

Synthetic Aperture Radar, Multi-Source Data Fusion, and Machine Learning for Aboveground Biomass and Carbon Stock Estimation: A Systematic Review and Bibliometric Analysis

Temitope Emmanuel Paul

u23944359@tuks.co.za

University of Pretoria

Adedayo Adeleke

University of Pretoria

Abel Ramoelo

University of Pretoria

Systematic Review

Keywords: Aboveground biomass, Carbon estimation, Synthetic aperture radar (SAR), Multi-source remote sensing, Machine learning, LiDAR, Urban trees

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11364089/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Synthetic Aperture Radar, Multi-Source Data Fusion, and Machine Learning for Aboveground Biomass and Carbon Stock Estimation: A Systematic Review and Bibliometric Analysis

Temitope Emmanuel Paul^{1,2}, Adedayo Adeleke¹, Abel Ramoelo^{1,3}

¹Department of Geography, Geoinformatics and Meteorology, Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa. (City: Pretoria)

²Department of Surveying and Geoinformatics, Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria. (City: Akungba-Akoko)

³Earth Observations Programme, South African National Space Agency, CSIR campus, Pretoria, South Africa. (City: Pretoria).

Temitope Emmanuel Paul (u23944359@tuks.co.za)

Adedayo Adeleke (adedayo.adeleke@up.ac.za)

Abel Ramoelo (abel.ramoelo@up.ac.za)

Corresponding Author: Temitope Emmanuel Paul

Abstract

Accurate estimation of aboveground biomass (AGB) and carbon stocks is fundamental for sustainable forest management, climate change mitigation, and ecosystem monitoring. Advances in synthetic aperture radar (SAR), multi-source remote sensing, and artificial intelligence have substantially improved biomass estimation; however, existing research remains methodologically fragmented, with limited synthesis of sensor fusion strategies, machine learning, and deep learning approaches, and their applicability to urban tree carbon estimation. This study presents a systematic review and bibliometric analysis of peer-reviewed literature published between 2004 and 2026 following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines. Literature was retrieved from Scopus, Web of Science, and Google Scholar, screened using Rayyan, and 68 eligible studies were identified for qualitative synthesis. The review examined publication trends, geographical distribution, remote sensing platforms, data fusion strategies, machine learning and deep learning algorithms, validation approaches, and emerging research gaps. The findings reveal a substantial increase in publications after 2018, driven by the widespread availability of Sentinel-1, Sentinel-2, LiDAR, cloud-based geospatial platforms, and advances in artificial intelligence. The reviewed literature demonstrates a clear transition from single-sensor approaches towards feature-level fusion of SAR, optical, and LiDAR data, which consistently improved biomass and carbon estimation compared with individual sensors. Among predictive models, ensemble learning algorithms, particularly Random Forest and Extreme Gradient Boosting, remained the most robust and widely adopted, while deep learning approaches showed increasing potential for high-resolution applications where sufficient training

data were available. Despite these advances, research focusing specifically on urban tree biomass and carbon estimation remains limited, highlighting challenges related to heterogeneous urban environments, limited reference datasets, SAR integration, and model transferability. This review provides a comprehensive synthesis of current methodological developments, identifies critical knowledge gaps, and proposes future research directions to develop robust, scalable, interpretable, and transferable frameworks for biomass and carbon estimation across natural and urban ecosystems.

Keywords: Aboveground biomass, Carbon estimation, Synthetic aperture radar (SAR), Multi-source remote sensing, Machine learning, LiDAR, Urban trees.

1.0 Introduction

Climate change has become one of the most pressing environmental challenges of the twenty-first century, with increasing atmospheric greenhouse gas concentrations driving global warming, biodiversity loss, and ecosystem degradation. Forest ecosystems play a pivotal role in regulating the global carbon cycle by acting as major carbon sinks that absorb and store atmospheric carbon dioxide (CO₂), thereby mitigating the impacts of climate change. Accurate estimation of aboveground biomass (AGB) and carbon stocks is therefore fundamental for understanding ecosystem productivity, assessing carbon sequestration potential, supporting sustainable forest management, and informing international climate policies such as the Paris Agreement and the United Nations Sustainable Development Goals (IPCC, 2023).

Reliable biomass estimation is equally important in urban environments, where trees provide critical ecosystem services beyond carbon sequestration, including temperature regulation, air pollution mitigation, stormwater management, biodiversity conservation, and improvements in human well-being (Livesley et al., 2016; Nowak & Greenfield, 2018). As cities continue to expand globally, urban forests are increasingly recognized as an essential component of climate adaptation and resilience strategies. Consequently, developing accurate, scalable, and transferable approaches for estimating tree biomass and carbon stocks across both natural and urban ecosystems has become a priority for researchers, forest managers, and policymakers.

Conventional field-based forest inventories remain the benchmark for biomass estimation because they provide direct measurements of tree characteristics such as diameter at breast height (DBH), tree height, and species composition. However, these methods are labour-intensive, time-consuming, costly, and often impractical over large or inaccessible areas. In addition, repeated field campaigns are difficult to implement, limiting their ability to support continuous monitoring of rapidly changing forest environments (Gibbs et al., 2007; Picard et al., 2012). These limitations have accelerated the adoption of remote sensing technologies that provide consistent, spatially explicit, and repetitive observations across multiple spatial and temporal scales.

Optical remote sensing has long been the most widely applied approach for biomass estimation because of its ability to capture vegetation spectral characteristics related to canopy condition,

photosynthetic activity, and vegetation health. Satellite missions such as Landsat and Sentinel-2 have substantially improved vegetation monitoring by providing freely accessible, high-resolution multispectral imagery suitable for regional and global assessments (Wulder et al., 2022). Nevertheless, optical sensors are constrained by cloud cover, atmospheric effects, and signal saturation in dense vegetation, which often reduce their effectiveness in estimating biomass in structurally complex forests (Mitchard, 2018).

To overcome these limitations, Synthetic Aperture Radar (SAR) has emerged as a powerful active remote sensing technology capable of acquiring observations independent of daylight and most weather conditions. Microwave signals penetrate vegetation canopies and interact with branches, stems, and trunks, providing structural information that complements the spectral information derived from optical imagery. Advances in spaceborne SAR missions, including Sentinel-1, ALOS PALSAR, TerraSAR-X, and RADARSAT, have significantly expanded opportunities for estimating forest structure, biomass, and carbon stocks across diverse ecosystems (Joshi et al., 2016; Sinha et al., 2015). However, SAR data are also affected by challenges such as speckle noise, geometric distortions, signal saturation in high-biomass forests, and sensitivity to imaging geometry, requiring appropriate preprocessing and feature extraction to maximize predictive performance.

Light Detection and Ranging (LiDAR) further complements optical and SAR observations by providing highly accurate three-dimensional information on vegetation structure, including canopy height, vertical complexity, and forest volume. Airborne and terrestrial LiDAR have consistently demonstrated superior performance in biomass estimation due to their direct representation of forest structure. More recently, the increasing availability of spaceborne LiDAR missions, such as the Global Ecosystem Dynamics Investigation (GEDI), has created new opportunities for integrating structural information with optical and SAR observations to improve biomass estimation across large geographic regions (Dubayah et al., 2020). Similarly, unmanned aerial vehicles (UAVs) equipped with optical, multispectral, hyperspectral, LiDAR, and SAR sensors have enabled high-resolution monitoring of vegetation characteristics, particularly within heterogeneous urban landscapes.

Rather than functioning as competing technologies, optical imagery, SAR, LiDAR, and UAV observations provide complementary spectral, structural, and geometric information. Consequently, increasing attention has been paid to multi-source remote sensing data fusion, in which information from multiple sensors is integrated to exploit their complementary strengths while mitigating the limitations of individual datasets. Depending on the application, data fusion may occur at the pixel, feature, or decision level, each providing different advantages for biomass estimation. Numerous studies have demonstrated that feature-level fusion, in particular, improves predictive performance by combining spectral, textural, and structural variables into unified modelling frameworks (Luo et al., 2025; Ma et al., 2019).

The rapid growth of multi-source remote sensing has been accompanied by significant advances in machine learning and deep learning algorithms. Traditional empirical and regression-based approaches often struggle to capture the complex, nonlinear relationships between remotely sensed variables and forest attributes. In contrast, machine learning algorithms including Random Forest (RF), Support Vector Regression (SVR), Extreme Gradient Boosting (XGBoost), Artificial Neural Networks (ANN), and deep learning architectures such as Convolutional Neural Networks (CNNs) have demonstrated superior capability in modelling high-dimensional, nonlinear interactions while effectively handling heterogeneous datasets (Belgiu & Drăguț, 2016; Maxwell et al., 2018). These methodological advances have shifted biomass estimation from conventional statistical modelling towards data-driven approaches capable of integrating diverse remote sensing datasets with increasing accuracy and computational efficiency.

Despite these substantial developments, the rapidly expanding literature has become increasingly fragmented. Existing review studies have typically focused on individual remote sensing technologies, specific machine learning algorithms, or general applications of remote sensing for biomass estimation, with comparatively limited attention given to the integrated use of SAR, optical imagery, LiDAR, multi-source data fusion, and machine learning within a single systematic framework. Furthermore, relatively few reviews have combined systematic evidence synthesis with bibliometric analysis to evaluate the evolution of research themes, scientific collaboration, methodological trends, and emerging research priorities. This fragmentation limits researchers' ability to identify methodological consensus, evaluate technological progress, and determine future directions for biomass and carbon stock estimation.

Table 1. Comparison of Previous Review Studies and the Present Review

Reference	Review focus	Main contribution	Remaining scope/gap relative to the present review	Contribution of the present review
Maxwell et al. (2018).	Machine-learning classification in remote sensing	Reviewed the implementation, strengths, limitations, and practical considerations of ML classifiers for remotely sensed imagery.	Broadly focused on remote-sensing classification rather than biomass/carbon estimation and multi-sensor biomass retrieval.	Extends the ML perspective specifically to biomass and carbon estimation using SAR, optical imagery, LiDAR, and multi-source fusion.
Ma et al. (2019).	Deep learning in remote sensing	Synthesized more than 200 studies and reviewed DL applications including classification, object detection, segmentation, image	Broad remote-sensing focus rather than a dedicated assessment of biomass/carbon retrieval and associated sensor-fusion workflows.	Compares conventional ML and emerging DL approaches specifically within biomass and carbon estimation.

		fusion, registration, and OBIA		
Ullah et al. (2025)	AGB monitoring across terrestrial ecosystems	Systematically assessed spatiotemporal research trends, study scales, remote-sensing technologies, sensor synergies, ML performance, and uncertainties in AGB estimation.	Provides a broad terrestrial-ecosystem synthesis rather than specifically centering on SAR-led fusion architectures, fusion levels, and transfer of biomass/carbon estimation approaches to heterogeneous urban-tree environments..	Provides a targeted synthesis of SAR, optical and LiDAR integration, fusion strategies, ML approaches, validation practices, bibliometric trends, and implications for urban-tree carbon estimation..
Luo et al. (2025)	Forest above-ground biomass dynamics	Reviewed remote-sensing approaches for monitoring temporal variations in forest AGB and highlighted the value of time-series observations for biomass change assessment.	Primarily concerned with forest AGB dynamics rather than an integrated assessment of bibliometric patterns, SAR-centred multi-source fusion, ML performance, and urban-tree carbon estimation.	Integrates systematic and bibliometric evidence with assessment of SAR-centred fusion, ML approaches, validation, and urban-tree carbon applications.
Ullah et al. (2025)	Aboveground biomass monitoring in terrestrial ecosystems	Systematically reviewed remote sensing approaches for aboveground biomass monitoring across terrestrial ecosystems, evaluating sensor technologies and emerging remote sensing applications.	Emphasized remote sensing technologies but provided limited comparison of data fusion strategies, machine learning algorithms, bibliometric evolution, and the transferability of methods to heterogeneous urban environments.	Provides a comprehensive synthesis of SAR, optical imagery, LiDAR, multi-source data fusion, machine learning, bibliometric trends, validation strategies, and research priorities for urban tree carbon estimation.
Present study	SAR, multi-source remote sensing, ML, and bibliometric analysis for biomass/carbon estimation	Integrates evidence on sensors, data-fusion strategies, ML algorithms, validation approaches, bibliometric patterns, research gaps, and future directions.	-	Provides an integrated methodological assessment with particular emphasis on SAR-centred multi-source data fusion and its relevance to urban-tree carbon estimation

				and operational carbon monitoring.
--	--	--	--	------------------------------------

133

134 As summarized in Table 1, previous reviews have made valuable contributions to understanding
135 remote sensing applications for biomass estimation; however, they generally evaluate individual
136 sensors or modelling techniques rather than comprehensively examining the integration of SAR,
137 multi-source remote sensing data fusion, machine learning, and bibliometric evidence. Moreover,
138 few studies explicitly discuss the implications of these methodological developments for urban
139 tree carbon estimation, despite increasing global interest in nature-based solutions for climate
140 change mitigation and sustainable urban development.

141 Accordingly, this study presents a PRISMA-guided systematic review and bibliometric analysis of
142 peer-reviewed literature published between 2004 and 2026 to synthesize methodological
143 developments in SAR, multi-source remote sensing data fusion, and machine learning for
144 estimating aboveground biomass and carbon stocks. Specifically, the review analyses publication
145 trends, geographical distribution, remote sensing platforms, data fusion strategies, predictive
146 modelling techniques, validation approaches, and emerging research themes while identifying
147 methodological challenges and future research priorities. Particular attention is given to the
148 potential application of these advances for urban tree carbon assessment, where robust and
149 transferable modelling frameworks remain comparatively underdeveloped.

150 The remainder of this paper is organized as follows. Section 2 describes the systematic review
151 methodology, including the literature search strategy, study selection process, data extraction
152 procedures, and bibliometric analysis. Section 3 presents the results of the systematic review and
153 bibliometric analyses. Section 4 discusses the major methodological advances, research gaps, and
154 implications for biomass and carbon stock estimation, while Section 5 summarizes the principal
155 conclusions and outlines future research directions.

156 The conceptual framework for this systematic review, which illustrates the integration of multi-
157 source remote sensing, data fusion, and machine learning for biomass and carbon estimation, is
158 given in Figure 1 below.

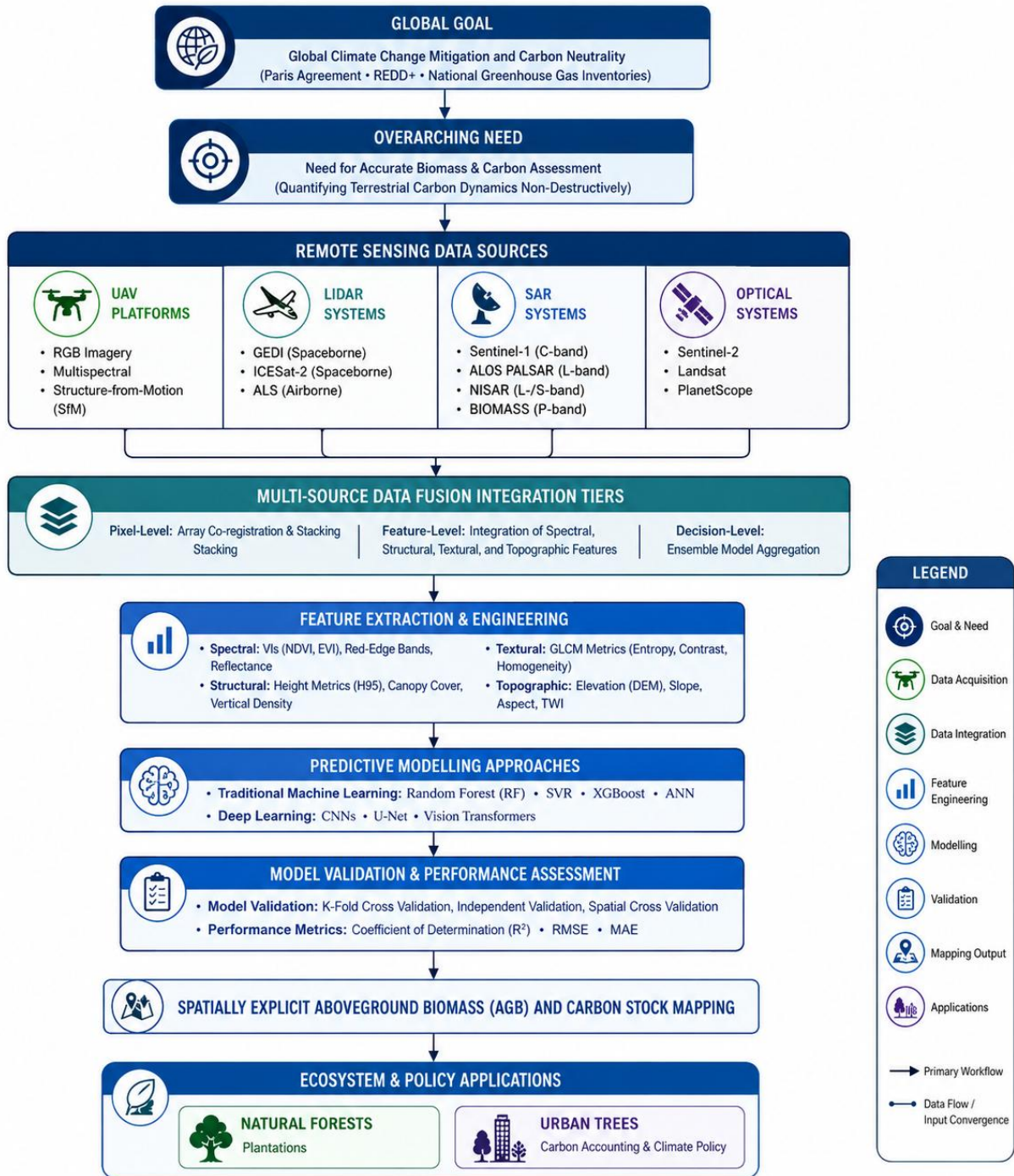


Figure 1. Conceptual framework illustrating the integration of multi-source remote sensing, data fusion, and machine learning for biomass and carbon estimation.

2.0 Materials and Methods

This study employed a systematic literature review guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework to ensure a transparent, reproducible, and methodologically rigorous synthesis of the literature. PRISMA provides internationally recognized guidelines for the identification, screening, eligibility assessment, and inclusion of scientific studies, thereby reducing selection bias and improving the transparency and reproducibility of evidence synthesis (Page et al., 2021). The PRISMA framework has become the preferred reporting standard for systematic reviews across environmental sciences, remote sensing, forestry, and geospatial research because it promotes consistency in study selection and documentation.

Prior to commencing the review, a structured review protocol was developed to ensure consistency throughout the research process. The protocol defined the review objectives, literature sources, search strategy, eligibility criteria, study selection procedures, data extraction variables, quality assessment criteria, and evidence synthesis approach. Establishing the protocol before conducting the database search reduced the likelihood of subjective decision-making during the review and ensured that all studies were evaluated using identical criteria.

The review was undertaken through six sequential stages: (i) literature identification, (ii) duplicate removal, (iii) title and abstract screening, (iv) full-text eligibility assessment, (v) data extraction and quality assessment, and (vi) qualitative and bibliometric synthesis. Each stage was designed to progressively refine the evidence base while ensuring that only methodologically relevant and scientifically robust studies were retained for detailed analysis. Unlike traditional narrative reviews, this systematic approach enables reproducibility, minimizes reviewer bias, and provides a transparent framework for comparing methodological developments across multiple studies.

Following the eligibility assessment, information from each included article was systematically extracted using a standardized data extraction template developed in Microsoft Excel. Extracted variables included publication year, country or region, ecosystem type, remote sensing datasets, SAR sensor characteristics, complementary data sources, data fusion strategy, machine learning or deep learning algorithm, validation approach, reported performance metrics, and principal findings. The compiled dataset subsequently formed the basis for descriptive statistics, bibliometric analysis, comparative methodological evaluation, and qualitative evidence synthesis.

Rather than providing a descriptive summary of previous studies, this review critically synthesizes methodological developments across diverse ecosystems, remote sensing technologies, sensor fusion strategies, and artificial intelligence approaches to identify emerging research trends, persistent methodological limitations, and future opportunities for improving tree biomass and carbon stock estimation.

2.1 Literature Search Strategy

A comprehensive literature search was conducted in accordance with the PRISMA 2020 guidelines to identify peer-reviewed studies that investigated the use of Synthetic Aperture Radar (SAR), multi-source remote sensing data, and machine learning techniques for estimating tree biomass and carbon stocks. To maximize coverage and minimize publication bias, three internationally recognized bibliographic databases, Scopus, Web of Science Core Collection, and Google Scholar (accessed using Harzing's Publish or Perish), were searched systematically. These databases were selected because they collectively provide extensive coverage of high-quality publications in remote sensing, forestry, environmental sciences, geospatial analysis, and artificial intelligence.

The search encompassed studies published between 2004 and 2026, representing a period characterized by rapid developments in satellite remote sensing, freely available Earth observation datasets, cloud computing platforms, and machine learning applications for environmental monitoring. Restricting the review to this period ensured that the synthesis captured contemporary methodological advances while avoiding the outdated techniques that are no longer widely applied.

To ensure a systematic and reproducible search process, the search strategy was organized into four thematic groups representing the principal concepts underpinning the review: (i) vegetation and forest resources, (ii) biomass and carbon estimation, (iii) remote sensing technologies, and (iv) machine learning and artificial intelligence methods. Within each thematic group, synonymous terms were combined using the Boolean operator OR to broaden retrieval, while the four thematic groups were connected using the Boolean operator AND to ensure that retrieved studies addressed all major components of the research objectives simultaneously as shown in table 2 below.

Table 2. Search terms and corresponding Boolean search strategy used for literature retrieval

Publish or Perish retrieval keywords	Boolean operator meaning
"Urban tree" "Urban vegetation" "Urban forest" Tree Forest "Aboveground biomass" Biomass "Carbon estimation" "Carbon stock" "Carbon storage" "Carbon sequestration" "Biomass measurement" "Carbon assessment" "Carbon evaluation" "Carbon inventory" "Machine learning" "Deep learning" "Artificial intelligence" "Random Forest" "Support Vector Machine" XGBoost "Synthetic Aperture Radar" SAR Sentinel-1 Sentinel-2 Landsat LiDAR Optical "Remote sensing" "Multi-source remote sensing" "Data fusion" UAV	("Urban tree" OR "Urban vegetation" OR "Urban forest" OR Tree OR Forest) AND ("Aboveground biomass" OR Biomass OR "Carbon estimation" OR "Carbon stock" OR "Carbon storage" OR "Carbon sequestration" OR "Biomass measurement" OR "Carbon assessment" OR "Carbon evaluation" OR "Carbon inventory") AND ("Machine learning" OR "Deep learning" OR "Artificial intelligence" OR "Random Forest" OR "Support Vector Machine" OR XGBoost) AND ("Synthetic Aperture Radar" OR SAR OR Sentinel-1 OR Sentinel-2 OR Landsat OR LiDAR OR Optical OR UAV) AND ("Remote sensing" OR "Multi-source remote sensing" OR "Data fusion")

Although the conceptual structure of the search strategy remained consistent across all databases, minor modifications to the syntax were made where necessary to comply with the indexing rules

and search functionalities of individual databases. Only peer-reviewed journal articles published in English were considered eligible. Conference proceedings, dissertations, theses, editorials, book chapters, review papers, and duplicate records were excluded during the screening process to ensure methodological consistency and maintain the scientific quality of the evidence base.

Google Scholar searches were conducted using Harzing's Publish or Perish software, which retrieves the highest-ranked records from Google Scholar while facilitating citation management and export. Because Google Scholar limits the number of records that can be retrieved per query, identical searches were also conducted in Scopus and Web of Science to maximize coverage and reduce potential retrieval bias. Records obtained from all databases were merged into a unified reference library, in which duplicate publications were automatically identified and manually verified before screening commenced.

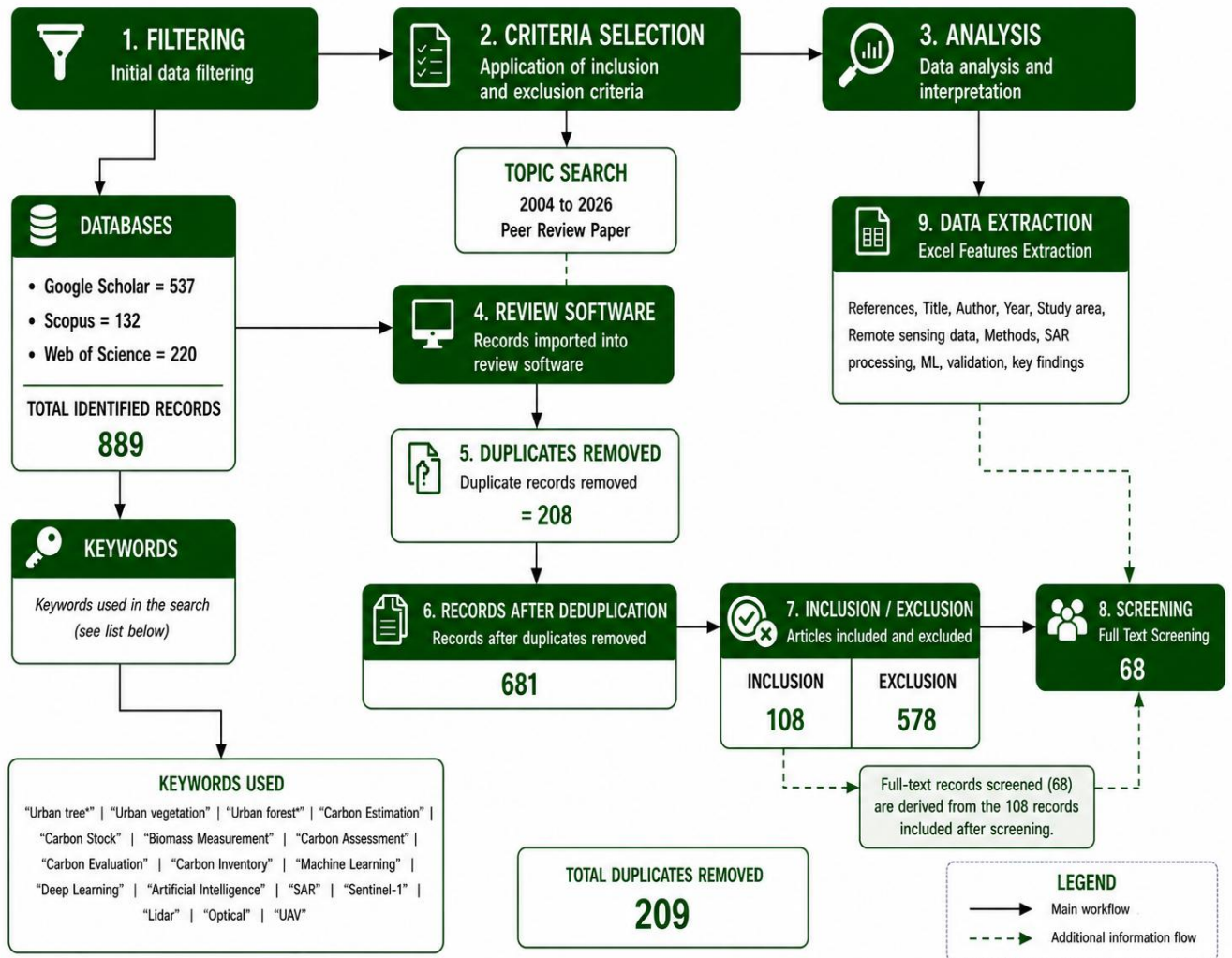


Figure 2. Systematic review workflow (PRISMA)

2.2 Study Selection and Eligibility Criteria

The study selection process followed the PRISMA 2020 framework and was managed using Rayyan, a web-based platform specifically developed to support systematic reviews by facilitating study screening, duplicate detection, reviewer collaboration, and conflict resolution (Ouzzani et al., 2016). All records retrieved from Scopus, Web of Science, and Google Scholar were exported in RIS format and imported into Rayyan, where duplicate records were automatically identified and removed. The remaining records were subsequently verified manually to ensure the accuracy of the duplicate removal process.

Following deduplication, the screening process was conducted in two sequential stages within the Rayyan platform. During the first stage, the titles and abstracts of all retrieved studies were independently screened against the predefined eligibility criteria to exclude publications that were clearly unrelated to the review's objectives. Studies considered potentially relevant progressed to the second stage, where their full texts were independently assessed for eligibility. Rayyan's labelling, filtering, and conflict identification features facilitated the systematic organization of screening decisions and enabled efficient comparison of reviewer assessments throughout the selection process.

The inclusion and exclusion criteria were established prior to screening to ensure consistency, minimize selection bias, and enhance the reproducibility of the review. Any disagreements regarding study eligibility were resolved through discussion among the reviewers until consensus was reached. The complete study selection procedure is summarized in the PRISMA flow diagram (Figure 2).

2.2.1 Inclusion Criteria

Studies were included if they:

- Investigated aboveground biomass (AGB), tree biomass, carbon stock, or carbon estimation in urban or natural forest ecosystems
- Utilized SAR data independently or in combination with complementary remote sensing datasets such as optical imagery, LiDAR, or UAV observations;
- Applied machine learning, deep learning, or artificial intelligence algorithms for biomass estimation, classification, or predictive modelling
- Reported at least one quantitative validation metric, including coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), overall accuracy (OA), or Kappa coefficient; and
- Were published as peer-reviewed journal articles in English between January 2004 and December 2024.

Recognizing the relatively limited number of studies focusing specifically on urban tree carbon estimation, forest biomass studies conducted in natural, plantation, and managed forest ecosystems

were also included where they provided methodological insights relevant to urban environments, particularly in relation to SAR preprocessing, multi-source data fusion, feature engineering, and machine learning model development.

2.2.2 Exclusion Criteria

Studies were excluded if they:

- Were not published in English;
- Were conference proceedings, theses, dissertations, editorials, book chapters, or review articles
- Did not incorporate SAR data within the analytical framework
- Did not employ machine learning, deep learning, or artificial intelligence techniques; or
- Lacked sufficient methodological detail or failed to report quantitative validation metrics.

The complete study selection procedure is presented in the PRISMA flow diagram in Figure 2 below.

2.3 Data Extraction and Quality Assessment

Following the full-text eligibility assessment, data extraction was undertaken using a standardized Microsoft Excel spreadsheet developed specifically for this review. Information was systematically extracted from each eligible study, including the publication year, country or region, ecosystem type (urban trees, urban forests, natural forests, plantations, mangroves, or other forested environments), SAR sensor and frequency band, complementary remote sensing datasets, data fusion strategy, machine learning or deep learning algorithm, field reference data, validation methodology, reported performance metrics (e.g., coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), overall accuracy (OA), and Kappa coefficient), together with the principal findings related to biomass and carbon stock estimation. Standardizing the extraction process ensured consistency across all included studies and facilitated comprehensive comparisons of methodological developments, remote sensing technologies, modelling approaches, and predictive performance.

To enhance methodological transparency and minimize selection bias, Rayyan was used throughout the study selection process, including duplicate identification, title and abstract screening, full-text eligibility assessment, reviewer conflict management, and documentation of inclusion and exclusion decision (Ouzzani et al., 2016). Three independent reviewers, comprising the principal researcher, the academic supervisor, and an independent PhD graduate, independently evaluated all retrieved studies according to the predefined eligibility criteria. Rayyan's duplicate detection, labelling, tagging, filtering, and conflict identification functions supported the systematic management of records throughout the review. Any disagreements arising during the screening or eligibility assessment stages were resolved through discussion until consensus was achieved.

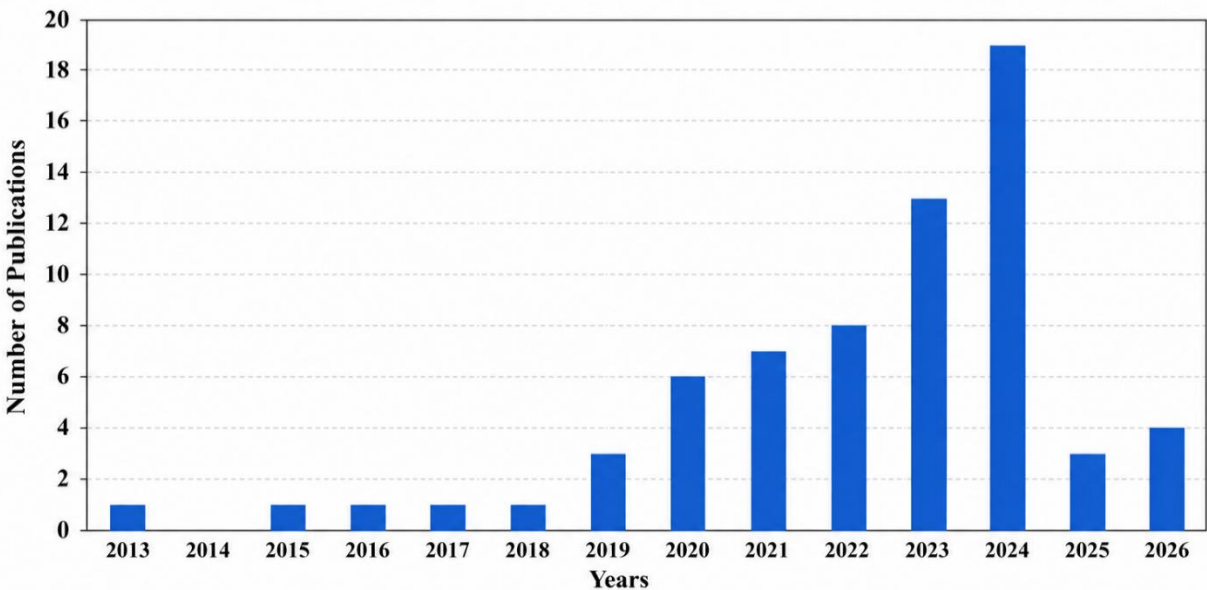
Rather than assigning numerical quality scores, the methodological quality assessment focused on evaluating the completeness, clarity, and reproducibility of the included studies. Particular attention was given to the description of the research objectives, remote sensing datasets, preprocessing procedures, feature extraction methods, machine learning workflow, validation strategy, and reporting of performance metrics. This qualitative appraisal ensured that the evidence synthesis was based on studies that provided sufficient methodological detail for robust comparison, while avoiding the introduction of subjective quality-weighting schemes.

The extracted dataset subsequently supported descriptive statistical analyses, bibliometric visualizations, comparative evaluations of remote sensing technologies, data fusion strategies, and machine learning algorithms, as well as the identification of methodological gaps, research challenges, and future directions presented in the subsequent Results and Discussion sections.

3.0 Results

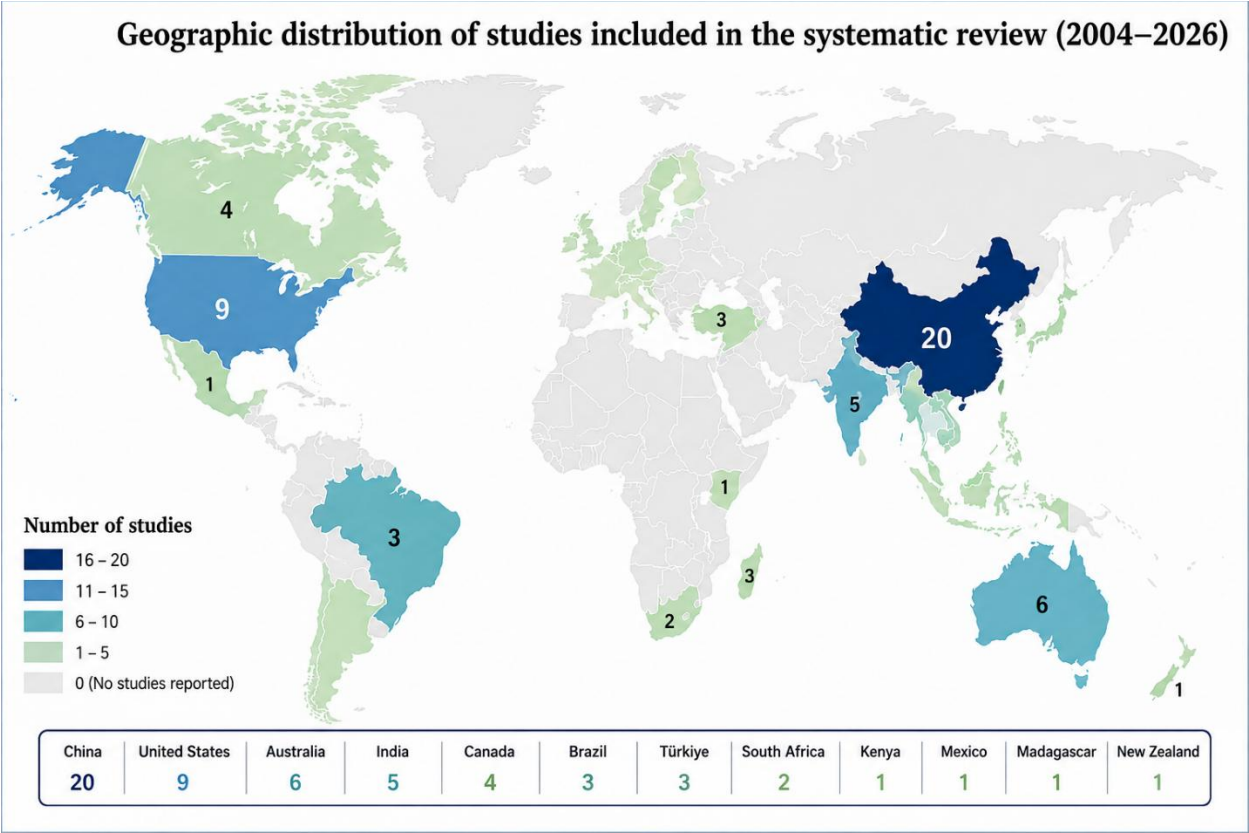
3.1 Publication Trends and Research Growth

The systematic review revealed a clear and sustained increase in research on integrating synthetic aperture radar (SAR), multi-source remote sensing, and machine learning for biomass and carbon estimation between 2004 and 2026 (Figure 3 & Figure 4). Although the search covered more than two decades, no eligible studies meeting the predefined inclusion criteria were identified before 2013. Publication output remained relatively low from 2013 to 2018, with only isolated studies investigating the combined use of SAR and complementary remote sensing datasets for biomass estimation. This limited research activity reflects the early stage of multi-source remote sensing integration, constrained by the limited availability of high-resolution Earth observation data, computational resources, and mature machine learning techniques during that period (Wulder et al., 2022; Zhu et al., 2026).



337

Figure 3. Annual publication trend (2004–2026)



338

339

Figure 4. Geographic distribution of studies

340

341

342

343

344

345

346

347

348

349

350

351

A marked increase in publication output emerged from 2019 onwards and continued through 2026, with the highest research activity occurring between 2021 and 2024, followed by sustained contributions from recent studies published in 2025 and 2026 (Figure 3). The additional studies published during 2025–2026 further demonstrate the continued expansion of research on urban tree inventory, biomass estimation, carbon stock mapping, and the integration of multi-source remote sensing. This growth has been driven by the widespread availability of freely accessible satellite missions such as Sentinel-1, Sentinel-2, and Landsat, together with increasing access to LiDAR, UAV platforms, cloud-based geospatial computing environments, and advances in machine learning and deep learning algorithms for Earth observation applications (Wulder et al., 2022). These technological developments have enabled researchers to integrate spectral, structural, and spatial information from multiple data sources, thereby improving the accuracy, scalability, and operational applicability of biomass and carbon estimation.

352

353

354

355

356

The studies published in 2025 and 2026 also indicate a continued shift towards advanced multi-source data fusion and artificial intelligence-based approaches. Recent investigations increasingly employed deep learning models, UAV–LiDAR integration, hyperspectral data fusion, and street-view imagery to improve urban tree species classification, biomass estimation, and carbon stock assessment. However, despite these methodological advances, the newly screened studies reveal

that SAR remains underutilized within urban tree carbon estimation, with most investigations relying predominantly on optical imagery, LiDAR, UAV-derived datasets, or combinations of these sensors. This finding highlights an important research gap and reinforces the need for more studies integrating SAR with complementary remote sensing datasets to exploit its all-weather imaging capability and structural sensitivity for urban vegetation mapping.

Despite encouraging growth in publication output, the distribution of the reviewed studies indicates that research remains concentrated primarily on natural and peri-urban forest ecosystems, whereas comparatively fewer studies have focused on urban trees and heterogeneous urban environments. Consequently, the transferability of many existing biomass and carbon estimation models to complex urban landscapes remains uncertain. The continued scarcity of studies integrating SAR with multi-source remote sensing and machine learning for urban tree carbon estimation underscores a significant knowledge gap and highlights an important direction for future research and methodological development. This gap further underscores the novelty and relevance of the present review, which synthesizes current evidence and identifies emerging opportunities to advance urban tree carbon estimation through integrated remote sensing approaches. The reviewed studies employed a diverse range of remote sensing data sources, reflecting the increasing use of complementary sensor technologies for biomass and carbon estimation. The distribution of the principal data types identified across the reviewed literature is presented in Figure 5

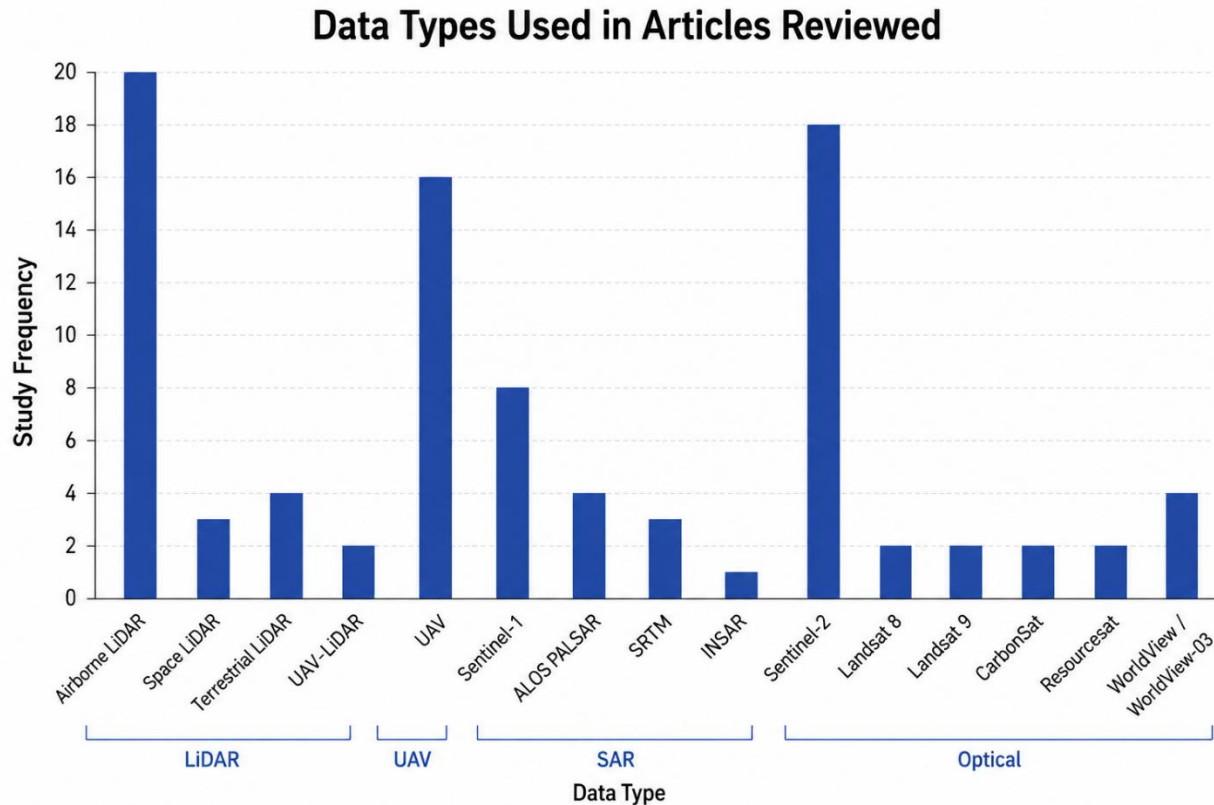


Figure 5. Distribution of remote sensing data types used in the articles reviewed

3.2 Bibliometric Analysis and Research Themes

The bibliometric analysis was conducted using VOSviewer (version 1.6.20) to examine the conceptual structure and thematic evolution of research on SAR-based multi-source remote sensing and machine learning for biomass and carbon estimation. The co-occurrence network was generated from terms extracted from the titles and abstracts of the reviewed studies using the full counting method. After applying a minimum occurrence threshold, frequently occurring terms were grouped into clusters based on their co-occurrence relationships, thereby identifying the dominant research themes in the literature (van Eck, 2011).

The resulting network (Figure 6) reveals a highly interconnected research landscape centered on four dominant themes. The first cluster primarily represents forest biomass estimation, with strong associations among terms such as aboveground biomass, forest, carbon, estimation, and model. These terms form the core of the network, reflecting the longstanding emphasis on quantifying forest biomass and carbon stocks using remotely sensed data. Their central position indicates that biomass estimation remains the principal application around which methodological developments in remote sensing and machine learning have evolved.

A second cluster is characterized by terms related to remote sensing technologies and Earth observation datasets, including synthetic aperture radar, Sentinel, LiDAR, remote sensing, and SAR. This cluster highlights the increasing integration of complementary remote sensing datasets to overcome the limitations of individual sensors. While optical imagery provides detailed spectral information, SAR contributes structural and moisture-related characteristics under all-weather conditions, and LiDAR supplies three-dimensional measurements of canopy architecture. The close spatial association among these terms demonstrates that multi-source data integration has become a defining feature of contemporary biomass estimation research.

The third cluster reflects the growing influence of machine learning and artificial intelligence, with prominent terms including machine learning, random forest, classification, prediction, and deep learning. The prominence of these terms indicates a clear methodological transition from conventional statistical approaches towards data-driven algorithms capable of modelling complex nonlinear relationships between remotely sensed variables and biomass. The increasing visibility of deep learning within the network further suggests that automated feature extraction and representation learning are becoming important research directions, although traditional ensemble learning methods remain more widely adopted across the reviewed studies.

The fourth cluster encompasses terms associated with environmental applications and ecosystem monitoring, linking biomass estimation with broader themes such as vegetation dynamics, ecosystem assessment, carbon cycling, and sustainable forest management. The connections between this cluster and those representing remote sensing technologies and machine learning demonstrate the increasingly interdisciplinary nature of biomass estimation research, where advances in sensor technology, artificial intelligence, and environmental science are converging to improve forest monitoring and carbon assessment.

Overall, the bibliometric network illustrates a clear evolution in research priorities over the past decade. Earlier studies primarily emphasized individual remote sensing datasets and empirical modelling approaches, whereas more recent investigations increasingly integrate SAR, optical imagery, LiDAR, and advanced machine learning techniques within unified analytical frameworks. At the same time, the comparatively limited prominence of terms related to urban trees, urban forests, and urban carbon indicates that urban applications remain underrepresented within the existing literature. This observation supports the central premise of the present review that, despite significant methodological advances, important knowledge gaps remain in adapting forest-based biomass estimation approaches to complex urban environments.

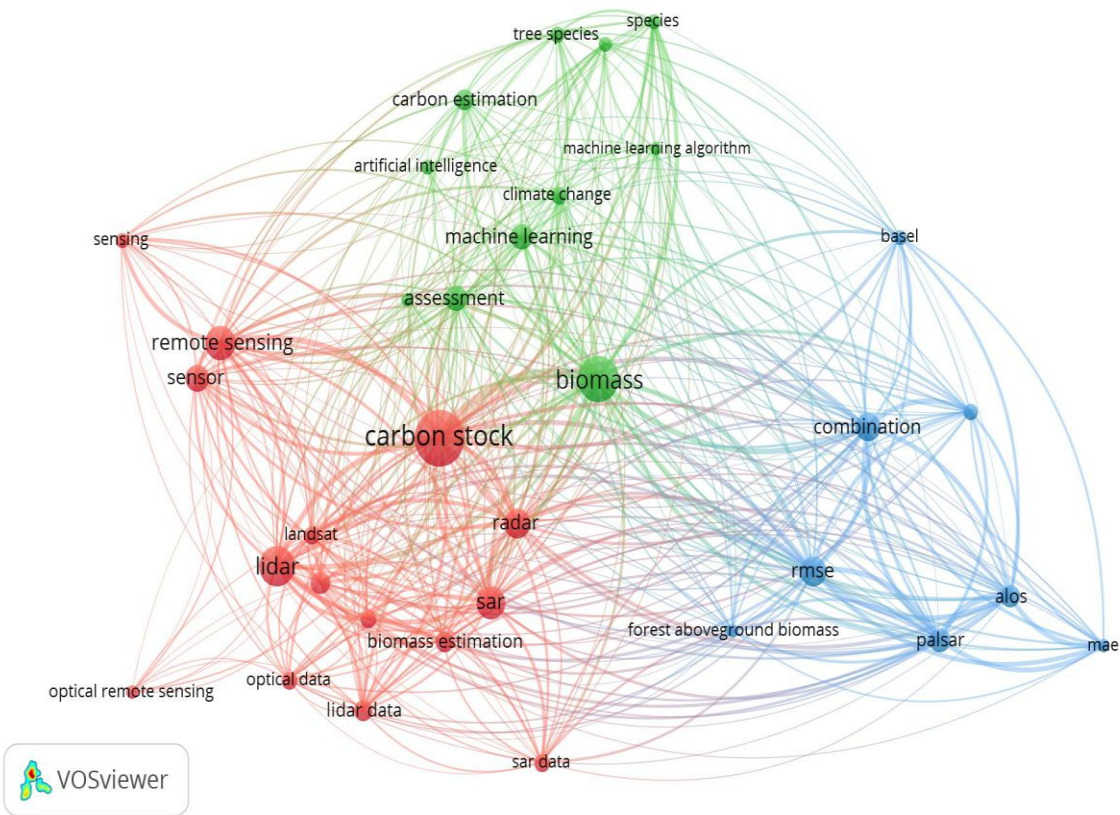


Figure 6. Keyword co-occurrence network of terms extracted from the titles and abstracts of the included studies using VOSviewer.

3.3 Remote Sensing Technologies and Data Fusion

Remote sensing technologies formed the foundation of all studies included in this review, although considerable variation was observed in the types of sensors, data combinations, and integration strategies adopted for biomass and carbon estimation, as shown in Figure 7 below. The reviewed literature demonstrates a clear progression from single-sensor datasets to multi-source remote sensing frameworks that integrate the complementary strengths of optical imagery, synthetic aperture radar (SAR), Light Detection and Ranging (LiDAR), unmanned aerial vehicles (UAVs), and, more recently, street-view imagery. This transition reflects the growing recognition that no single remote sensing technology can comprehensively characterize the spectral, structural, and biophysical properties required for accurate biomass and carbon estimation across diverse forest and urban environments.

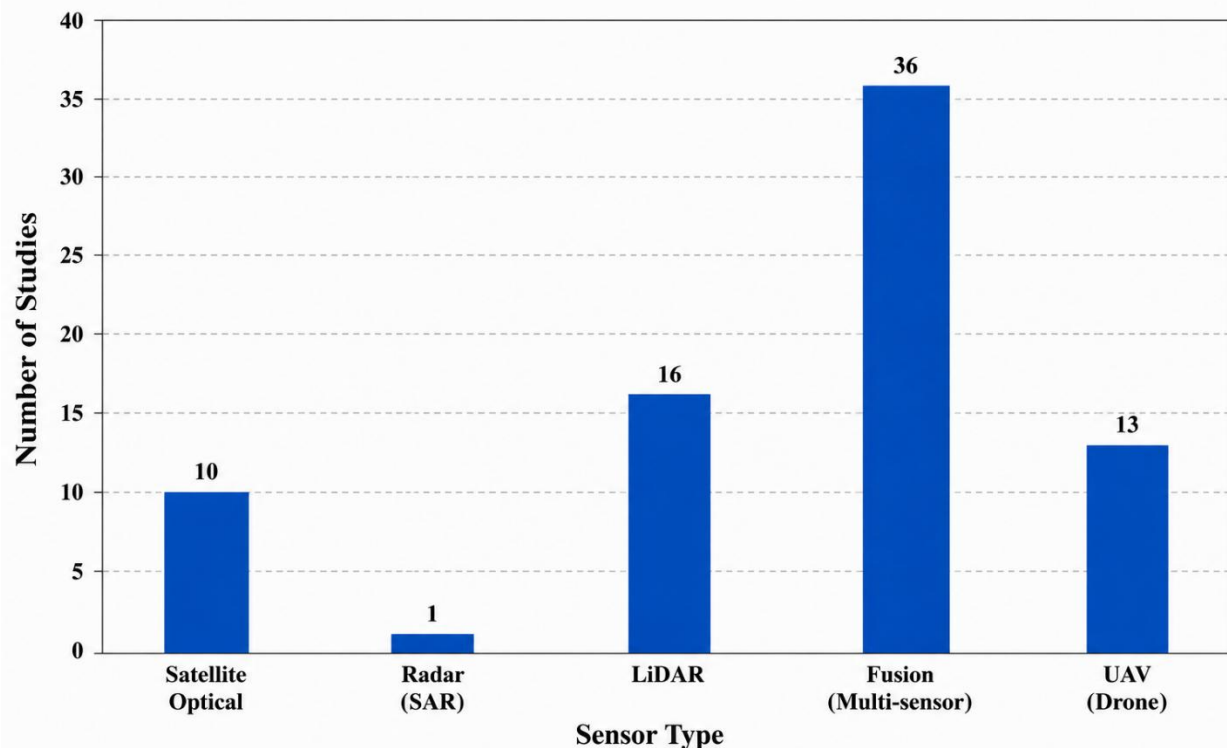


Figure 7. Sensor type and studies reviewed

Optical remote sensing remained the most widely used data source because of its global availability, relatively high spatial resolution, and ability to derive vegetation indices closely associated with canopy condition and photosynthetic activity. Satellite missions such as Landsat and Sentinel-2 were frequently used to generate spectral variables, including the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Normalized Difference Water Index (NDWI), which consistently demonstrated strong relationships with vegetation productivity and biomass. The inclusion of recent studies published in 2025 and 2026 further confirms the continued dominance of optical imagery, particularly when integrated with high-resolution UAV imagery and street-view datasets for urban tree species classification, canopy mapping, and carbon stock assessment. Nevertheless, several studies reported reduced predictive performance when optical imagery was used independently, especially in densely vegetated forests where spectral saturation and persistent cloud cover limited the accurate representation of canopy structure. These limitations have encouraged the growing integration of optical imagery with complementary remote sensing datasets that capture structural information unavailable from spectral measurements alone.

SAR remained the second most frequently applied remote sensing technology among the reviewed studies. Unlike passive optical sensors, SAR actively transmits microwave signals that penetrate cloud cover and operate independently of solar illumination, making the technology particularly valuable for continuous vegetation monitoring under varying weather conditions. The reviewed literature demonstrates that C-band SAR, particularly Sentinel-1, has become the dominant

microwave data source due to its global coverage and open-access availability, while L-band SAR datasets such as ALOS PALSAR continue to provide improved sensitivity to woody biomass owing to their longer wavelengths. However, the review indicates that relatively few recent studies published during 2025 to 2026 explicitly integrated SAR within urban tree carbon estimation frameworks, with most recent investigations favouring LiDAR, UAVs, and optical imagery. Several studies have also reported that SAR backscatter becomes progressively less sensitive as biomass increases, particularly in mature forests, where signal saturation limits discrimination of high-biomass stands. Geometric distortions, including layover and radar shadow over complex terrain and dense urban landscapes, remain important challenges affecting SAR-based biomass estimation.

LiDAR provided unique three-dimensional structural information unavailable from either optical or SAR observations. Measurements of canopy height, crown architecture, and vertical vegetation structure consistently improved biomass estimation, particularly when integrated with spectral and microwave variables. The review provides evidence of increasing adoption of airborne LiDAR (as shown in Figure 8 below), UAV-mounted LiDAR, and terrestrial handheld LiDAR for urban tree inventory, species classification, and crown-level carbon mapping. Several recent studies demonstrated that UAV–LiDAR integration substantially improves the characterization of individual trees and urban canopy structure, while airborne LiDAR remains the dominant platform for large-area biomass assessment. Despite these advantages, LiDAR datasets continue to be used less frequently than optical satellite imagery because of relatively high acquisition costs, limited spatial coverage, and lower temporal availability. Consequently, LiDAR is increasingly employed as a complementary data source within multi-sensor frameworks rather than as a standalone solution.

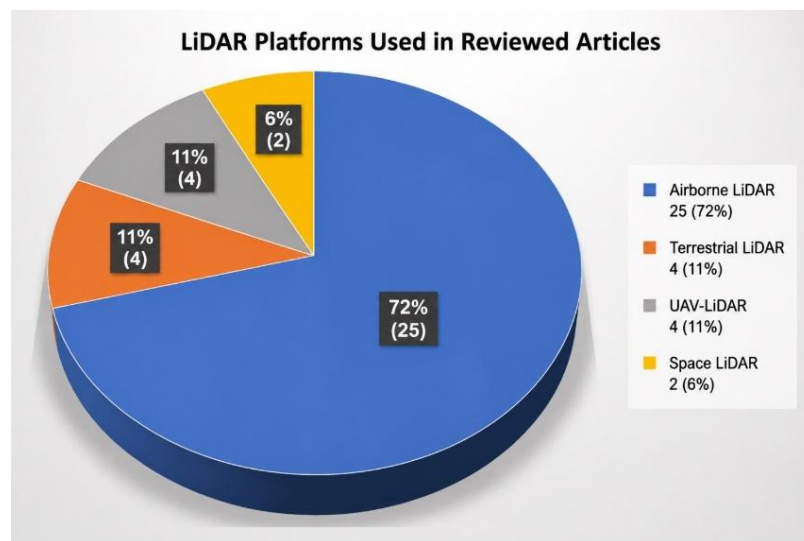


Figure 8. LiDAR platforms

The review indicates that improvements in biomass and carbon estimation have been driven not only by the integration of multiple sensors but also by advances in data fusion strategies. The literature identified several fusion combinations, with LiDAR–optical integration remaining the most frequently adopted approach, followed by SAR–optical, UAV–LiDAR, UAV–SAR, UAV–optical, LiDAR–SAR, and the emerging integration of optical imagery with street-view data for urban tree applications, as shown in Figure 9 and also compared in Table 3. These combinations exploit the complementary strengths of spectral, structural, and contextual information, thereby improving model robustness and prediction accuracy across heterogeneous landscapes. Feature-level fusion remained the predominant strategy because it integrates vegetation indices, SAR backscatter, texture measures, terrain attributes, canopy height metrics, and other derived variables into a unified feature space prior to model development. Pixel-level fusion was used less frequently due to stringent geometric co-registration requirements, whereas decision-level fusion was primarily adopted when predictions from independent sensor-specific models were combined to improve overall estimation performance.

Overall, the reviewed studies demonstrate a clear shift from single-source remote sensing towards integrated multi-source frameworks. Rather than relying on individual sensors, recent research increasingly combines optical imagery, SAR, LiDAR, UAV-derived observations, hyperspectral data, and ancillary datasets to exploit their complementary strengths in representing vegetation spectral responses, canopy structure, and three-dimensional forest architecture. Studies published in 2025 and 2026 further confirm that urban tree research is increasingly adopting advanced data fusion strategies and deep learning techniques. However, recent reviews also highlight that relatively few studies have incorporated SAR into these urban-focused frameworks, indicating that integrating SAR with optical, LiDAR, and other complementary datasets remains an important research opportunity for improving urban tree carbon estimation. This methodological evolution provides a strong foundation for the increasing adoption of advanced machine learning approaches discussed in the following section.

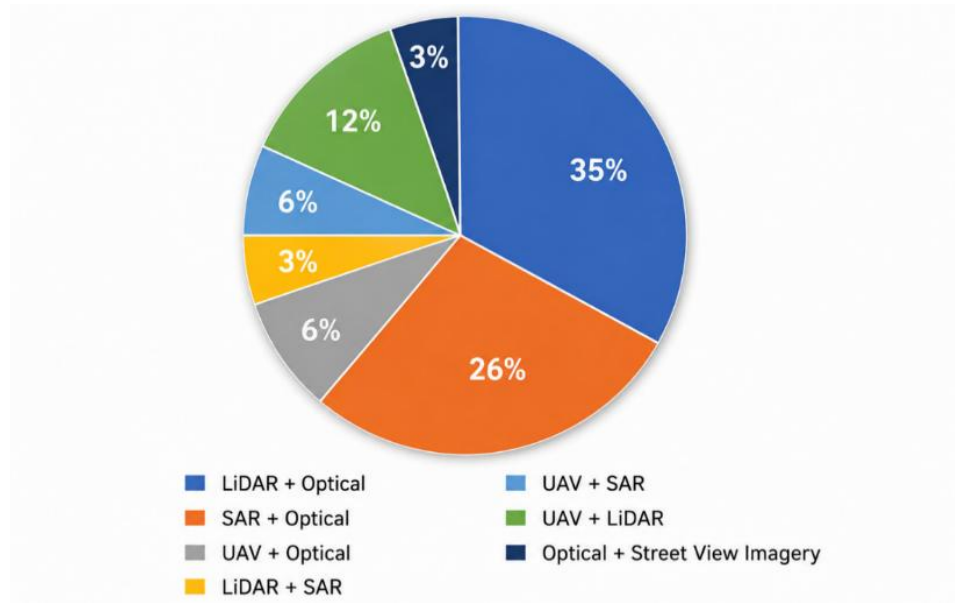


Figure 9. Articles with fusion of data

Table 3

Comparison of remote sensing technologies identified in the reviewed studies

Technology	Strengths	Limitations	Typical Application
Optical	Rich Spectral Information	Cloud Contamination	Vegetation Indices
SAR	All-weather imaging	Signal Saturation	Biomass Estimation
LiDAR	3D canopy structure	Expensive	Forest Inventory
UAV	Very high Resolution	Limited Coverage	Local Mapping

3.4 Machine Learning and Deep Learning Approaches for Biomass and Carbon Estimation

Machine learning has become one of the principal analytical approaches for estimating biomass and carbon from remote sensing data because it can effectively model the complex, nonlinear relationships between remotely sensed variables and vegetation characteristics. The reviewed studies show a diverse range of machine learning and deep learning algorithms, as illustrated in Figure 10, reflecting the increasing availability of multi-source remote sensing datasets and the growing demand for accurate, scalable, and transferable prediction models. Studies published in 2025 and 2026 further demonstrate the rapid evolution of artificial intelligence within remote sensing, particularly for urban tree inventory, species classification, biomass estimation, and

carbon stock mapping. Although several algorithms demonstrated satisfactory predictive performance, their effectiveness varied according to the characteristics of the input data, vegetation structure, spatial scale, and modelling objectives.

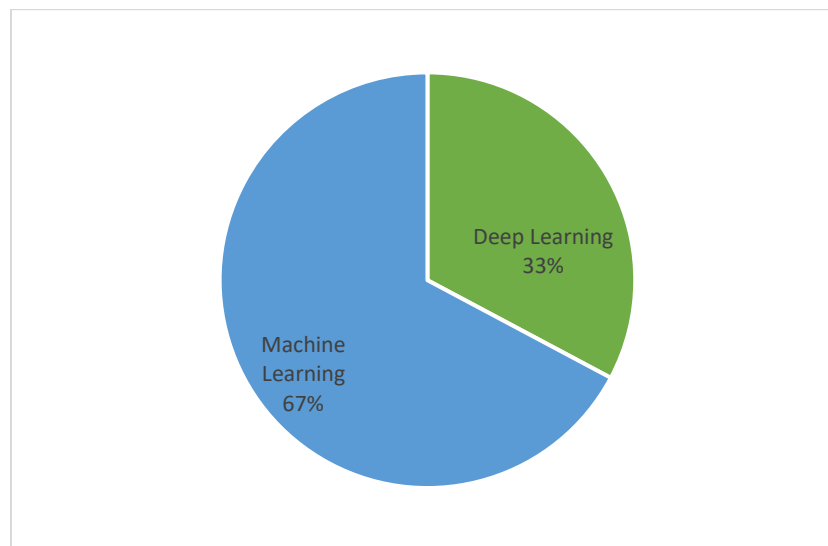


Figure 10. Distribution of machine-learning and deep-learning approaches in the reviewed studies

Random Forest (RF) remained the most frequently applied machine learning algorithm across the reviewed studies as shown in figure 11. Its widespread adoption is largely attributed to its ability to model nonlinear relationships, accommodate high-dimensional datasets, and remain relatively robust to multicollinearity and noisy predictor variables. Unlike many conventional statistical approaches, RF requires relatively little parameter tuning while providing measures of variable importance that facilitate model interpretation. These characteristics make it particularly suitable for integrating heterogeneous variables derived from optical imagery, SAR backscatter, LiDAR metrics, terrain attributes, vegetation indices, and other environmental variables. Consequently, many studies consistently reported that RF achieved high predictive accuracy while maintaining stable performance across diverse forest ecosystems and urban landscapes.

Support Vector Regression (SVR) was another commonly applied algorithm shown in figure 11, particularly in studies involving moderate-sized datasets. SVR effectively models nonlinear relationships using kernel functions and often performs well when training data are limited. Several studies reported predictive accuracies comparable to those obtained using Random Forest, particularly when predictor variables were carefully selected and model parameters were optimized. However, the reviewed literature indicates that SVR becomes increasingly sensitive to parameter selection and computational complexity as the dimensionality and size of remote sensing datasets increase, reducing its suitability for recent large-scale studies employing extensive multi-source datasets.

Recent years have seen increasing adoption of ensemble learning algorithms beyond Random Forest, particularly Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine

(LightGBM), CatBoost, and other boosting frameworks as explain in table 4. These algorithms iteratively minimize prediction errors by sequentially improving weak learners, often producing higher predictive accuracy than single-tree methods. Their ability to capture complex interactions among heterogeneous remote sensing variables has made them increasingly attractive for biomass and carbon estimation, especially where large and diverse predictor datasets are available. Nevertheless, boosting algorithms generally require more extensive parameter tuning than Random Forest and may be more susceptible to overfitting if model complexity is not carefully controlled.

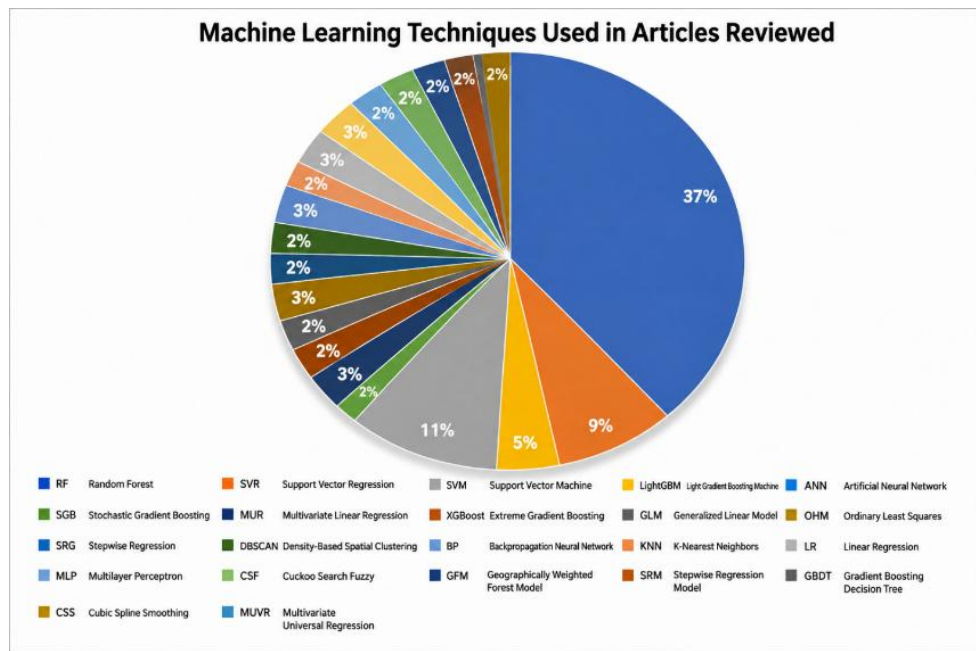


Figure 11. Machine learning techniques

Deep learning methods have become increasingly prominent in the reviewed literature, particularly in studies published in 2025 and 2026. Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), and other deep learning architectures have demonstrated considerable potential for automatically extracting hierarchical features from high-dimensional remote sensing data illustrated in figure 12. Recent studies have introduced advanced object detection and feature extraction models, including YOLOv11, for the identification of individual urban tree species and biomass estimation using UAV–LiDAR data. Similarly, deep learning frameworks have been successfully applied for crown-level urban carbon stock mapping, enabling automated delineation of individual trees within complex urban environments. These developments demonstrate a shift from conventional pixel-based modelling towards object-based and instance-level analysis capable of capturing detailed structural characteristics of urban vegetation.

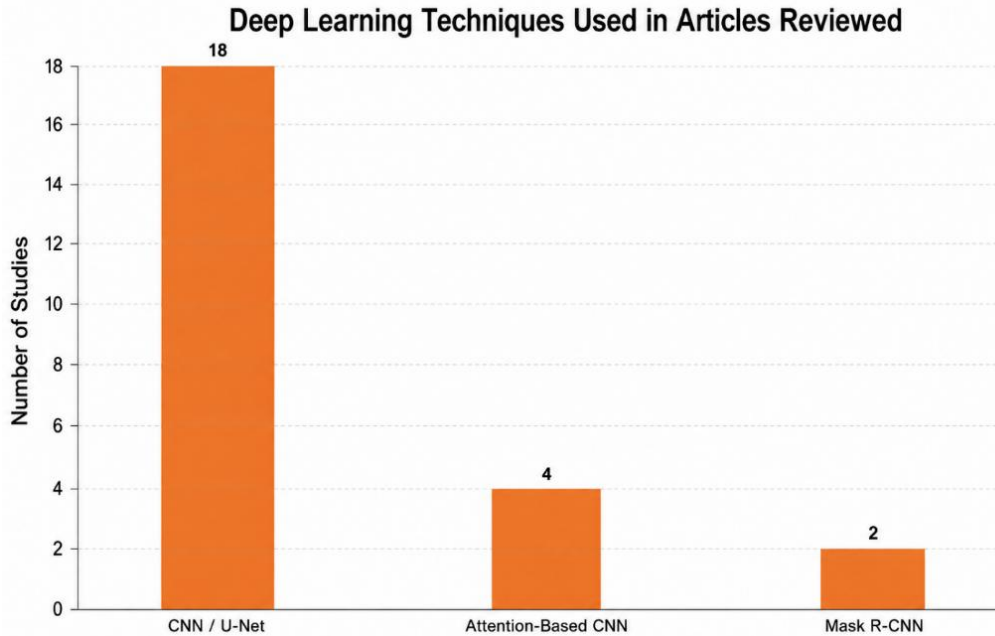


Figure 12. Deep learning techniques

Despite these advances, the reviewed studies consistently reported that deep learning models require substantially greater computational resources, larger training datasets, and careful hyperparameter optimization than conventional machine learning algorithms. Their predictive performance is highly dependent on the availability of extensive labelled datasets, which remain scarce for many urban environments and developing regions. Consequently, although deep learning frequently outperformed traditional machine learning algorithms under data-rich conditions, its operational application remains constrained by computational demands, data availability, and reduced model interpretability.

The review also indicates that machine learning is increasingly being integrated with advanced multi-source data fusion strategies. Rather than relying solely on handcrafted predictor variables, recent studies combined optical imagery, SAR, LiDAR, UAV-derived datasets, hyperspectral imagery, and street-view observations within unified modelling frameworks. Feature-level fusion remained the predominant approach because it enables algorithms such as Random Forest, XGBoost, LightGBM, and deep neural networks to exploit complementary spectral, structural, and spatial information simultaneously. This integration has substantially improved biomass estimation, species classification, and urban carbon stock mapping while enhancing model robustness across heterogeneous landscapes.

Overall, the reviewed literature demonstrates a clear methodological evolution from conventional statistical modelling towards increasingly sophisticated machine learning and deep learning frameworks. While Random Forest continues to provide a favourable balance between predictive accuracy, computational efficiency, and interpretability, recent studies increasingly demonstrate the potential of boosting algorithms and deep learning architectures for extracting complex

relationships from multi-source remote sensing data. Nevertheless, the review highlights that relatively few studies have explored integrating SAR with advanced deep learning models for urban tree carbon estimation. This gap represents an important opportunity for future research and further reinforces the novelty of the present review, which synthesizes current evidence on integrating SAR, multi-source remote sensing, and machine learning to improve urban tree carbon estimation.

Table 4. Comparative characteristics of machine learning and deep learning algorithms used for biomass and carbon estimation.

Algorithm	Strengths	Limitations	Typical Application
Random Forest	Robust to noise and multicollinearity; handles nonlinear relationships; interpretable variable importance	May underestimate extreme values; limited extrapolation	Multi-source remote sensing; biomass estimation
Support Vector Regression	Effective with small datasets; captures nonlinear relationships	Sensitive to kernel selection and parameter tuning; computationally demanding for large datasets	Moderate-sized datasets
XGBoost	High predictive accuracy; models complex interactions	Requires extensive parameter optimization; risk of overfitting	Large multi-source datasets
Artificial Neural Networks	Learns nonlinear relationships	Requires large training datasets; reduced interpretability	Biomass prediction
Convolutional Neural Networks	Automatic spatial feature extraction; effective with imagery	Computationally intensive; large, labelled datasets required	High-resolution optical, SAR and LiDAR imagery
Light Gradient Boosting Machine	Fast training speed; efficient with high-dimensional datasets; reduced memory consumption; excellent scalability	Performance depends on parameter tuning; may overfit small datasets	Large-scale biomass and carbon estimation using multi-source remote sensing
CatBoost	Handles categorical variables efficiently; minimizes overfitting; requires	Longer training time than LightGBM; computationally intensive for very large datasets	Biomass and carbon prediction using heterogeneous environmental datasets

	limited data preprocessing		
YOLOv11	Real-time object detection; highly accurate individual tree detection and species identification; effective with UAV imagery and LiDAR integration	Requires extensive labelled training datasets; high computational requirements; relatively new with limited ecological applications	Individual urban tree detection, species classification, biomass estimation using UAV–LiDAR data
Deep Learning (Hybrid Architectures)	Learns complex spectral–spatial–structural relationships automatically; suitable for multi-source data fusion	Reduced interpretability; high computational cost; performance strongly depends on training data quality	Urban tree carbon mapping, multi-source remote sensing fusion, crown-level carbon estimation

614

615 The preceding sections demonstrate that substantial progress has been achieved through advances
616 in multi-source remote sensing, data fusion strategies, and machine learning algorithms. However,
617 despite these developments, several methodological, computational, and operational limitations
618 continue to constrain the accuracy, scalability, and transferability of biomass and carbon estimation
619 models. These challenges are discussed in the following section.

620 **3.5 Challenges and Research Gaps**

621 The rapid advancement of SAR-based multi-source remote sensing and machine learning has
622 substantially improved biomass and carbon estimation over the past decade. Nevertheless, the
623 reviewed studies reveal several methodological and operational challenges that continue to limit
624 model accuracy, generalizability, and large-scale implementation. These challenges relate not only
625 to the physical characteristics of remote sensing data but also to the transferability of machine
626 learning models across different ecosystems, the availability of high-quality reference data, and
627 the increasing complexity of multi-source data fusion frameworks. Studies published in 2025 and
628 2026 further confirm that, despite significant methodological progress, important knowledge gaps
629 remain, particularly for urban tree carbon estimation.

630 One of the most frequently reported limitations concerns the behaviour of SAR backscatter in
631 dense vegetation. Although microwave signals provide valuable structural information and can
632 penetrate cloud cover, their sensitivity to aboveground biomass decreases progressively in high-
633 biomass forests because of signal saturation. Beyond a certain biomass threshold, increases in
634 vegetation volume produce only small changes in radar backscatter, reducing the ability of SAR
635 observations to discriminate mature forest stands. The magnitude of this limitation varies with
636 radar wavelength, polarization, vegetation structure, and environmental conditions, with longer

wavelengths such as L-band generally exhibiting greater sensitivity than C-band. Consequently, several studies recommended combining SAR observations with optical imagery or LiDAR-derived structural metrics to improve prediction accuracy in densely vegetated environments (Santoro et al., 2024; Tamiminia et al., 2022). However, the review indicates that relatively few recent studies have adopted SAR within urban tree carbon estimation frameworks, despite its clear advantages for all-weather monitoring and structural characterization.

Another important challenge is the occurrence of geometric distortions inherent to side-looking radar systems. Terrain-induced effects such as foreshortening, layover, and radar shadow can significantly influence SAR backscatter, particularly in mountainous terrain and highly built-up urban environments. In cities, buildings often generate strong double-bounce reflections that complicate the separation of tree canopies from surrounding infrastructure, while shadows created by tall structures may obscure vegetation entirely. These effects introduce additional uncertainty into biomass estimation and reduce the transferability of models developed for relatively homogeneous natural forests to heterogeneous urban landscapes. Although terrain correction and advanced preprocessing techniques reduce these distortions, they cannot eliminate them completely, particularly in areas with complex urban morphology.

The review also highlights important challenges associated with multi-source data fusion. Integrating optical imagery, SAR, LiDAR, UAV-derived observations, hyperspectral imagery, street-view imagery, and ancillary geospatial datasets increases the amount of information available for biomass and carbon estimation, but also introduces practical difficulties related to differences in spatial resolution, acquisition dates, sensor calibration, and data compatibility. Accurate geometric co-registration remains essential because even minor positional inconsistencies can propagate through feature extraction and reduce model performance. Furthermore, the increasing dimensionality of fused datasets may introduce redundant variables, increase computational costs, and elevate the risk of model overfitting when effective feature selection is not performed. The studies further demonstrate that although advanced fusion strategies improve prediction accuracy, they also require more sophisticated preprocessing workflows and greater computational resources.

Machine learning algorithms present additional challenges. While Random Forest, XGBoost, LightGBM, and other ensemble learning methods consistently demonstrated strong predictive performance, they often function as black-box models, making it difficult to explain how individual predictor variables influence biomass and carbon estimates. The recent adoption of deep learning approaches, including convolutional neural networks and object detection models such as YOLOv11, has further improved individual tree detection, species classification, and crown-level carbon mapping. However, these methods require extensive labelled datasets, substantial computational resources, and careful hyperparameter optimization. The scarcity of comprehensive field measurements and high-quality urban tree inventories remains a major limitation, particularly in developing countries where reference datasets are often incomplete or unavailable.

Consequently, many models remain geographically specific and require further validation before being transferred to other cities or ecological regions.

One of the most significant findings emerging from the review is the continued imbalance between studies conducted in natural forests and those focusing on urban trees and urban forests shown in figure 13. Although recent studies published in 2025 and 2026, represented in Figure 7, demonstrate growing interest in urban tree inventory, species classification, and crown-level carbon estimation, the overall number of urban-focused investigations remains relatively small compared with those conducted in natural forest ecosystems. Moreover, the screened studies reveal that urban tree carbon estimation continues to rely predominantly on optical imagery, LiDAR, UAV platforms, and deep learning approaches, while the integration of SAR with these datasets remains uncommon. This imbalance highlights an important methodological gap and suggests that SAR's potential to improve urban biomass estimation has yet to be fully explored.

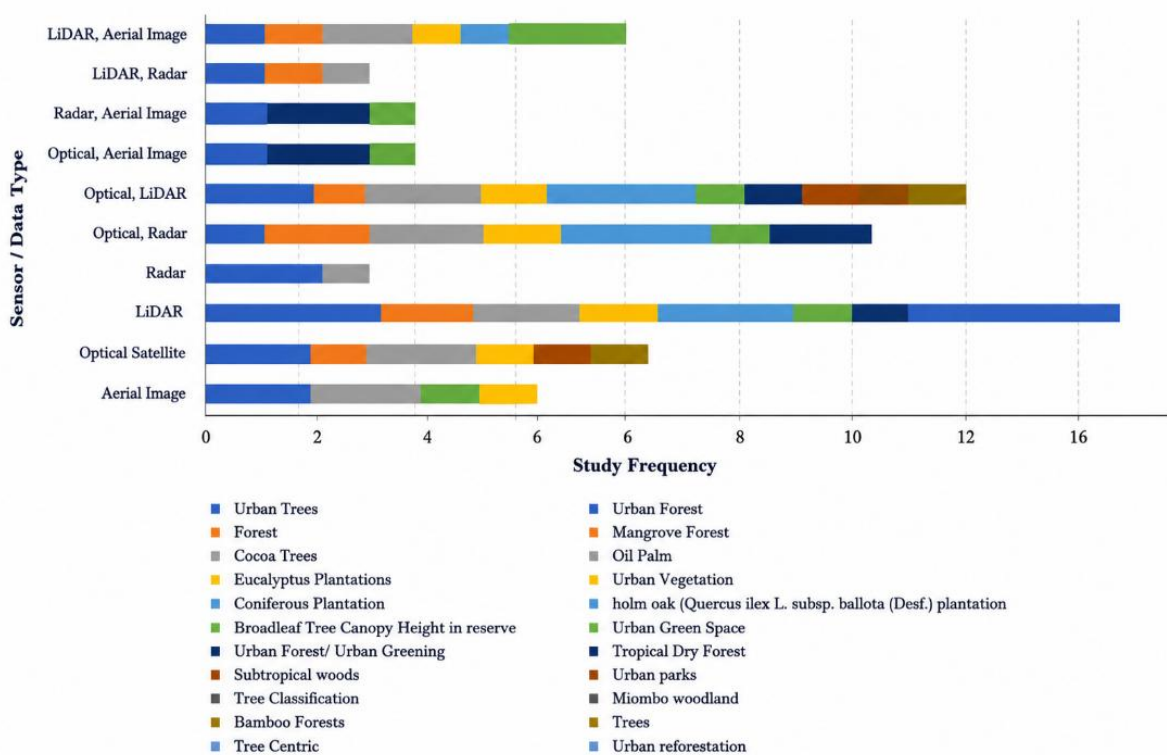


Figure 13. Sensor type and study environment

Another emerging research gap concerns the limited evaluation of model transferability across different geographic regions, climatic conditions, and urban forms. Most studies developed and validated models within a single study area, with relatively few assessing their robustness under contrasting environmental conditions. Standardized benchmark datasets, harmonized validation protocols, and cross-regional model evaluation remain limited, making it difficult to compare algorithms objectively or identify the most transferable modelling approaches. Addressing these

limitations will be essential for developing operational carbon monitoring systems applicable across diverse urban environments.

Looking ahead, several technological developments are likely to shape future research. The increasing availability of next-generation SAR missions, including NISAR and the ESA BIOMASS mission, is expected to improve the structural characterization of vegetation at regional and global scales. Advances in explainable artificial intelligence (XAI), transformer-based deep learning architectures, multimodal learning, foundation models for Earth observation, and cloud-based geospatial computing platforms are also expected to enhance the integration of heterogeneous remote sensing datasets while improving model transparency and scalability. Equally important will be the development of standardized benchmark datasets, rigorous uncertainty quantification frameworks, and transferable machine learning models that support consistent biomass and carbon estimation across natural and urban ecosystems.

Overall, the reviewed evidence indicates that future progress will depend less on introducing individual sensors or algorithms than on developing integrated, transparent, and transferable analytical frameworks that effectively combine complementary remote sensing technologies with robust machine learning methods. The review further demonstrates that, despite rapid advances in deep learning and multi-source data fusion, SAR remains underrepresented in urban tree carbon estimation. Addressing this limitation represents a significant opportunity for future research and reinforces the present review's contribution in identifying methodological gaps and directions for advancing urban tree carbon estimation using SAR, multi-source remote sensing, and machine learning.

4.0 Discussion

This systematic review synthesized evidence from studies published between 2004 and 2026 to evaluate advances in SAR-based remote sensing, multi-source data fusion, and machine learning for biomass and carbon estimation. The results reveal substantial technological and methodological progress, accompanied by important challenges relating to data integration, model transferability, and urban applications. The following sections interpret these findings in relation to previous research, discuss their implications for biomass estimation, and identify priorities for future investigation.

4.1 Evolution of Remote Sensing Technologies for Biomass and Carbon Estimation

The evidence synthesized in this review demonstrates a substantial evolution in the application of remote sensing technologies for biomass and carbon estimation over the past two decades. Early studies relied predominantly on optical satellite imagery because of its widespread availability, established vegetation indices, and relatively mature processing techniques. Optical sensors,

particularly Landsat and Sentinel-2, have provided valuable information on vegetation condition, canopy cover, and photosynthetic activity through spectral indices such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). However, the findings of this review indicate that optical imagery alone is often insufficient for accurate biomass estimation in densely vegetated environments because of spectral saturation, atmospheric interference, and cloud contamination. Similar limitations have been reported in previous reviews, which concluded that optical data become increasingly insensitive to biomass variations beyond moderate canopy densities (Reisi Gahrouei et al., 2024; Wulder et al., 2022). Studies published in 2025 and 2026 further demonstrate that optical imagery remains the dominant data source, particularly when integrated with UAV imagery and street-view observations for urban tree species classification and carbon stock assessment, rather than being used as a standalone dataset.

The increasing integration of synthetic aperture radar (SAR) observed across the reviewed studies reflects the growing recognition that microwave remote sensing provides structural information that complements spectral observations. Unlike passive optical sensors, SAR systems operate independently of daylight and are largely unaffected by cloud cover, making them particularly suitable for vegetation monitoring in regions with frequent cloud cover or prolonged rainy seasons. Several studies included in this review demonstrated that SAR-derived backscatter and texture measures significantly improved biomass prediction when combined with optical variables, particularly in structurally complex forests where optical indices alone became saturated. These findings agree with previous investigations showing that integrating Sentinel-1 with Sentinel-2 improves the representation of vegetation structure and reduces uncertainty in biomass estimation compared with single-sensor approaches (Persson & Fransson, 2017; Wulder et al., 2022). Nevertheless, the evidence from the review reveals that relatively few recent urban tree studies have incorporated SAR into multi-source frameworks. Most studies published during 2025 to 2026 relied primarily on optical imagery, LiDAR, UAV-derived datasets, and deep learning, suggesting that the potential of SAR for urban tree carbon estimation remains underexplored.

LiDAR remains the benchmark technology for capturing three-dimensional vegetation structure because it directly measures canopy height, crown geometry, and vertical vegetation distribution. Consistent with previous studies, this review found that LiDAR-derived metrics generally produced the highest prediction accuracies when available, particularly when incorporated into multi-source modelling frameworks. The review also demonstrates an increasing adoption of airborne LiDAR, UAV-mounted LiDAR, and terrestrial handheld LiDAR for urban tree inventory, species classification, and crown-level carbon mapping. Recent studies have shown that integrating UAV-LiDAR with hyperspectral imagery and deep learning substantially improves individual tree detection, biomass estimation, and urban carbon stock mapping. Despite these advantages, relatively high acquisition costs, limited temporal coverage, and inconsistent data availability continue to constrain the widespread operational use of LiDAR over large geographical areas. Consequently, LiDAR is increasingly employed as a complementary structural data source within integrated remote sensing frameworks rather than as a standalone technology.

An important finding emerging from this review is that technological progress has been accompanied by a fundamental shift in research priorities. Earlier investigations focused primarily on evaluating the performance of individual sensors, whereas recent studies increasingly emphasize integrating heterogeneous datasets within unified analytical frameworks supported by advanced machine learning and deep learning algorithms. This transition has been facilitated by the widespread availability of freely accessible satellite missions, improvements in UAV technology, advances in cloud-based geospatial computing, and the rapid development of artificial intelligence for Earth observation. The studies published during 2025 to 2026 further illustrate this transition by introducing advanced approaches such as UAV–LiDAR synergy, street-view imagery integration, hyperspectral data fusion, and deep learning-based crown-level mapping for urban carbon estimation. These developments demonstrate that biomass and carbon estimation has evolved from a sensor-specific problem to a data integration challenge in which complementary spectral, structural, spatial, and contextual information are simultaneously exploited to improve prediction accuracy.

The review also highlights an important difference between methodological developments in natural forests and urban environments. Although multi-source remote sensing and machine learning have substantially improved biomass estimation across forest ecosystems, relatively few studies have specifically addressed the complexity of urban trees. Urban landscapes are characterized by fragmented vegetation, diverse species composition, impervious surfaces, and heterogeneous canopy structures, presenting challenges not encountered in continuous forest environments. The recent emergence of studies integrating UAV imagery, LiDAR, street-view imagery, and deep learning demonstrates growing interest in addressing these challenges. However, the continued scarcity of studies integrating SAR with these emerging data sources indicates that urban tree carbon estimation remains an evolving research field with considerable scope for methodological innovation. Consequently, future research should prioritize developing robust SAR-enabled multi-source frameworks that leverage the complementary strengths of optical imagery, LiDAR, UAV observations, and machine learning to enable accurate and transferable urban tree carbon estimation across diverse urban environments.

4.2 Critical Comparison of Machine Learning and Deep Learning Algorithms

The findings of this review demonstrate that the increasing adoption of machine learning has fundamentally transformed biomass and carbon estimation by enabling the modelling of complex, nonlinear relationships that are difficult to capture using conventional statistical approaches. Although a wide range of algorithms has been applied, their performance is strongly influenced by the characteristics of the remote sensing data, vegetation complexity, sample size, feature engineering, predictor selection, and model parameterization. The inclusion of studies published during 2025 and 2026 further confirms that algorithm development is evolving alongside advances in multi-source remote sensing, particularly for urban tree inventory, species classification, and carbon stock estimation. Consequently, no single algorithm consistently outperformed all others across all studies, supporting previous observations that model selection should be guided by both

data characteristics and the intended application rather than algorithm popularity alone (Belgiu & Drăguț, 2016; Maxwell et al., 2018).

Among the reviewed studies, Random Forest (RF) remained the most widely adopted algorithm. This finding is consistent with previous reviews that describe RF as one of the most robust ensemble learning techniques for remote sensing applications because of its ability to model nonlinear relationships, accommodate multicollinearity, manage noisy predictor variables, and process high-dimensional datasets without extensive parameter optimization (Belgiu & Drăguț, 2016). The consistent performance of RF observed across the reviewed studies is largely attributable to its ensemble structure, in which multiple decision trees improve model stability, reduce prediction variance, and enhance generalization. Furthermore, the variable importance measures generated by RF facilitate interpretation of the relative contributions of optical, SAR, LiDAR, terrain, and environmental variables. These characteristics explain why RF continues to dominate biomass and carbon estimation studies despite the emergence of more sophisticated machine learning algorithms.

Support Vector Regression (SVR) also demonstrated strong predictive capability, particularly in studies involving relatively small training datasets. Unlike RF, SVR employs kernel functions to transform nonlinear relationships into higher-dimensional feature spaces, enabling complex interactions between remotely sensed variables and biomass to be modelled effectively (Drucker et al., 1996; Mountrakis et al., 2011). Several studies included in this review reported prediction accuracies comparable to those obtained using RF when kernel functions and regularization parameters were carefully optimized. However, SVR proved more sensitive to parameter selection, feature scaling, and computational complexity, particularly when applied to large multi-source datasets containing numerous predictor variables. Consequently, while SVR remains an effective option for moderate-sized datasets, its operational applicability decreases as dataset size and model complexity increase.

The review also highlights the growing importance of gradient boosting algorithms, particularly Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and CatBoost. These algorithms improve predictive performance by sequentially correcting the errors of preceding decision trees, enabling them to capture complex interactions among heterogeneous predictor variables while reducing prediction bias (Chen & Guestrin, 2016). The study indicates that these boosting algorithms achieved predictive accuracies comparable to, and occasionally exceeding, those of Random Forest, particularly when integrating optical imagery, LiDAR metrics, UAV-derived products, terrain variables, and other environmental predictors. Nevertheless, these improvements generally required more extensive hyperparameter optimization and greater computational resources, increasing the risk of overfitting when model complexity was not carefully controlled. Consequently, boosting algorithms appear particularly advantageous for information-rich datasets, whereas Random Forest remains the preferred choice for operational applications because of its balance between predictive accuracy, computational efficiency, and interpretability.

Deep learning approaches have become increasingly prominent within the reviewed literature, reflecting recent advances in artificial intelligence for Earth observation. Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs) have demonstrated considerable potential for automatically extracting hierarchical spectral, spatial, and structural features directly from remote sensing imagery, reducing the reliance on manually engineered variables (Ma et al., 2019). The study further demonstrates the growing adoption of advanced deep learning architectures, including YOLOv11, for individual urban tree detection, species classification, and biomass estimation using UAV–LiDAR data. Similarly, deep learning frameworks have enabled crown-level carbon stock mapping in dense urban environments by automating the delineation of individual trees. These advances represent an important shift from conventional pixel-based modelling towards object-based and instance-level analysis capable of exploiting high-resolution imagery and three-dimensional structural information.

Despite these advances, the reviewed literature consistently indicates that deep learning models require substantially larger labelled datasets, greater computational resources, and more extensive hyperparameter optimization than conventional machine learning algorithms. These requirements remain a significant limitation for biomass and carbon estimation, particularly in urban environments where comprehensive field inventories are often unavailable. In addition, deep learning models generally exhibit lower interpretability than ensemble learning methods, which limits their acceptance in operational environmental monitoring and policy applications, where transparent decision-making is increasingly important.

An important observation emerging from this review is that algorithm performance depends more on the quality and diversity of the available remote sensing data than on the choice of algorithm alone. Random Forest consistently produced stable and reliable predictions when integrating heterogeneous variables derived from optical imagery, SAR, LiDAR, UAV observations, and topographic datasets, whereas SVR performed well with carefully selected predictors and relatively small datasets. XGBoost, LightGBM, and CatBoost demonstrated considerable potential for exploiting complex multi-source datasets, while deep learning approaches generally achieved the highest predictive performance where abundant labelled data and high-performance computing resources were available. However, the review also indicates that relatively few studies have integrated SAR with advanced deep learning architectures for urban tree carbon estimation, suggesting that the complementary strengths of microwave remote sensing and modern artificial intelligence remain insufficiently exploited.

Another notable trend identified in this review is the transition from algorithm comparison towards the development of integrated modelling frameworks that combine multi-source remote sensing, feature-level data fusion, and artificial intelligence within a unified analytical workflow. Rather than seeking a universally superior algorithm, recent studies increasingly focus on optimizing the interplay among sensor characteristics, feature engineering, and learning architectures. The studies published during 2025 and 2026 further reinforce this trend by integrating UAV imagery, LiDAR, hyperspectral data, street-view imagery, and deep learning for individual-tree analysis and urban

carbon estimation. These developments demonstrate that future methodological advances are more likely to arise from improved integration of complementary remote sensing datasets, explainable artificial intelligence, and transferable machine learning frameworks than from incremental improvements to individual algorithms alone. Moreover, the limited incorporation of SAR into these emerging frameworks highlights a clear research opportunity to advance urban tree carbon estimation through more comprehensive fusion of multi-source data.

4.3 Urban Trees Versus Natural Forests: Implications for Model Transferability

One of the clearest patterns emerging from this review is the pronounced imbalance between studies conducted in natural forests and those focusing specifically on urban trees and urban forests. Most of the reviewed studies developed and validated biomass and carbon estimation models within relatively continuous forest ecosystems, where canopy structure, species composition, and land-cover conditions are comparatively homogeneous. In contrast, considerably fewer investigations have focused on urban environments, despite the growing recognition of urban forests as critical components of climate change mitigation, carbon sequestration, biodiversity conservation, and sustainable urban development (IPCC, 2023; Nowak & Greenfield, 2018). This disparity indicates that much of the current methodological evidence has been derived from ecological conditions that differ fundamentally from the heterogeneous landscapes typical of urban areas.

The transferability of machine learning models from natural forests to urban environments remains one of the major challenges identified in the reviewed literature. Forest-based models are generally calibrated using training datasets in which vegetation dominates the remotely sensed signal, allowing relatively direct relationships to be established between sensor-derived variables and biomass. Urban landscapes, however, consist of fragmented tree cover interspersed with buildings, roads, impervious surfaces, and other anthropogenic features that create highly complex spectral, structural, and spatial mixtures (Nowak & Greenfield, 2018; Raciti et al., 2014). Consequently, predictor variables that perform well in continuous forest ecosystems often exhibit reduced predictive capability in urban settings, where tree characteristics are strongly influenced by surrounding infrastructure, management practices, and species diversity (Zhu et al., 2026). Similar observations have been reported by Nowak & Greenfield, (2018) and Raciti et al., (2014), who emphasized that urban vegetation displays greater spatial heterogeneity and lower spectral separability than vegetation in natural forests.

These challenges are particularly evident in SAR-based biomass estimation. Within natural forests, radar backscatter is primarily controlled by vegetation structure, moisture content, and canopy architecture. In urban environments, however, SAR signals are simultaneously influenced by vegetation and complex scattering mechanisms associated with buildings, roads, and other vertical structures. Double-bounce scattering, layover, radar shadow, and variations in building orientation frequently complicate the interpretation of microwave observations and reduce the ability to distinguish tree canopies from surrounding infrastructure (Balzter et al., 2015; Wulder et al., 2022).

As a result, SAR-based relationships developed in forest ecosystems cannot be directly transferred to urban landscapes without careful calibration. This limitation partly explains why many urban tree studies have relied more heavily on optical imagery, LiDAR, UAV observations, and deep learning techniques than on SAR data, despite the well-established advantages of microwave remote sensing for all-weather vegetation monitoring (Ferreira et al., 2024; Gong et al., 2022). The relatively limited integration of SAR within urban tree carbon estimation therefore represents an important opportunity for future methodological development.

Another important distinction concerns the availability and quality of reference data used for model calibration and validation. Forest biomass studies frequently benefit from standardized national forest inventories containing consistent measurements of tree diameter, height, species composition, and stand characteristics (Dubayah et al., 2020). In contrast, urban tree inventories are often incomplete, spatially inconsistent, or developed primarily for municipal management rather than scientific research (Nowak & Greenfield, 2018). Consequently, many urban studies rely on relatively small, geographically constrained reference datasets, thereby limiting model robustness and transferability across cities. This limitation is particularly challenging for deep learning approaches, including convolutional neural networks and recent object detection models, which require extensive labelled datasets to achieve reliable performance (Luo et al., 2025; Ma et al., 2019). The lack of standardized urban benchmark datasets therefore remains a major obstacle to the broader adoption of advanced artificial intelligence techniques for urban tree carbon estimation.

The reviewed literature also demonstrates that urban environments increasingly require modelling approaches specifically designed to capture fine-scale vegetation characteristics. High-resolution satellite imagery, UAV observations, airborne and terrestrial LiDAR, hyperspectral imagery, and street-view imagery have shown considerable potential for representing individual trees and complex urban canopy structure (Velasquez-Camacho et al., 2021; Wang et al., 2025). When combined through feature-level data fusion and advanced machine learning algorithms, these complementary datasets substantially improve species classification, crown delineation, biomass estimation, and carbon stock mapping (Alonzo et al., 2014; Seely et al., 2025). Recent advances in explainable artificial intelligence (XAI), object-based image analysis, and multimodal deep learning further offer opportunities to improve model transparency and to exploit the complementary strengths of diverse remote sensing data sources (Zhu et al., 2026).

From a broader perspective, the limited emphasis on urban trees within the reviewed literature represents both a methodological gap and a significant research opportunity. As cities continue to expand and adopt nature-based solutions to address climate change, accurate quantification of urban tree carbon stocks will become increasingly important for greenhouse gas accounting, climate reporting, ecosystem service assessment, and evidence-based urban planning (IPCC, 2023; Nowak & Greenfield, 2018). Future research should therefore prioritize the development of transferable modelling frameworks that integrate SAR, optical imagery, LiDAR, UAV-derived observations, and advanced machine learning within unified analytical workflows. Equally

important is the need for standardized urban tree reference datasets, harmonized validation protocols, and cross-city model evaluation to improve the reliability and scalability of urban tree carbon estimation across diverse environmental and socio-economic settings (Bressant et al., 2025).

4.4 Future Research Directions and Emerging Opportunities

The findings of this review indicate that future advances in biomass and carbon estimation will depend less on the development of individual sensors or machine learning algorithms and more on integrating complementary technologies into robust, transferable analytical frameworks. Although substantial progress has been achieved through the combination of SAR, optical imagery, LiDAR, and machine learning, several methodological limitations remain unresolved. Addressing these limitations will require improvements in remote sensing data integration, model interpretability, uncertainty quantification, and the availability of representative reference datasets. Consequently, future research should move beyond simply improving prediction accuracy towards developing reliable, transparent, and operational workflows that support long-term monitoring of both natural forests and urban ecosystems (Wulder et al., 2022; Zhu et al., 2026).

One important research priority is the development of generalizable machine learning models that perform consistently across forest types, climatic regions, and urban environments. Most studies have developed and validated models within a single study area, making it difficult to assess their transferability to different geographic regions. Variations in vegetation structure, species composition, environmental conditions, management practices, and sensor characteristics often degrade model performance when algorithms are applied outside their calibration environments. Future investigations should therefore emphasize multi-regional datasets, cross-site validation, and standardized benchmarking protocols to improve model robustness and facilitate operational implementation at regional and national scales (Ma et al., 2019; Maxwell et al., 2018).

Another important research direction is the integration of explainable artificial intelligence (XAI) into biomass and carbon estimation. While ensemble learning and deep learning algorithms often achieve high predictive accuracy, their decision-making processes are difficult to interpret, which limits user confidence in applications such as carbon accounting, climate reporting, and environmental policy. Explainability techniques such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) provide valuable insights into variable importance and model behaviour, enabling researchers to better understand the relationships between remotely sensed variables and biomass estimates. Incorporating these approaches into future modelling frameworks would improve transparency, facilitate model validation, and support more informed scientific and policy decisions (Duncanson et al., 2019; Ullah et al., 2025).

Recent advances in artificial intelligence also present new opportunities for remote sensing applications. Transformer-based architectures and foundation models have demonstrated considerable success in computer vision and are increasingly being adopted for Earth observation

due to their ability to capture long-range spatial relationships and learn generalized feature representations from large, heterogeneous datasets. These models show considerable promise for integrating optical imagery, SAR, LiDAR, and other complementary datasets within unified analytical frameworks, thereby reducing the need for extensive manual feature engineering while improving model scalability and transferability. Nevertheless, their application to biomass and carbon estimation remains in its early stages and is constrained by substantial computational requirements and the need for large, well-annotated training datasets (Yun et al., 2024).

Future progress will also be supported by advances in Earth observation systems. The forthcoming NASA-ISRO Synthetic Aperture Radar (NISAR) mission and the European Space Agency BIOMASS mission are expected to provide unprecedented information on vegetation structure through enhanced radar capabilities and near-global coverage. These missions will complement existing observations from Sentinel-1, Sentinel-2, Landsat, and GEDI, creating new opportunities for multi-source data fusion, improved structural characterization of vegetation, and more consistent biomass estimation across diverse ecosystems (Dubayah et al., 2020; Santoro et al., 2024). Their integration with advanced machine learning algorithms is expected to strengthen regional and global carbon monitoring programmes and improve support for climate mitigation initiatives.

The reviewed literature also demonstrates that urban tree carbon estimation remains comparatively underrepresented relative to studies conducted in natural forest ecosystems. Given the increasing importance of urban forests for climate adaptation, ecosystem service provision, biodiversity conservation, and sustainable city planning, future research should prioritize modelling frameworks specifically designed for heterogeneous urban environments. High-resolution optical imagery, airborne and terrestrial LiDAR, UAV observations, object-based image analysis, and feature-level data fusion offer considerable potential for improving the representation of individual trees and small urban forest patches. Greater emphasis should also be placed on integrating SAR within these multi-source frameworks to exploit its all-weather imaging capability and sensitivity to vegetation structure, thereby improving the robustness of urban tree carbon estimation.

Finally, greater attention should be given to uncertainty quantification, reproducibility, and open science practices. Although many studies report conventional performance metrics such as the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE), comparatively few explicitly quantify prediction uncertainty or evaluate model transferability using independent datasets. Future studies should incorporate uncertainty analysis, standardized benchmark datasets, open-access reference data, and reproducible analytical workflows to facilitate objective comparisons among competing approaches and improve confidence in biomass and carbon estimates used for environmental management, climate policy, and national greenhouse gas accounting (Shane et al., 2026; Ullah et al., 2025).

4.5 Strengths and Limitations of the Review

This systematic review provides a comprehensive synthesis of research published between 2004 and 2026 on the integration of synthetic aperture radar (SAR), multi-source remote sensing, and machine learning techniques for biomass and carbon estimation. By following a transparent and reproducible review protocol based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, employing Rayyan for article screening and duplicate removal, and applying predefined inclusion and exclusion criteria, the review minimizes selection bias and enhances methodological transparency (Ouzzani et al., 2016; Page et al., 2021). Furthermore, integrating bibliometric analysis with a critical synthesis of remote sensing technologies, data fusion strategies, and machine learning algorithms provides a broader perspective than conventional narrative reviews, enabling both quantitative and qualitative assessments of developments in the field.

Another strength of this review lies in its emphasis on urban tree carbon estimation, an area that has received comparatively limited attention within the existing literature. While previous reviews have predominantly focused on natural forests or generalized biomass estimation, the present review critically evaluates the applicability of existing methodologies to heterogeneous urban environments. By identifying differences between natural forests and urban tree applications, this review highlights key methodological gaps and provides a framework for future investigations to improve urban carbon monitoring by integrating remote sensing and artificial intelligence techniques.

Despite these strengths, several limitations should be acknowledged. First, the review was restricted to peer-reviewed publications indexed in Scopus, Google Scholar, and Web of Science, and only English-language studies were considered. Consequently, relevant research available in other bibliographic databases or published in other languages may not have been captured. Second, although systematic search procedures and predefined eligibility criteria were applied, the synthesis remains dependent on the methodological quality and reporting standards of the included studies. Variations in study design, field sampling protocols, remote sensing datasets, feature engineering approaches, and accuracy assessment methods limited direct quantitative comparison among studies and precluded formal meta-analysis.

Another limitation relates to the rapid evolution of remote sensing technologies and machine learning methods. The review covers literature published up to 2026, providing a comprehensive assessment of the evidence available during the review period. Nevertheless, new satellite missions, deep learning architectures, and explainable artificial intelligence techniques continue to emerge, meaning that future systematic reviews will be required to capture ongoing methodological developments. This limitation is inherent to systematic reviews in rapidly evolving research fields and does not diminish the relevance of the present synthesis.

Overall, the strengths of this review outweigh its limitations. The systematic methodology, critical evaluation of emerging technologies, and emphasis on urban tree applications provide a robust foundation for understanding current research trends while identifying priority areas for future

investigation. Taken together, the findings demonstrate that integrating SAR, complementary remote sensing datasets, and advanced machine learning provides a promising pathway towards more accurate, scalable, and transferable biomass and carbon estimation. At the same time, addressing the methodological gaps identified in this review, particularly those relating to urban environments, model interpretability, and standardized validation, will be essential for supporting reliable carbon accounting and climate-informed environmental management

5.0 Conclusion

This systematic review demonstrates that aboveground biomass and carbon stock estimation has evolved from predominantly single-sensor approaches to integrated frameworks that combine SAR, optical imagery, LiDAR, UAV observations, and machine learning. The reviewed evidence shows that these data sources provide complementary information on vegetation spectral characteristics, canopy structure, and biomass-related properties, with multi-source and feature-level fusion offering important advantages over reliance on individual sensors. Random Forest remains one of the most widely applied and robust machine learning approaches because of its ability to handle nonlinear relationships and heterogeneous predictors, while XGBoost, artificial neural networks, convolutional neural networks, and other deep learning methods are increasingly extending the analytical capabilities available for biomass and carbon estimation.

Despite these advances, important challenges remain. SAR signal saturation in high-biomass environments, geometric distortions such as layover and shadow, differences in sensor resolution and acquisition timing, limited field reference data, and uncertainty in model transferability continue to constrain reliable large-scale implementation. These challenges are particularly important in urban environments, where fragmented tree cover, buildings, impervious surfaces, and complex scattering mechanisms make relationships established in continuous natural forests difficult to transfer directly. The comparatively limited representation of urban trees in the reviewed literature, therefore, identifies a significant research gap and demonstrates the need for models specifically developed and independently validated under heterogeneous urban conditions.

Future progress will depend on moving beyond improvements in individual sensors or algorithms towards transparent, transferable, and reproducible multi-source modelling frameworks. The growing availability of advanced Earth observation data, together with explainable artificial intelligence, multimodal learning, improved uncertainty quantification, and standardized reference datasets, provides opportunities to strengthen biomass and carbon monitoring across both natural and urban ecosystems. Greater emphasis on cross-region validation and harmonized urban tree datasets will be particularly important for translating methodological advances into operational applications. Ultimately, integrating complementary remote sensing observations with robust and interpretable machine learning offers a practical pathway towards more reliable carbon assessment and can strengthen the evidence required for forest management, urban planning, climate mitigation, and carbon accounting.

References

- 1115 Alonzo, M., Bookhagen, B., & Roberts, D. A. (2014). Urban tree species mapping using
1116 hyperspectral and lidar data fusion. *Remote Sensing of Environment*, 148, 70–83.
- 1117 Balzter, H., Cole, B., Thiel, C., & Schmullius, C. (2015). Mapping CORINE land cover from
1118 Sentinel-1A SAR and SRTM digital elevation model data using random forests. *Remote*
1119 *Sensing*, 7(11), 14876–14898.
- 1120 Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and
1121 future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31.
- 1122 Bressant, C., Wenger, R., Michéa, D., & Puissant, A. (2025). Urban tree species benchmark
1123 dataset for time series classification. *Data in Brief*, 61, 111777.
- 1124 Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the*
1125 *22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*,
1126 785–794.
- 1127 Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector
1128 regression machines. *Advances in Neural Information Processing Systems*, 9.
- 1129 Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G.,
1130 Kellner, J., & Luthcke, S. (2020). The Global Ecosystem Dynamics Investigation: High-
1131 resolution laser ranging of the Earth’s forests and topography. *Science of Remote Sensing*, 1,
1132 100002.
- 1133 Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave,
1134 J., Herold, M., & Crowther, T. W. (2019). The importance of consistent global forest
1135 aboveground biomass product validation. *Surveys in Geophysics*, 40(4), 979–999.
- 1136 Ferreira, M. P., dos Santos, D. R., Ferrari, F., Coelho, L. C. T., Martins, G. B., & Feitosa, R. Q.
1137 (2024). Improving urban tree species classification by deep-learning based fusion of digital
1138 aerial images and LiDAR. *Urban Forestry & Urban Greening*, 94, 128240.
- 1139 Gibbs, H. K., Brown, S., Niles, J. O., & Foley, J. A. (2007). Monitoring and estimating tropical
1140 forest carbon stocks: making REDD a reality. *Environmental Research Letters*, 2(4),
1141 045023.
- 1142 Gong, Y., Li, X., Du, H., Zhou, G., Mao, F., Zhou, L., Zhang, B., Xuan, J., & Zhu, D. (2022).
1143 Tree species classifications of urban forests using UAV-LiDAR intensity frequency data.
1144 *Remote Sensing*, 15(1), 110.
- 1145 Intergovernmental Panel on Climate Change. (2023). *Climate Change 2023: Synthesis Report*.
- 1146 IPCC: Intergovernmental Panel on Climate Change. (2023). *Climate Change 2023: Synthesis*
1147 *Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the*
1148 *Intergovernmental Panel on Climate Change*.

- 1149 Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R.,
1150 Kuemmerle, T., Meyfroidt, P., & Mitchard, E. T. A. (2016). A review of the application of
1151 optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote*
1152 *Sensing*, 8(1), 70.
- 1153 Livesley, S. J., Escobedo, F. J., & Morgenroth, J. (2016). The biodiversity of urban and peri-
1154 urban forests and the diverse ecosystem services they provide as socio-ecological systems.
1155 In *Forests* (Vol. 7, Number 12, p. 291). MDPI.
- 1156 Luo, P., Zhang, Y., Ruan, J., Zhang, G., Tan, J., Wang, Q., & Shang, K. (2025). Estimation of
1157 individual tree biomass for three tree species using LiDAR and multispectral data in
1158 megacity Shanghai. *Frontiers in Remote Sensing*, 6, 1697927.
- 1159 Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote
1160 sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and*
1161 *Remote Sensing*, 152, 166–177.
- 1162 Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning
1163 classification in remote sensing: An applied review. *International Journal of Remote*
1164 *Sensing*, 39(9), 2784–2817.
- 1165 Mitchard, E. T. A. (2018). The tropical forest carbon cycle and climate change. *Nature*,
1166 559(7715), 527–534.
- 1167 Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A
1168 review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259.
- 1169 Nowak, D. J., & Greenfield, E. J. (2018). US urban forest statistics, values, and projections.
1170 *Journal of Forestry*, 116(2), 164–177.
- 1171 Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and
1172 mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210.
- 1173 Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D.,
1174 Shamseer, L., Tetzlaff, J. M., & Moher, D. (2021). Updating guidance for reporting
1175 systematic reviews: development of the PRISMA 2020 statement. *Journal of Clinical*
1176 *Epidemiology*, 134, 103–112.
- 1177 Persson, H. J., & Fransson, J. E. S. (2017). Comparison between TanDEM-X-and ALS-based
1178 estimation of aboveground biomass and tree height in boreal forests. *Scandinavian Journal*
1179 *of Forest Research*, 32(4), 306–319.
- 1180 Picard, N., Saint-André, L., & Henry, M. (2012). *Manual for building tree volume and biomass*
1181 *allometric equations: from field measurement to prediction*.

- 1182 Raciti, S. M., Hutyra, L. R., & Newell, J. D. (2014). Mapping carbon storage in urban trees with
1183 multi-source remote sensing data: Relationships between biomass, land use, and
1184 demographics in Boston neighborhoods. *Science of the Total Environment*, 500, 72–83.
- 1185 Reisi Gahrouei, O., Côté, J.-F., Bournival, P., Giguère, P., & Béland, M. (2024). Comparison of
1186 deep and machine learning approaches for quebec tree species classification using a
1187 combination of multispectral and liDAR data. *Canadian Journal of Remote Sensing*, 50(1),
1188 2359433.
- 1189 Santoro, M., Cartus, O., Quegan, S., Kay, H., Lucas, R. M., Araza, A., Herold, M., Labrière, N.,
1190 Chave, J., & Rosenqvist, Å. (2024). Design and performance of the Climate Change
1191 Initiative Biomass global retrieval algorithm. *Science of Remote Sensing*, 10, 100169.
- 1192 Seely, H., Coops, N. C., White, J. C., Montwé, D., & Ragab, A. (2025). Forest aboveground
1193 biomass estimation using deep learning data fusion of ALS, multispectral, and topographic
1194 data. *International Journal of Remote Sensing*, 46(10), 3874–3912.
- 1195 Shane, T. L., Waaswa, A., Williams, P. J., Reeves, M. C., Washington-Allen, R. A., & Perryman,
1196 B. L. (2026). Estimating the Aboveground Biomass of Shrubland and Savanna Ecosystems
1197 Using High-Resolution Small UAV Systems: A Systematic Review. *Remote Sensing*, 18(6),
1198 942.
- 1199 Sinha, S., Jeganathan, C., Sharma, L. K., & Nathawat, M. S. (2015). A review of radar remote
1200 sensing for biomass estimation. *International Journal of Environmental Science and*
1201 *Technology*, 12(5), 1779–1792.
- 1202 Tamiminia, H., Salehi, B., Mahdianpari, M., Beier, C. M., & Johnson, L. (2022). Evaluating
1203 pixel-based and object-based approaches for forest above-ground biomass estimation using
1204 a combination of optical, Sar, and an extreme gradient boosting model. *ISPRS Annals of the*
1205 *Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 485–492.
- 1206 Ullah, S., Nazeer, M., Wong, M. S., & Amin, G. (2025). Remote sensing for aboveground
1207 biomass monitoring in terrestrial ecosystems: A systematic review. *Remote Sensing*
1208 *Applications: Society and Environment*, 39, 101635.
- 1209 van Eck, N. J. (2011). *Methodological advances in bibliometric mapping of science* (Number
1210 EPS-2011-247-LIS).
- 1211 Velasquez-Camacho, L., Cardil, A., Mohan, M., Etxegarai, M., Anzaldi, G., & de-Miguel, S.
1212 (2021). Remotely sensed tree characterization in urban areas: a review. *Remote Sensing*,
1213 13(23), 4889.
- 1214 Wang, S., Liu, C., Wei, S., & Tang, H. (2025). Urban tree species classification using multisource
1215 satellite remote sensing data and street view imagery. *Geo-Spatial Information Science*,
1216 28(1), 164–184.

- 1217 Wulder, M. A., Roy, D. P., Radeloff, V. C., Loveland, T. R., Anderson, M. C., Johnson, D. M.,
1218 Healey, S., Zhu, Z., Scambos, T. A., & Pahlevan, N. (2022). Fifty years of Landsat science
1219 and impacts. *Remote Sensing of Environment*, 280, 113195.
- 1220 Yun, T., Li, J., Ma, L., Zhou, J., Wang, R., Eichhorn, M. P., & Zhang, H. (2024). Status,
1221 advancements and prospects of deep learning methods applied in forest studies.
1222 *International Journal of Applied Earth Observation and Geoinformation*, 131, 103938.
- 1223 Zhu, X. X., Xiong, Z., Wang, Y., Stewart, A. J., Heidler, K., Wang, Y., Yuan, Z., Dujardin, T., Xu,
1224 Q., & Shi, Y. (2026). On the foundations of Earth foundation models. *Communications*
1225 *Earth & Environment*, 7(1), 103.

1226

1227

1228 **Funding Declaration**

1229 This work was supported by the University of Pretoria (UP DOCTORAL FUNDING)

1230 **Competing Interest**

1231 The authors declare that they have no known competing financial or personal relationships that
1232 may appear to influence the work reported in this paper

1233 **Authors Contributions**

1234 Temitope Emmanuel Paul: Conceptualization; Methodology, Systematic review design, Data
1235 curation, formal analysis, visualization, and original draft writing.

1236 Adedayo Adeleke: Supervisor; Methodology, investigation, review, and editing, and project
1237 administration

1238 Abel Ramoelo: Supervisor: Review and editing, validation, and project administration.

1239 **Data Availability Statement**

1240 All data supporting the findings of this study are available within the article, and additional
1241 materials, including the data extraction sheets and analysis scripts used in the study, will be made
1242 available on reasonable request.

1243 **Clinical trial number:** not applicable

1244 **Consent to Publish:** Not applicable

1245 **Ethics Approval and Consent to Participate:** Not applicable