.	Supports correspondence between the display and the real forest scene.	[13–15,47,53,64]
Operator confirmation and override	Allow confirm, reject, or override actions.	Maintains human responsibility and reduces automation overreliance.	[5,13,64]
Logging of decisions and overrides	Store recommendations, operator actions, and context metadata.	Supports later model evaluation, training-data improvement, and auditability.	[63–67,75–84]

4.5. Forecasting-Oriented Extension

The reviewed literature suggests that selective-harvesting assistance should eventually extend beyond the immediate cut/keep recommendation. A harvesting action changes the residual stand, modifies competition patterns, and influences such important parameters such as future growth, biomass, resilience, and harvesting potential. Forecasting-oriented smart forestry therefore represents a natural extension of the proposed architecture, although it should be introduced cautiously because mature, real-time forest Digital Twin systems are still limited [6,75–83].

In the reference architecture, the forecasting layer receives aggregated outputs from the local tree map and the decision-support workflow. These outputs may include such information as the current detected tree distribution, retained and removed tree sets, diameter or height distributions, tree-density indicators, and different types of metadata describing the applied harvesting prescription. The forecasting layer then updates a simplified stand-state representation and estimates future outcomes over a medium-term horizon, such as five to ten years. Possible outputs include major parameters such as expected harvestable volume, expected number of harvestable trees, stand-density indicators or scenario comparisons under different thinning intensities.

The literature supports several possible modeling approaches. Classical growth and yield models remain useful where long-term empirical relationships are available. Machine-learning approaches can emulate or approximate forest growth dynamics when suitable historical or simulation data are available [78,79]. Remote-sensing time series and LiDAR-optical synergies can support broader monitoring of forest structure and harvesting practices [80–82]. Digital Twin-oriented studies offer a longer-term vision in which physical forest stands, data streams, simulations, and interactive decision environments are linked within a continuously updated system [75–77,83,84].

Based on this analysis, a proposed forecasting and digital-forestry-oriented extension workflow is depicted in Figure 5.

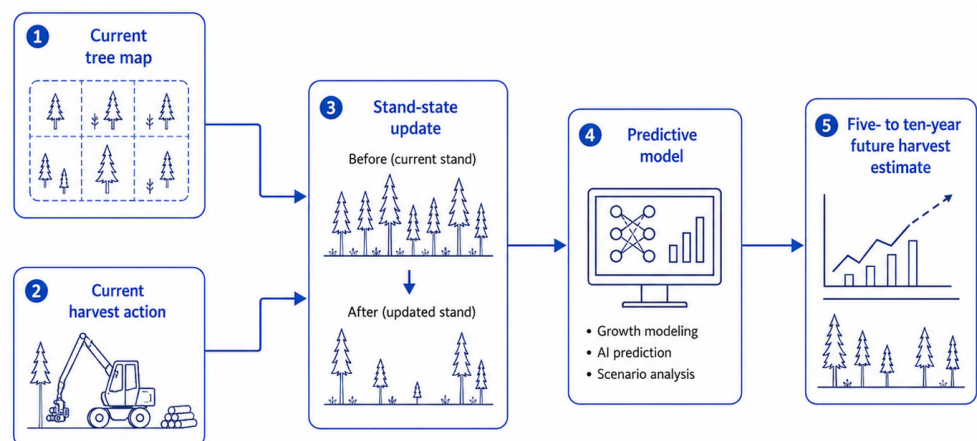


Figure 5. Forecasting- and digital-forestry-oriented extension. The subsystem links the current tree map, current harvest action, stand-state update, predictive model, and five- to ten-year future harvest estimate.

For the present review, the most defensible interpretation is that future harvester systems should be digital-forestry-ready rather than presented as complete Digital Twins from the outset. This means that data structures, local tree maps, and harvesting-action records should be designed so that they can later feed forecasting models and stand-level planning tools. Such a staged approach avoids overclaiming technological maturity while still connecting selective harvesting to long-term smart forest management.

4.6. Deployment Architecture

The deployment architecture (Figure 6) follows a split between onboard field processing and offline model development. Onboard processing is required because forest operations often occur in environments with limited or unreliable connectivity, and operator-support outputs must be produced with low latency. The edge subsystem should perform sensor acquisition, AI inference, LiDAR processing, RGB–LiDAR association, local tree-map generation, decision-support computation, and operator-interface rendering directly on the harvester [13–15,63,67].

Offline processing has a different purpose. It should support dataset curation, annotation, model training, benchmarking, hyperparameter tuning, validation, and model packaging. The offline workstation or data-center subsystem can also aggregate field data from multiple machines, store model versions, and support longer-term forecasting or scenario analysis. This separation is consistent with edge-AI and forestry-robotics studies, where deployment performance must be evaluated separately from model-training capacity [47,48,63,66,67].

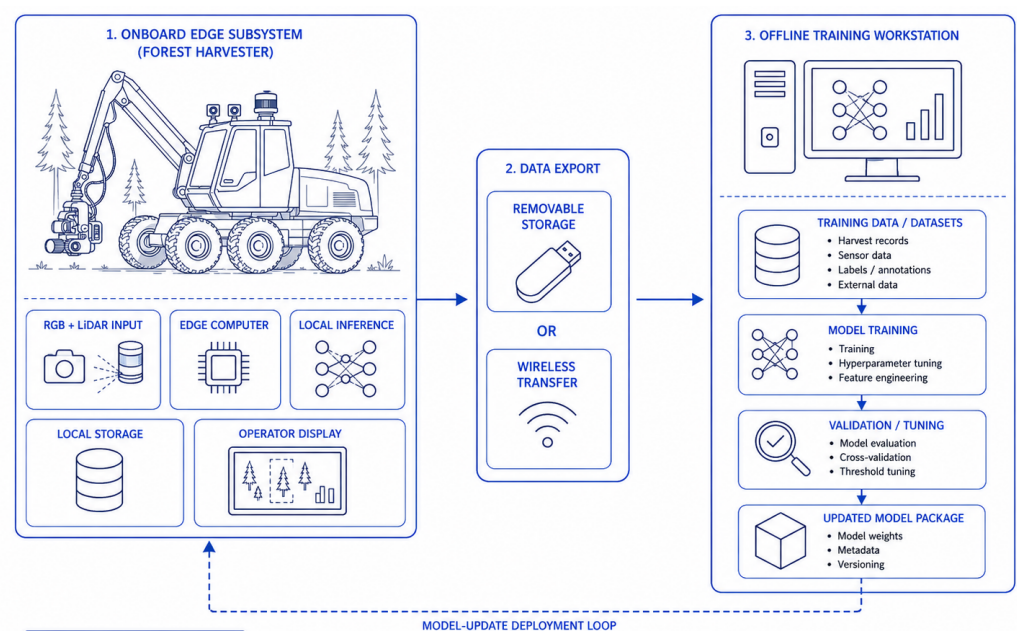


Figure 6. Deployment architecture showing the onboard edge subsystem on the forest harvester, local inference and storage, removable or wireless data export, offline training workstation, and model-update deployment loop.

The resulting deployment loop has two directions. The first direction transfers field data from the harvester to the offline environment for annotation, analysis, model improvement, and data archiving. The second direction returns validated model updates, calibration settings, and decision-support configurations to the harvester to be used in a field scenario. In operational practice, this loop may be implemented through different media. Possible implementation options include removable storage, wired transfer, local wireless networks, and cloud synchronization where connectivity and data-governance requirements allow.

This incremental pathway is also important from a validation perspective. The first prototype should not attempt to solve tree perception, multimodal fusion, decision support, forecasting, and full Digital Twin integration simultaneously. A more defensible experimental sequence is to validate local RGB–LiDAR perception and tree-map generation first, then test operator-facing cut/keep recommendations, and only then connect recorded actions to forecasting or stand-state update models. This staged pathway reflects the ma-

turity differences identified in the reviewed literature and avoids overstating the readiness of fully autonomous or fully integrated Digital Twin-based harvesting systems [1–8,13–20,63–84].

5. Discussion: Research Gaps, Practical Implications, and Future Experimental Agenda

5.1. Main Research Gaps

The synthesis in Sections 3 and 4 shows that the field has advanced rapidly but unevenly. Smart forestry, Forest 4.0 concepts, precision forestry, and IoT-enabled sensing have created a broad digitalization context for forest operations [1–8,10–12]. At the same time, the literature still tends to develop sensing, AI perception, forest-machine assistance, optimization, and forecasting as partially separate research streams rather than as components of a single operational workflow. For AI-assisted selective harvesting, this separation is a central gap because the practical value of the system depends on whether tree-level observations can be transformed into transparent and timely management recommendations.

A related gap concerns feature adaptation for lightweight models. Existing visual models provide important baselines, but many of them are trained and evaluated on aerial, unmanned aerial vehicle, satellite, or static image data [25–38]. Harvester-mounted perception requires a different feature emphasis: stems, lower crowns, partial tree visibility, vertical structures, distance variation, and occlusion patterns become more important than crown-only appearance. Therefore, future datasets should not only contain more field images; they should also support model adaptation toward logging-specific features and should be evaluated under machine-motion, vibration, shadow, and understory conditions [33,34,63–67]. Harvester-mounted LiDAR and mobile laser scanning studies directly address this setting, but they remain much fewer than general forest-inventory or aerial mapping studies [13–16,47,53].

Another gap lies in the transition from accurate sensing to reliable multimodal fusion. RGB-based models provide semantic information, and LiDAR or point-cloud workflows provide geometry, stem positions, and structural attributes [25–55]. Several recent studies demonstrate the value of combining LiDAR, RGB, hyperspectral, or multi-source data for individual-tree mapping and classification [56–63,65]. However, the literature still offers limited evidence on real-time RGB–LiDAR fusion under actual harvesting-machine constraints, where calibration stability, synchronization, latency, and operator interpretability are as important as offline accuracy.

Another unresolved issue is the absence of a controlled benchmark for the trade-off between geometric point-cloud extraction, deep learning-based 3D segmentation, and hybrid workflows. The reviewed point-cloud literature includes geometric, mobile-mapping, and learning-based methods [39–54], but the studies differ substantially in scanner type, point density, forest structure, evaluation metrics, and deployment setting. As a result, this review cannot report a universal numerical comparison of accuracy, latency, and field robustness. This should be treated as a research gap rather than as a weakness of a particular method. For selective harvesting, future studies should compare these approaches under the same harvester-oriented dataset and hardware conditions.

A further gap concerns the connection between perception and silvicultural decision-making. Decision-support systems, multi-criteria methods, and optimization approaches have already been well established in forestry [68–74]. Recent studies on thinning-tree selection and tree-level optimization show that individual-tree decisions can be formalized algorithmically [16–21,70–74]. Yet these approaches often rely on inventory or planning data rather than continuous real-time observations from a working harvester. Con-

versely, perception-oriented papers often stop at detection, localization, or attribute extraction without specifying how the result should alter a thinning or selective-harvesting decision.

The final gap concerns the still immature link between harvesting actions and future stand development. Digital Twin and forecasting-oriented forestry studies indicate that data-driven and simulation-based tools can connect current observations with future management outcomes [76–84]. Nevertheless, full Digital Twin integration remains rare in operational harvesting contexts. For this reason, the near-term research target should not be an overstated fully autonomous forest Digital Twin, but a digital-forestry-ready forecasting layer that can gradually connect tree-level harvesting decisions with future stand-state and harvest-potential estimates.

5.2. Practical Implications for Forest Operations

For forest operations, the most immediate implication of this review is that AI-assisted selective harvesting should be framed as an operator-support technology rather than as a replacement for professional judgment. This framing is consistent with human-centered AI discussions in Forestry 5.0 and with studies of harvester-oriented assistance systems [4,13–15,64]. The operator remains responsible for interpreting stand conditions, machine limitations, and local silvicultural prescriptions, while the AI system can help maintain awareness of local tree density, candidate removals, tree attributes, and compliance with predefined constraints.

Selective harvesting also requires more than accurate object detection. A technically successful detector may still be operationally weak if it does not produce information that supports decisions such as which tree should be retained, which tree is eligible for removal, how many trees can be removed in the local neighborhood, and whether the resulting stand structure remains consistent with management objectives. Studies on thinning-density assistance, onboard LiDAR, tree-selection optimization, and tree-level planning show that this decision layer is where sensing becomes relevant to actual harvesting practice [13–20].

Another important implication concerns sustainability and resilience. Thinning and selective harvesting influence competition, stand structure, habitat retention, and long-term productivity. Evidence on thinning effects and broader harvesting regimes indicates that harvesting decisions must be evaluated not only in terms of immediate productivity but also in terms of resilience, disturbance risk, and stand-development trajectory [23,24]. Accordingly, an AI-assisted harvester should not simply identify harvestable stems; it should support a controlled decision process where preservation rules, ecological constraints, and operator review are explicit.

Organizational implications also should be addressed. A harvester-mounted sensing system can become a mobile source of forest data if the workflow records local tree maps, operator decisions, and post-harvest stand information. Such data could strengthen inventory updating, training-data generation, and future decision-support models, especially if combined with mobile LiDAR, tree-map matching, and standardized point-cloud processing pipelines [15,39,49,52,53]. However, the value of this data stream depends on technical factors such as calibration quality, consistent metadata, data governance, and the ability to integrate field data with forest-management systems, which should be addressed carefully.

The operator-interface implications are also important. A useful system should not overload the operator with all intermediate detections or low-confidence model outputs. Instead, the display should prioritize operationally relevant tree candidates, show confidence or uncertainty, provide short reasons for each cut/keep suggestion, and allow the operator to confirm, reject, or override recommendations. This approach is consistent with

the human-in-the-loop framing of the reference architecture and reduces the risk that the system will be interpreted as an automatic command generator rather than as a decision-support tool [5,13–15,64].

Finally, the findings suggest that transparency is not optional. Forestry decisions involve ecological, economic, and operational trade-offs, and operators are unlikely to trust recommendations that appear as “black-box” outputs. Constraint-aware ranking, interpretable thinning rules, confidence-aware visualization, and clear feedback to the operator are, therefore, more suitable approaches for near-term deployment than fully opaque recommendation models [68–75].

5.3. Future Experimental Agenda

The next stage of research should begin with the creation of a harvester-perspective multimodal dataset. Such a dataset should contain synchronized RGB imagery, LiDAR or point-cloud observations, machine-state metadata, calibration parameters, tree-instance annotations, and, where possible, expert-defined cut/keep or thinning-priority labels. Existing tree-detection and LiDAR datasets provide important methodological precedents [25–45], but they do not fully represent the viewpoint, motion, occlusion, and decision context of a working forest harvester.

After dataset creation, model benchmarking should compare one-stage object detectors, lightweight trunk-detection models, convolutional instance-segmentation workflows, and emerging transformer-based approaches. These models should be evaluated not only by detection accuracy, but also by robustness under occlusion, lighting variation, stand density, and viewpoint changes [25–38]. For this application, model selection should use combined criteria including detection quality, suitability for RGB–LiDAR association, inference speed, and compatibility with edge deployment.

In parallel, LiDAR and point-cloud workflows should be benchmarked under the same operational assumptions. Geometric stem extraction, clustering with model fitting, deep learning-based 3D instance segmentation, mobile mapping pipelines, and hybrid geometric–visual approaches should be compared using common test data. The evaluation should include stem-localization error, instance association accuracy, end-to-end latency, dropped-frame rate, robustness under occlusion and understory clutter, and sensitivity to calibration drift. This would directly address the current lack of comparable accuracy–latency–robustness evidence for selective-harvesting conditions [39–54].

Once perception models have been benchmarked, fusion strategies should be compared experimentally. Late fusion should be treated as the initial baseline because it is interpretable and compatible with limited early-stage datasets. The system can then progress toward learned or feature-level fusion after sufficient aligned multimodal data become available. Evaluation should include not only classification or detection accuracy, but also tree-association accuracy, coordinate error, robustness under calibration drift, and the stability of tree identifiers across time [10–12,39–63].

The decision-support layer should then be validated. This requires comparing AI-generated recommendations with expert thinning decisions and with explicit management prescriptions. Experiments should evaluate whether the system respects target thinning ratios, avoids protected or undersized trees, maintains spatial distribution requirements, and reduces operator workload without oversimplifying silvicultural judgment [13–20,68–74]. Such validation should include both simulated scenarios and field trials, because offline agreement with expert labels does not necessarily guarantee usability during real harvesting.

Edge deployment and human–machine interaction should also be tested. The system should be evaluated on onboard computing hardware under realistic latency, vibration, lighting, temperature, and connectivity conditions. Metrics should include inference time,

end-to-end cycle time, dropped-frame rate, memory use, thermal stability, and operator comprehension of the displayed recommendations [63–67]. Human-in-the-loop experiments should also evaluate whether confidence cues, explanation labels, candidate prioritization, and override options reduce cognitive load and prevent overreliance on automation. Such tests are especially important because the interface may determine whether the technology is perceived as useful support or distracting overload [5,13–15,64].

A final step is forecasting-oriented integration. Once reliable tree-level observations and harvesting decisions are available, these outputs can feed growth-model surrogates, stand-state updates, scenario analysis, or Digital Twin-inspired planning modules [76–84]. At this stage, the key research question changes from “Can the system identify and rank trees?” to “Can the system help evaluate how today’s selective harvesting decision affects future stand condition and harvest potential?”.

The research gaps, operational implications and future experimental priorities, along with the references to the key supporting literature, are summarized in Table 8.

Table 8. Research gaps, operational implications, and future experimental priorities for AI-assisted selective harvesting.

Research Gap	Operational Implication	Future Experimental Priority	Key Supporting Literature
Machine-perspective data scarcity	Models trained on aerial or static data may not generalize to harvester-cabin views.	Create synchronized RGB–LiDAR datasets from moving harvesters with expert cut/keep labels.	[13–16,25–38,47,53]
Cross-perspective domain shift and feature adaptation	Aerial crown-oriented features may not capture stems, lower crowns, partial visibility, and machine-perspective occlusion.	Adapt lightweight models using ground-level data, occlusion-aware augmentation, multi-scale features, and LiDAR-guided constraints.	[25–38,63–67]
Limited real-time multi-modal fusion evidence	Tree recognition and localization may remain disconnected or unstable in dense stands.	Compare late, feature-level, and learned fusion using association accuracy, coordinate error, and latency.	[10–12,39–63,65]
Insufficient quantitative trade-off evidence for point-cloud workflows	It remains unclear which combination of geometric preprocessing, 3D segmentation, and hybrid perception is best under field constraints.	Benchmark geometric extraction, clustering, deep 3D segmentation, mobile mapping, and hybrid workflows on the same harvester-oriented dataset.	[39–54]
Weak perception-to-decision coupling	Accurate detection does not automatically produce silviculturally valid recommendations.	Validate cut/keep outputs against expert prescriptions, thinning ratios, and retention constraints.	[16–20,68–74]
Field deployment and operator interface uncertainty	A technically accurate model may fail if latency, visualization, or trust are inadequate.	Test edge hardware, operator displays, confidence cues, and workload effects in field trials.	[5,13–15,63–68]
Immature forecasting integration	Immediate harvesting assistance remains weakly connected to future stand development.	Link tree maps and harvest actions to growth models, scenario analysis, and digital-forestry-ready databases.	[75–84]

5.4. Limitations of the Review

This review has several limitations that should be considered when interpreting the synthesis. First, although the search followed a PRISMA-informed strategy, the topic spans several research communities with different terminologies. Forestry, robotics, remote sensing, computer vision, optimization, and Digital Twin research do not always use the same vocabulary. As a result, some relevant studies may have remained outside

the search strings, particularly if they framed the same problem through more specialized or regional terminology.

Second, the review intentionally emphasized the English-language peer-reviewed literature from the period 2015–2026, supplemented by selected foundational works. This decision made the review manageable and focused on the modern deep learning and smart-forestry period, but it may underrepresent older silvicultural decision-support traditions or non-English regional studies.

Third, the reviewed evidence is heterogeneous: the included papers differ substantially in sensing modality, spatial scale, species composition, evaluation metrics, deployment setting, and maturity level. For that reason, a statistical meta-analysis would not be meaningful at this stage.

Fourth, the reference architecture proposed in this paper is a synthesis outcome, not a validated commercial system. Its value lies in organizing the reviewed evidence into a coherent development pathway, but its performance must still be tested through datasets, prototype implementations, and operational trials.

A further limitation is that the reviewed studies do not provide a directly comparable quantitative basis for ranking all sensing and perception approaches under identical selective-harvesting conditions. Reported metrics vary across datasets, forest types, platforms, sensors, and evaluation protocols. Therefore, the accuracy–latency–robustness comparisons in this review should be interpreted as review-based trade-off synthesis rather than as a universal numerical benchmark.

Finally, because this is a scoping review across heterogeneous research domains, thematic classification and interpretation required author judgment. A structured extraction matrix and predefined thematic categories were used to make the synthesis transparent, but no inter-rater reliability analysis was performed. This should be considered when interpreting the classification of studies across overlapping categories such as multimodal fusion, edge deployment, and decision support.

Despite these limitations, the review provides a practical basis for moving from fragmented research components toward integrated, transparent, and field-oriented AI assistance for selective harvesting and smart forest management.

6. Conclusions

This review examined how recent advances in artificial intelligence, multimodal sensing, and decision-support methods can contribute to AI-assisted selective harvesting and smart forest management. Based on a PRISMA-informed scoping process, 82 studies were retained and synthesized across the main research streams relevant to future harvester-mounted systems, including smart forestry, forest operations, RGB and LiDAR sensing, computer vision, multimodal fusion, edge deployment, decision support, and forecasting-oriented digital forestry.

The main finding is that the field already provides several technically mature components, but they are not yet sufficiently integrated for operational selective harvesting. Computer-vision methods support tree and trunk recognition, while LiDAR and point-cloud processing provide the spatial and structural information needed for local tree mapping. However, much of the existing work remains based on aerial, static, or offline workflows. Selective harvesting requires a different setting: ground-level, machine-mounted perception that can operate under occlusion, vibration, changing illumination, limited connectivity and continuous operator workload.

The review therefore supports RGB–LiDAR sensing as the most realistic basis for near-term harvester assistance. RGB imagery provides semantic information that is easy to interpret, whereas LiDAR contributes metric geometry, local tree position, and structural context. Their combination is especially important when moving from visual detec-

tion toward actionable tree-level assessment. At the same time, the reviewed literature indicates that perception alone is not enough. A useful system must connect detected and localized trees with transparent decision-support logic, including thinning objectives, retention rules, accessibility constraints, operator-interface cues and professional judgement.

The proposed reference architecture should be understood as a review-derived development pathway rather than as a finalized technical solution. It organizes the available evidence into a staged system concept: multimodal sensing, AI-based perception, LiDAR-supported localization, local tree mapping, operator-facing recommendations, and forecasting-oriented data use. This structure reflects the current maturity of the field and avoids overstating the readiness of fully autonomous harvesting or fully operational forest Digital Twins. In the near term, the most defensible approach is a human-in-the-loop system that provides interpretable cut/keep recommendations rather than automatic harvesting commands.

Future work should focus on harvester-perspective RGB–LiDAR datasets, controlled benchmarking of geometric, deep learning-based, and hybrid point-cloud workflows, field validation of real-time fusion methods, edge-compatible perception models, and decision-support modules tested against expert silvicultural prescriptions. Operator-interface studies should also evaluate whether confidence cues, explanation labels, candidate prioritization, and override options reduce cognitive load and prevent overreliance on automation. Forecasting components should be introduced gradually, using recorded tree-level observations and harvesting actions to support stand-state updates and medium-term harvest estimates. In this form, AI-assisted selective harvesting can become a practical human-in-the-loop technology that strengthens operator situational awareness, improves consistency of tree-level decisions, and links harvesting operations more closely with long-term smart forest management.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The author used ChatGPT (GPT-5; OpenAI, San Francisco, CA, USA) during manuscript preparation to assist with minor language editing and translation. All AI-generated content was thoroughly reviewed and revised by the author, who takes full responsibility for the accuracy and integrity of the final work.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Ehrlich-Sommer, F.; Hoenigsberger, F.; Gollob, C.; Nothdurft, A.; Stampfer, K.; Holzinger, A. Sensors for Digital Transformation in Smart Forestry. *Sensors* **2024**, *24*, 798. <https://doi.org/10.3390/s24030798>.
2. Damaševičius, R.; Mozgeris, G.; Kurti, A.; Maskeliūnas, R. Digital transformation of the future of forestry: An exploration of key concepts in the principles behind Forest 4.0. *Front. For. Glob. Change* **2024**, *7*, 1424327. <https://doi.org/10.3389/ffgc.2024.1424327>.
3. Vatandaslar, C.; Boston, K.; Ucar, Z.; Narine, L.L.; Madden, M.; Akay, A.E. Precision Forestry Revisited. *Remote Sens.* **2025**, *17*, 3465. <https://doi.org/10.3390/rs17203465>.
4. Holzinger, A.; Schweier, J.; Gollob, C.; Nothdurft, A.; Hasenauer, H.; Kirisits, T.; Häggström, C.; Visser, R.; Cavalli, R.; Spinelli, R.; et al. From industry 5.0 to forestry 5.0: Bridging the gap with human-centered artificial intelligence. *Curr. For. Rep.* **2024**, *10*, 442–455. <https://doi.org/10.1007/s40725-024-00231-7>.

5. Wang, G.G.; Lu, D.; Gao, T.; Zhang, J.; Sun, Y.; Teng, D.; Yu, F.; Zhu, J. Climate-smart forestry: An AI-enabled sustainable forest management solution for climate change adaptation and mitigation. *J. For. Res.* **2024**, *36*, 7. <https://doi.org/10.1007/s11676-024-01802-x>.
6. Arion, I.D.; Morar, I.M.; Truta, A.M.; Chereches, I.A.; Isarie, V.I.; Arion, F.H. Internet of Things (IoT)-Based Applications in Smart Forestry: A Conceptual and Technological Analysis. *Forests* **2026**, *17*, 44. <https://doi.org/10.3390/f17010044>.
7. Bastos, T.; Teixeira, L.C.; Nunes, L.J. Forest 4.0: Technologies and digitalization to create the residual biomass supply chain of the future. *J. Clean. Prod.* **2024**, *467*, 143041. <https://doi.org/10.1016/j.jclepro.2024.143041>.
8. Subedi, B.; Gautam, S.H.; LeBel, L. Artificial intelligence in the forest products supply chain: Current applications and open challenges. *Forestry* **2026**, *99*, cpag010. <https://doi.org/10.1093/forestry/cpag010>.
9. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. <https://doi.org/10.1136/bmj.n71>.
10. Balestra, M.; Marselis, S.; Sankey, T.T.; Cabo, C.; Liang, X.; Mokroš, M.; Peng, X.; Singh, A.; Stereńczak, K.; Vega, C.; et al. LiDAR data fusion to improve forest attribute estimates: A review. *Curr. For. Rep.* **2024**, *10*, 281–297. <https://doi.org/10.1007/s40725-024-00223-7>.
11. Kulicki, M.; Cabo, C.; Trzciński, T.; Będkowski, J.; Stereńczak, K. Artificial intelligence and terrestrial point clouds for forest monitoring. *Curr. For. Rep.* **2024**, *11*, 5. <https://doi.org/10.1007/s40725-024-00234-4>.
12. Murtiyoso, A.; Cabo, C.; Singh, A.; Obaya, D.P.; Cherlet, W.; Stoddart, J.; Fol, C.R.; Beloiu Schwenke, M.; Rehush, N.; Sterenczak, K.; et al. A review of software solutions to process ground-based point clouds in Forest applications. *Curr. For. Rep.* **2024**, *10*, 401–419. <https://doi.org/10.1007/s40725-024-00228-2>.
13. Pohjala, J.; Kankare, V.; Hyyppä, J.; Kärhä, K. Effects of a mobile LiDAR-based thinning density assistant (TDA) system on harvester operator performance. *Eur. J. For. Res.* **2025**, *144*, 1301–1322. <https://doi.org/10.1007/s10342-025-01808-y>.
14. Sagar, A.; Kärhä, K.; Einola, K.; Koivusalo, A. Assessing the Potential of Onboard LiDAR-Based Application to Detect the Quality of Tree Stems in Cut-to-Length (CTL) Harvesting Operations. *Forests* **2024**, *15*, 818. <https://doi.org/10.3390/f15050818>.
15. Fäitli, T.; Hyyppä, E.; Hyyti, H.; Hakala, T.; Kaartinen, H.; Kukko, A.; Muhojoki, J.; Hyyppä, J. Integration of a Mobile Laser Scanning System with a Forest Harvester for Accurate Localization and Tree Stem Measurements. *Remote Sens.* **2024**, *16*, 3292. <https://doi.org/10.3390/rs16173292>.
16. Lee, Y.; Sim, W.; Lee, S.; Lee, J. Development and Evaluation of a Thinning Tree Selection System Using Optimization Techniques Based on Multi-Platform LiDAR. *Forests* **2025**, *16*, 1776. <https://doi.org/10.3390/f16121776>.
17. Yang, K.; Wu, H.; Shen, Y.; Wang, S.; Lai, R.; Liu, C. GEOIMIA: A new automated forest thinning framework for selective harvesting at the individual tree level in mixed-species forests. *ISPRS J. Photogramm. Remote Sens.* **2026**, *236*, 397–420. <https://doi.org/10.1016/j.isprsjprs.2026.04.006>.
18. Pukkala, T.; Nuutinen, Y.; Muhonen, T. A two-level optimization approach to tree-level planning in continuous cover forest management. *J. For. Res.* **2025**, *36*, 79. <https://doi.org/10.1007/s11676-025-01876-1>.
19. Sforza, F.; Starke, M.; Dietsch, P.; Thür, P.; Lingua, E.; Ziesak, M. Maximizing the economic benefit for cable yarding timber harvesting operations by spatially optimizing tree selection. *Eur. J. For. Res.* **2025**, *144*, 831–850. <https://doi.org/10.1007/s10342-024-01700-1>.
20. Sforza, F.; Ziesak, M.; Lingua, E.; Starke, M. Optimizing tree selection for planning cable yarding operations: A multi-objective modelling approach. *For. Ecol. Manag.* **2025**, *578*, 122489. <https://doi.org/10.1016/j.foreco.2024.122489>.
21. Arnič, D.; Krč, J.; Diaci, J. Modeling of time consumption for selective and situational precommercial thinning in mountain beech forest stands. *iForest—Biogeosci. For.* **2021**, *14*, 137–143. <https://doi.org/10.3832/for3556-014>.
22. Eberhard, B.; Trailovic, Z.; Magagnotti, N.; Spinelli, R. A GIS-Based Decision Support Model (DSM) for Harvesting System Selection on Steep Terrain: Integrating Operational and Silvicultural Criteria. *Forests* **2025**, *16*, 854. <https://doi.org/10.3390/f16050854>.
23. Chagnon, C.; Dumont, S.; Morin-Bernard, A.; Jactel, H.; Achim, A.; Moreau, G. Potential of thinning to increase forest resilience and resistance to drought, pest, windstorm and fire: A meta-analysis. *For. Ecol. Manag.* **2025**, *590*, 122788. <https://doi.org/10.1016/j.foreco.2025.122788>.
24. Suvanto, S.; Esquivel-Muelbert, A.; Schelhaas, M.-J.; Astigarraga, J.; Astrup, R.; Cienciala, E.; Fridman, J.; Henttonen, H.M.; Kunstler, G.; Kändler, G.; et al. Understanding Europe's Forest Harvesting Regimes. *Earth's Future* **2025**, *13*, e2024EF005225. <https://doi.org/10.1029/2024EF005225>.

25. Zhao, H.; Morgenroth, J.; Pearse, G.; Schindler, J. A Systematic Review of Individual Tree Crown Detection and Delineation with Convolutional Neural Networks (CNN). *Curr. For. Rep.* **2023**, *9*, 149–170. <https://doi.org/10.1007/s40725-023-00184-3>.
26. Abreu-Dias, R.; Santos-Gago, J.M.; Martín-Rodríguez, F.; Álvarez-Sabucedo, L.M. Advances in the Automated Identification of Individual Tree Species: A Systematic Review of Drone- and AI-Based Methods in Forest Environments. *Technologies* **2025**, *13*, 187. <https://doi.org/10.3390/technologies13050187>.
27. Abbas, I.; Damaševičius, R. Tree Detection in RGB Satellite Imagery Using YOLO-Based Deep Learning Models. *Comput. Mater. Contin.* **2025**, *85*, 483–502. <https://doi.org/10.32604/cmc.2025.066578>.
28. Sengun, E.; Aksoy, S.; Sertel, E.; Fransson, J.E. Comparative Analysis of YOLOv8 and YOLOv11 on Tree Detection Using UAV RGB and Laser Scanning Data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2025**, *10*, 173–179. <https://doi.org/10.5194/isprs-annals-X-2-W2-2025-173-2025>.
29. Ball, J.G.; Hickman, S.H.; Jackson, T.D.; Koay, X.J.; Hirst, J.; Jay, W.; Archer, M.; Aubry-Kientz, M.; Vincent, G.; Coomes, D.A. Accurate delineation of individual tree crowns in tropical forests from aerial RGB imagery using Mask R-CNN. *Remote Sens. Ecol. Conserv.* **2023**, *9*, 641–655. <https://doi.org/10.1002/rse2.332>.
30. Fu, H.; Zhao, H.; Jiang, J.; Zhang, Y.; Liu, G.; Xiao, W.; Du, S.; Guo, W.; Liu, X. Automatic detection tree crown and height using Mask R-CNN based on unmanned aerial vehicles images for biomass mapping. *For. Ecol. Manag.* **2024**, *555*, 121712. <https://doi.org/10.1016/j.foreco.2024.121712>.
31. Lucas, M.; Pukrop, M.; Beckschäfer, P.; Waske, B. Individual tree detection and crown delineation in the Harz National Park from 2009 to 2022 using Mask R-CNN and aerial imagery. *ISPRS Open J. Photogramm. Remote Sens.* **2024**, *13*, 100071. <https://doi.org/10.1016/j.ophoto.2024.100071>.
32. Yao, S.; Hao, Z.; Post, C.J.; Mikhailova, E.A.; Lin, L. Individual Tree Crown Detection and Classification of Live and Dead Trees Using a Mask Region-Based Convolutional Neural Network (Mask R-CNN). *Forests* **2024**, *15*, 1900. <https://doi.org/10.3390/f15111900>.
33. Zhang, S.; Zheng, Y.; Bi, R.; Chen, Y.; Chen, C.; Tian, X.; Liao, B. Trunk Detection in Complex Forest Environments Using a Lightweight YOLOv11-TrunkLight Algorithm. *Sensors* **2025**, *25*, 6170. <https://doi.org/10.3390/s25196170>.
34. Badgujar, C.; Flippo, D.; Gunturu, S.; Baldwin, C. Tree Trunk Detection of Eastern Red Cedar in Rangeland Environment with Deep Learning Technique. *Croat. J. For. Eng.* **2023**, *44*, 357–368. <https://doi.org/10.5552/crojfe.2023.2012>.
35. Sivanandam, P.; Lucieer, A. Tree Detection and Species Classification in a Mixed Species Forest Using Unoccupied Aircraft System (UAS) RGB and Multispectral Imagery. *Remote Sens.* **2022**, *14*, 4963. <https://doi.org/10.3390/rs14194963>.
36. Bennett, L.; Yu, Z.; Wasowski, R.; Selland, S.; Otway, S.; Boisvert, J. Individual tree detection and classification from RGB satellite imagery with applications to wildfire fuel mapping and exposure assessments. *Int. J. Wildland Fire* **2024**, *33*, WF24008. <https://doi.org/10.1071/WF24008>.
37. Pacheco-Prado, D.; Bravo-López, E.; Martínez, E.; Ruiz, L.Á. Urban Tree Species Identification Based on Crown RGB Point Clouds Using Random Forest and PointNet. *Remote Sens.* **2025**, *17*, 1863. <https://doi.org/10.3390/rs17111863>.
38. Weinstein, B.G.; Marconi, S.; Bohlman, S.; Zare, A.; White, E. Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks. *Remote Sens.* **2019**, *11*, 1309. <https://doi.org/10.3390/rs11111309>.
39. Shao, J.; Lin, Y.C.; Wingren, C.; Shin, S.Y.; Fei, W.; Carpenter, J.; Habib, A.; Fei, S. Large-scale inventory in natural forests with mobile LiDAR point clouds. *Sci. Remote Sens.* **2024**, *10*, 100168. <https://doi.org/10.1016/j.srs.2024.100168>.
40. Ma, S.; Chen, Y.; Li, Z.; Chen, J.; Zhong, X. Improved Cylinder-Based Tree Trunk Detection in LiDAR Point Clouds for Forestry Applications. *Sensors* **2025**, *25*, 714. <https://doi.org/10.3390/s25030714>.
41. Lee, S.; Sim, W.; Lee, Y.; Park, J.; Kang, J.; Lee, J. Comparative Analysis of Machine Learning and Deep Learning Models for Individual Tree Structure Segmentation Using Terrestrial LiDAR Point Cloud Data. *Remote Sens.* **2025**, *17*, 2245. <https://doi.org/10.3390/rs17132245>.
42. Henrich, J.; van Delden, J.; Seidel, D.; Kneib, T.; Ecker, A.S. TreeLearn: A deep learning method for segmenting individual trees from ground-based LiDAR forest point clouds. *Ecol. Inform.* **2024**, *84*, 102888. <https://doi.org/10.1016/j.ecoinf.2024.102888>.
43. Wielgosz, M.; Puliti, S.; Xiang, B.; Schindler, K.; Astrup, R. SegmentAnyTree: A sensor and platform agnostic deep learning model for tree segmentation using laser scanning data. *Remote Sens. Environ.* **2024**, *313*, 114367.
44. Puliti, S.; Pearse, G.; Surový, P.; Wallace, L.; Hollaus, M.; Wielgosz, M.; Astrup, R. For-instance: A UAV laser scanning benchmark dataset for semantic and instance segmentation of individual trees. *arXiv Preprint* **2023**, arXiv:2309.01279. <https://doi.org/10.5281/zenodo.8287792>.

45. Laino, D.; Cabo, C.; Ordóñez, C.; Bolanos, R.; Janvier, R.; Giulioni, F.; Herrmann, M.; Hudak, A.; Parsons, R.; Santin, C. SegmentedForests: A labelled dataset of terrestrial LiDAR point clouds for semantic segmentation of forests. *For. Int. J. For. Res.* **2026**, *99*, cpaf062. <https://doi.org/10.1093/forestry/cpaf062>.
46. Maeda, E.E.; Brede, B.; Calders, K.; Disney, M.; Herold, M.; Lines, E.R.; Nunes, M.H.; Raunonen, P.; Rautiainen, M.; Saarinen, N.; et al. Expanding forest research with terrestrial LiDAR technology. *Nat. Commun.* **2025**, *16*, 8853. <https://doi.org/10.1038/s41467-025-63946-6>.
47. Martín-Cortés, C.; Porkhun, O.; García-Pascual, B.; Zhou, X.; Hormazabal, C.; Acuna, M.; Ruano, A.; Selkimäki, M.; Mola-Yudego, B.; Kymäläinen, H.; et al. 2D-CNN for tree stem detection and segmentation using data collected with a harvester-mounted mobile laser scanner. *Int. J. For. Eng.* **2026**, 1–15. <https://doi.org/10.1080/14942119.2026.2653269>.
48. Freißmuth, L.; Mattamala, M.; Chebrolu, N.; Schaefer, S.; Leutenegger, S.; Fallon, M. Online tree reconstruction and forest inventory on a mobile robotic system. In *Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, United Arab Emirates, 14–18 October 2024; IEEE: New York, NY, USA, 2024; pp. 11765–11772. <https://doi.org/10.1109/IROS58592.2024.10802455>.
49. de Paula Pires, R.; Lindberg, E.; Persson, H.J.; Olofsson, K.; Holmgren, J. Mobile laser scanning as reference for estimation of stem attributes from airborne laser scanning. *Remote Sens. Environ.* **2024**, *315*, 114414. <https://doi.org/10.1016/j.rse.2024.114414>.
50. Zhou, M.; Li, C.; Li, Z. Extraction of individual tree attributes using ultra-high-density point clouds acquired by low-cost UAV-LiDAR in Eucalyptus plantations. *Ann. For. Sci.* **2025**, *82*, 20. <https://doi.org/10.1186/s13595-025-01291-w>.
51. Zhang, Y.; Tan, Y.; Onda, Y.; Hashimoto, A.; Gomi, T.; Chiu, C.; and Inokoshi, S. A tree detection method based on trunk point cloud section in dense plantation forest using drone LiDAR data. *For. Ecosyst.* **2023**, *10*, 100088. <https://doi.org/10.1016/j.fecs.2023.100088>.
52. Liikonen, L.; Yrttimaa, T.; Erkkilä, A.; Paakkari, J.; Pitkänen, T.; Kotivuori, E.; Vastaranta, M. Measuring forest inventory attributes using Faro Orbis Mobile laser scanner in managed boreal forests. *For. Int. J. For. Res.* **2026**, *99*, cpag028. <https://doi.org/10.1093/forestry/cpag028>.
53. Seppälä, W.; Muhojoki, J.; Fäitli, T.; Hyypä, E.; Kaartinen, H.; Kukko, A.; Hyypä, J. Lidar-Enabled Tree Map Matching for Real-Time and Drift-Free Harvester Positioning. *Remote Sens.* **2026**, *18*, 1243. <https://doi.org/10.3390/rs18081243>.
54. Xu, D.; Chen, G.; Zhang, S.; Jing, W. An Automated Pipeline for Extracting Forest Structural Parameters by Integrating UAV and Ground-Based LiDAR Point Clouds. *Forests* **2023**, *14*, 2179. <https://doi.org/10.3390/f14112179>.
55. Qin, H.; Zhou, W.; Yao, Y.; Wang, W. Individual tree segmentation and tree species classification in subtropical broadleaf forests using UAV-based LiDAR, hyperspectral, and ultrahigh-resolution RGB data. *Remote Sens. Environ.* **2022**, *280*, 113143. <https://doi.org/10.1016/j.rse.2022.113143>.
56. Wang, J.; Zhang, H.; Qiu, H.; Lei, K.; Yu, H.; Wang, X. Species-specific tree structural parameters extraction via UAV RGB-LiDAR data and multimodal instance segmentation. *Plant Phenomics* **2026**, *8*, 100171. <https://doi.org/10.1016/j.plaphe.2026.100171>.
57. Wang, L.; Zhang, H.; Fu, R.; Lei, K.; Liu, Y.; Yang, T.; Zhang, J.; Ge, X. MTSCFNet: A novel framework for improving tree species classification in a subtropical forest using RGB, LiDAR-derived, and GF-2 data. *J. For. Res.* **2026**, *37*, 22. <https://doi.org/10.1007/s11676-025-01964-2>.
58. Huo, Z.; Fang, L.; Chu, Y.; Dang, S.; Yang, J.; Li, L.; Li, X.; Ren, S.; Chen, J.; Peng, Y.; et al. Precise urban tree species identification and biomass estimation using UAV–Handheld LiDAR synergy and YOLOv11 deep learning. *Int. J. Appl. Earth Obs. Geoinf.* **2026**, *146*, 105049. <https://doi.org/10.1016/j.jag.2025.105049>.
59. Kwon, R.; Ryu, Y.; Yang, T.; Zhong, Z.; Im, J. Merging multiple sensing platforms and deep learning empowers individual tree mapping and species detection at the city scale. *ISPRS J. Photogramm. Remote Sens.* **2023**, *206*, 201–221. <https://doi.org/10.1016/j.isprsjprs.2023.11.011>.
60. Wang, M.; Zhang, Q.; Liu, X.; Zhang, J.; Yu, F.; Zhang, X.; Zhao, R. Research progress on multimodal data fusion in forest resource monitoring. *Front. Plant Sci.* **2025**, *16*, 1710618. <https://doi.org/10.3389/fpls.2025.1710618>.
61. Wang, A.; Shi, S.; Yang, J.; Luo, Y.; Tang, X.; Du, J.; Bi, S.; Qu, F.; Gong, C.; Gong, W. Integration of lidar and hyperspectral imagery for tree species identification at the individual tree level. *Photogramm. Rec.* **2025**, *40*, e70007. <https://doi.org/10.1111/phor.70007>.
62. Cao, Y.; Coops, N.C.; Murray, B.A.; Sinclair, I.; Geordie, R.M. M3FNet: Multi-modal multi-temporal multi-scale data fusion network for tree species composition mapping. *ISPRS J. Photogramm. Remote Sens.* **2026**, *231*, 797–814. <https://doi.org/10.1016/j.isprsjprs.2025.11.026>.

63. da Silva, D.Q.; dos Santos, F.N.; Filipe, V.; Sousa, A.J.; Oliveira, P.M. Edge AI-Based Tree Trunk Detection for Forestry Monitoring Robotics. *Robotics* **2022**, *11*, 136. <https://doi.org/10.3390/robotics11060136>.
64. Kymäläinen, H.; García-Pascual, B.; Martín-Cortés, C.; Rogai, M.; Selkimäki, M.; Zhou, X.; Arboix, E.; Acuna, M.; Mola-Yudego, B. March. Real-time Assistance with AR HUD Screen in the Forest Harvester Machine. In *Proceedings of the 10th Annual International GamiFIN Conference 2026, Saariselkä, Finland, 23–27 March 2026*; Association for Computing Machinery: New York, NY, USA, 2026; pp. 1–5. <https://doi.org/10.1145/3789624.3789637>.
65. Zhong, H.; Zhang, Z.; Liu, H.; Wu, J.; Lin, W. Individual Tree Species Identification for Complex Coniferous and Broad-Leaved Mixed Forests Based on Deep Learning Combined with UAV LiDAR Data and RGB Images. *Forests* **2024**, *15*, 293. <https://doi.org/10.3390/f15020293>.
66. Aral, A. The promise of neuromorphic edge AI for rural environmental monitoring. *Environ. Data Sci.* **2024**, *3*, e34. <https://doi.org/10.1017/eds.2024.36>.
67. Aguiar, A.S.; dos Santos, F.N.; Cunha, J.B.; Sobreira, H.; Sousa, A.J. Localization and Mapping for Robots in Agriculture and Forestry: A Survey. *Robotics* **2020**, *9*, 97. <https://doi.org/10.3390/robotics9040097>.
68. Yadav, N.; Rakholia, S.; Yosef, R. Decision Support Systems in Forestry and Tree-Planting Practices and the Prioritization of Ecosystem Services: A Review. *Land* **2024**, *13*, 230. <https://doi.org/10.3390/land13020230>.
69. Kpadé, C.P.; Tamini, L.D.; Pepin, S.; Khasa, D.P.; Abbas, Y.; Lamhamedi, M.S. Evaluating Multi-Criteria Decision-Making Methods for Sustainable Management of Forest Ecosystems: A Systematic Review. *Forests* **2024**, *15*, 1728. <https://doi.org/10.3390/f15101728>.
70. Perry, N.; Schweier, J.; Bont, L.G.; Kushwaha, S.; Heli, P.; Eyvindson, K.; Astrup, R.; Chapman, M.; Blatter, C. Towards Precision Forestry: A Systematic Review of Optimisation Methods for Individual-Tree Decisions in Forest Management. *Trees For. People* **2026**, *24*, 101226. <https://doi.org/10.1016/j.tfp.2026.101226>.
71. Pascual, A.; Guerra-Hernández, J. Spatial connectivity in tree-level decision-support models using mathematical optimization and individual tree mapping. *For. Policy Econ.* **2022**, *139*, 102732. <https://doi.org/10.1016/j.forpol.2022.102732>.
72. Piedra-Jimenez, F.; Broz, D.; Novas, J.M.; Grossmann, I.E.; Rodriguez, M.A. Optimizing forest planning: Balancing timber production and carbon sequestration through a multi-objective disjunctive formulation. *For. Policy Econ.* **2025**, *178*, 103575. <https://doi.org/10.1016/j.forpol.2025.103575>.
73. Niemi, M.T.; Vauhkonen, J.; Uusitalo, J. Trade-off analysis for multi-objective boreal forest thinning. *Forestry* **2026**, *99*, cpaf027. <https://doi.org/10.1093/forestry/cpaf027>.
74. Sedmák, R.; Scheer, L.; Kúdela, P.; Vencúrik, J.; Modranský, J.; Daniš, D.; Fabrika, M. Multicriteria optimization of ecosystem services as a base for participative forest management promoting recreation near tourist centers and cities. *Trees For. People* **2025**, *21*, 100905. <https://doi.org/10.1016/j.tfp.2025.100905>.
75. Tagarakis, A.C.; Benos, L.; Kyriakarakos, G.; Pearson, S.; Sørensen, C.G.; Bochtis, D. Digital Twins in Agriculture and Forestry: A Review. *Sensors* **2024**, *24*, 3117. <https://doi.org/10.3390/s24103117>.
76. Qiu, H.; Zhang, H.; Lei, K.; Zhang, H.; Hu, X. Forest digital twin: A new tool for forest management practices based on Spatio-Temporal Data, 3D simulation Engine, and intelligent interactive environment. *Comput. Electron. Agric.* **2023**, *215*, 108416. <https://doi.org/10.1016/j.compag.2023.108416>.
77. Sasaki, N.; Abe, I. A Digital Twin Architecture for Forest Restoration: Integrating AI, IoT, and Blockchain for Smart Ecosystem Management. *Future Internet* **2025**, *17*, 421. <https://doi.org/10.3390/fi17090421>.
78. Astola, H.; Kangas, A.; Minunno, F.; Möttus, M. Emulating a forest growth and productivity model with deep learning. *Silva Fenn.* **2026**, *60*, 25012. <https://doi.org/10.14214/sf.25012>.
79. Schwartz, M.; Ciais, P.; Sean, E.; De Truchis, A.; Vega, C.; Besic, N.; Fayad, I.; Wigneron, J.P.; Brood, S.; Pelissier-Tanon, A.; et al. Retrieving yearly forest growth from satellite data: A deep learning based approach. *Remote Sens. Environ.* **2025**, *330*, 114959. <https://doi.org/10.1016/j.rse.2025.114959>.
80. Tanase, M.; Martini, J.P.; Miranda, P.; Garcia, D.; Wilke, V.; Miguel, S.; Mihai, C.; Diez, J.; Natal, S.; San Martin, D.; et al. Long-term forest structure trends in the peninsular Spain from lidar-optical sensors synergies. *Remote Sens. Environ.* **2026**, *334*, 115196. <https://doi.org/10.1016/j.rse.2025.115196>.
81. Thapa, N.; Narine, L.L.; Wilson, A.E. Forest aboveground biomass estimation using airborne LiDAR: A systematic review and meta-analysis. *J. For.* **2025**, *123*, 389–412. <https://doi.org/10.1007/s44392-025-00029-w>.
82. Giambelluca, A.L.; Hermosilla, T.; Alvarez-Mozos, J.; Gonzalez-Audicana, M. Identifying forest harvesting practices: Clear-cutting and thinning in diverse tree species using dense Landsat time series. *For. Ecol. Manag.* **2025**, *578*, 122442. <https://doi.org/10.1016/j.foreco.2024.122442>.

83. Varveris, D.; Basdekidou, V.; Basdekidou, C.; Xofis, P. Smart Forest Modeling Behavioral for a Greener Future: An AI Text-by-Voice Blockchain Approach with Citizen Involvement in Sustainable Forestry Functionality. *FinTech* **2025**, *4*, 47. <https://doi.org/10.3390/fintech4030047>.
84. Peksa, J.; Mamchur, D. State-of-the-Art on Digital Twin Technologies for Industrial Applications and the Federated Digital Twin Lifecycle Model (F-DTLM). *Automation* **2026**, *7*, 77. <https://doi.org/10.3390/automation7030077>.
85. Zamora-Cristales, R.; Kogler, C.; Sessions, J.; Väättäinen, K.; Contreras, M.; Alzamora, R.M.; Spinelli, R. Simulation in Forest Engineering: Planning the Next Generation of Forest Operations and Logistics for the Challenges of the XXI Century. *Curr. For. Rep.* **2026**, *12*, 8. <https://doi.org/10.1007/s40725-026-00268-w>.
86. Latterini, F.; Camarretta, N.; Watt, M.S. Remote Sensing for Planning Harvesting Operations and Monitoring Their Effects on the Forest Ecosystem: State of the Art and Future Perspectives. *For. Ecol. Manag.* **2025**, *597*, 123175. <https://doi.org/10.1016/j.foreco.2025.123175>.
87. Forkuo, G.O.; Picchio, R.; Proto, A.R.; Borz, S.A. Applications of Artificial Intelligence in Forest Operations Engineering Research: A Systematic Review. *Curr. For. Rep.* **2026**, *12*, 14. <https://doi.org/10.1007/s40725-026-00275-x>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.