A SYSTEMATIC REVIEW OVER MACHINE AND DEEP LEARNING APPLICATIONS ON REMOTE SENSING FOREST CARBON MONITORING

Alessandra de Oliveira Alves Correia
Universidade do Oeste Paulista – UNOESTE
Vagner Souza Machado
Universidade do Oeste Paulista - UNOESTE
Ana Paula Marques Ramos
Universidade Estadual Paulista - UNESP
Lucas Prado Osco
Universidade do Oeste Paulista – UNOESTE

ABSTRACT

Objective: Forest carbon monitoring has become critical for climate mitigation under international agreements, driving demand for accurate remote sensing-based assessment methods. This systematic review analyzes machine learning and deep learning applications to forest carbon monitoring using remote sensing data. Following systematic methodology, 59 studies were selected from Scopus, Web of Science, Google Scholar, and SciELO databases using structured Boolean search strategies. Results reveal exponential publication growth from 2019 onwards, correlating with the Paris Agreement implementation. Geographic distribution shows concentration in Asia, Brazil, and North America. Random Forest algorithms dominate all sensor types, followed by Support Vector Machine approaches. Local-scale studies using multispectral sensors represent the most frequent combination, reflecting validation requirements against forest inventories. Critical limitations include limited open-access carbon forest datasets, with only two freely available platforms identified. The trade-off between data accessibility and spatial resolution constrains operational implementation, while local-scale focus highlights challenges in scaling to global monitoring requirements. Machine learning techniques, particularly Random Forest, have matured into robust methodologies for forest carbon assessment, yet significant gaps remain in data standardization and global-scale application. Enhanced international collaboration and standardized data products are essential for bridging local research capabilities with global climate monitoring requirements.

Keyword: literature review, carbon monitoring, machine learning, deep learning.

INTRODUCTION

Accurate forest carbon monitoring has become essential for climate mitigation strategies, as forests contribute approximately 11% of global CO_2 emissions through deforestation while simultaneously acting as important carbon sinks, capable of storing 861 ± 66 GtC globally. Traditional ground-based monitoring approaches face significant limitations in spatial coverage, temporal frequency, and cost-effectiveness, particularly across large forest areas where consistent monitoring is crucial for REDD+ implementation and carbon market verification. Remote sensing technologies provide the spatial and temporal coverage necessary for operational forest carbon monitoring, but translating satellite observations into accurate carbon estimates remains a technical challenge due to the complex relationships between spectral signatures and forest structural parameters, often bringing complexity and lack of detail to the GHG monitoring system, which can lead to confusion about the role of forests in climate change and discourage the achievement of global climate targets (HARRIS et al., 2021).

Thus, machine learning approaches have emerged as powerful tools for addressing these remote sensing challenges by enabling automated pattern recognition and prediction from multi-dimensional satellite data. Traditional machine learning algorithms, including Random Forest, Support Vector Machines, and linear regression, have demonstrated capability in forest parameter estimation by learning relationships between remote sensing variables and field measurements. However, these shallow learning approaches often struggle with the complex, non-linear relationships inherent in forest ecosystems, particularly when integrating multi-sensor data streams or scaling across diverse forest types and environmental conditions, as well as in the work of Tu (2025).

Deep learning techniques, utilizing multiple hidden layers in neural network architectures, offer enhanced capability for hierarchical feature extraction and complex pattern recognition in remote sensing applications. Convolutional Neural Networks (CNNs) have shown particular promise for forest carbon monitoring by automatically extracting spatial features from optical and radar imagery, while recent advances in Vision Transformers and multi-sensor fusion approaches demonstrate improved accuracy in aboveground biomass estimation. These deep learning methods can integrate diverse data sources—including optical imagery from Sentinel-2 and Landsat, synthetic aperture radar from Sentinel-1, and LiDAR measurements from GEDI—to capture the multi-dimensional nature of forest structure and its relationship to carbon storage across different forest types and environmental gradients.

Despite these technological advances, significant research gaps remain in operational forest carbon monitoring using AI-driven remote sensing approaches. Current methods primarily focus on detecting complete forest loss rather than subtle degradation processes, belowground carbon quantification remains largely unaddressed, and model transferability across different biomes and geographic regions presents ongoing challenges. This paper presents a comprehensive review of state-of-the-art remote sensing approaches for forest carbon monitoring, synthesizing recent advances in machine learning and deep learning techniques while identifying critical research directions for improving accuracy, operational scalability, and global applicability of AI-driven forest carbon assessment systems.

METHODS

This systematic literature review was conducted following established guidelines for comprehensive evidence synthesis in remote sensing research. Four complementary databases were systematically searched to ensure comprehensive coverage of the interdisciplinary literature spanning remote sensing, forest science, and machine learning domains: Scopus (comprehensive science database with strong remote sensing coverage), Web of Science Core Collection (citation-indexed scientific literature), Google Scholar (broad academic coverage including conference proceedings and technical reports), and SciELO Brazil (regional focus on South American forest research).

The search strategy employed a structured Boolean approach designed to capture the intersection of three primary concept domains: forest carbon monitoring, remote sensing technologies, and artificial intelligence methodologies. The complete search string was formulated as follows:

("Forest Carbon" OR "Vegetation Carbon" OR "Forest Carbon Monitoring" OR "Forest Carbon Credit" OR "Forest Carbon Balance" OR "Forest Carbon Compensation" OR "Decarbonation") AND ("Remote Sensing" OR "Orbital Image" OR "Aerial Image" OR "Satellite Image" OR "Hyperspectral Image" OR "Multispectral Image") AND ("Machine Learning" OR "Random Forest" OR "Decision Tree" OR "Support Vector Machine" OR "Multilayer Perceptron" OR "K-Nearest Neighbour" OR "Artificial Neural Network" OR "Deep Learning" OR "Deep Neural Network" OR "Convolutional Neural

Network" OR "Recurrent Neural Network" OR "Long-Short Term Memory" OR "Generative Adversarial Network" OR "Vision Transformer")

This search strategy was designed to maximize sensitivity while maintaining relevance, incorporating both traditional machine learning algorithms (Random Forest, Support Vector Machine, Decision Tree) and advanced deep learning architectures (CNNs, RNNs, LSTMs, GANs, Vision Transformers) that have emerged as prominent techniques in recent forest carbon monitoring applications.

Inclusion criteria encompassed: (1) peer-reviewed journal articles and conference proceedings published in English; (2) studies demonstrating direct application of machine learning or deep learning techniques to forest carbon estimation, biomass quantification, or carbon stock assessment using remote sensing data; (3) research incorporating satellite, airborne, or drone-based remote sensing platforms; (4) publications providing quantitative results or methodological frameworks applicable to forest carbon monitoring; and (5) studies published between 2010-2024 to capture the evolution of AI applications in this domain.

Exclusion criteria included: (1) studies focusing solely on agricultural crops or non-forest vegetation without forest carbon applications; (2) theoretical machine learning papers without remote sensing implementation; (3) remote sensing studies without AI/ML components; (4) review articles without original research contributions; (5) studies exclusively using ground-based measurements without remote sensing integration; and (6) publications not available in full text or with insufficient methodological detail for evaluation.

The systematic screening process followed a three-stage approach. Initial screening involved automated duplicate removal across databases and title-based relevance filtering. Abstract screening was conducted by two independent reviewers using the predetermined inclusion/exclusion criteria, with disagreements resolved through consensus discussion. Full-text screening evaluated methodological rigor, data quality, and alignment with review objectives.

After applying the systematic inclusion and exclusion criteria, 59 articles were selected that demonstrated clear alignment with the thematic scope and research objectives of this review. The selected studies span diverse forest ecosystems, remote sensing platforms, and AI methodologies, providing comprehensive coverage of current approaches to forest carbon monitoring using machine learning and deep learning techniques.

To identify dominant research themes and methodological trends within the selected literature, a comprehensive bibliometric analysis was conducted using automated text processing techniques. Key terms were extracted from titles, abstracts, and keywords across all 59 selected articles, followed by frequency analysis and term co-occurrence mapping.

From this compilation, a word cloud visualization was generated to illustrate the relative frequency and prominence of key terminology within the research corpus. As shown in Figure 1, terms such as "forest" and "biomass" emerged as the most frequent concepts, reflecting the primary focus on forest ecosystem carbon quantification. The visualization reveals the dominance of core remote sensing terminology ("satellite," "LiDAR," "optical"), machine learning methodologies ("random forest," "neural network," "deep learning"), and forest carbon parameters ("biomass," "carbon stock," "AGB"), confirming the comprehensive coverage of relevant literature and the coherent thematic focus of the selected studies.

Figure 1. Word cloud highlighting key terms from studies on carbon monitoring using remote sensing

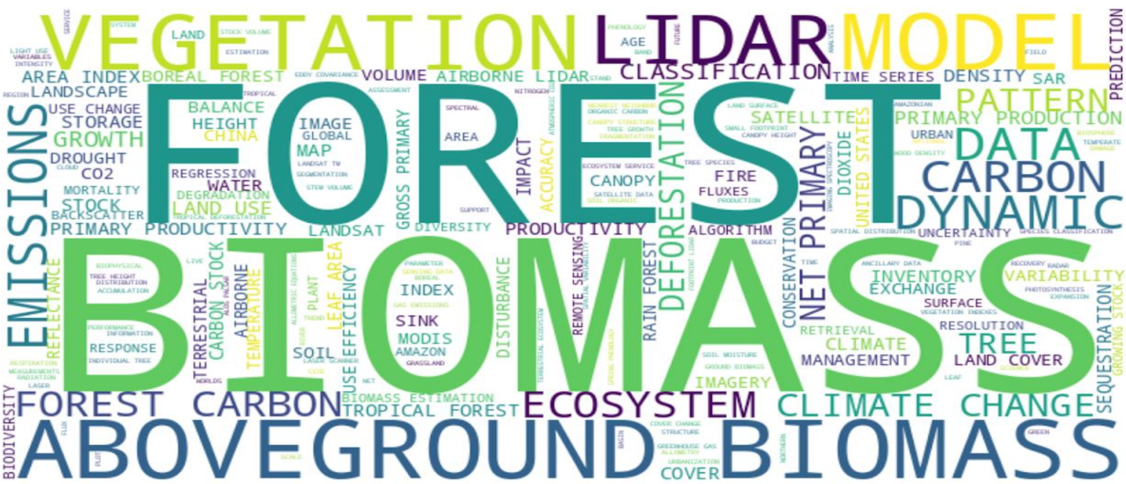
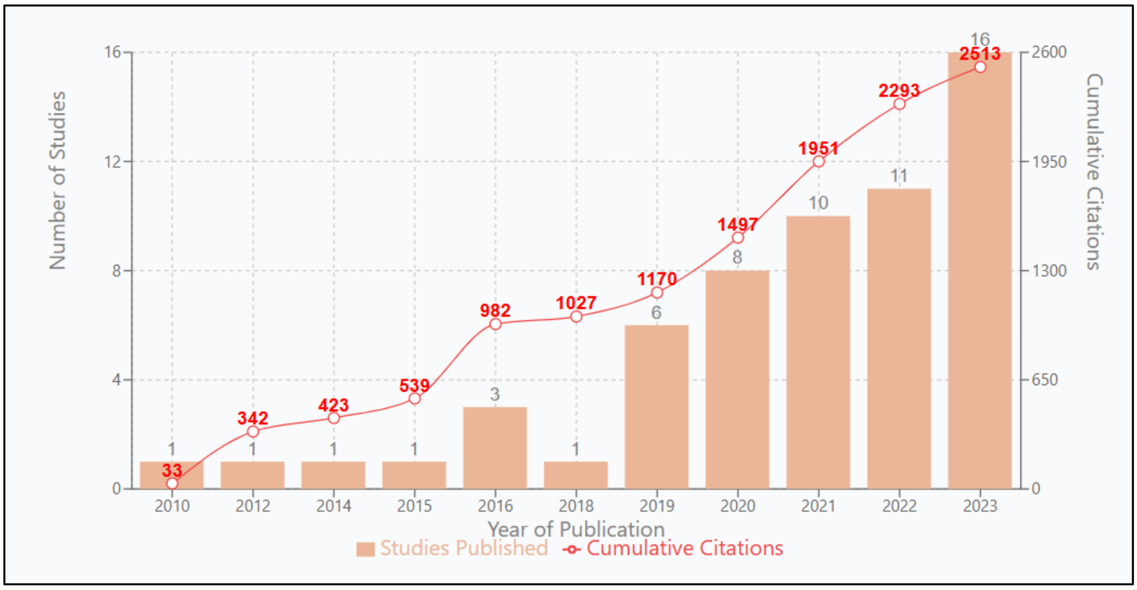


Figure 1 provides valuable insight into the current research landscape, demonstrating the convergence of remote sensing technologies with advanced machine learning approaches for forest carbon assessment. The prominence of specific sensor systems (Sentinel-2, Landsat, GEDI) and methodological approaches (Random Forest, CNN) in the word cloud reflects the current state-of-the-art in operational forest carbon monitoring applications.

RESULTS AND DISCUSSIONS

From the literature review conducted, it was observed that the earliest publication on the subject dates to 2010, according to the search platforms used. Regarding the geographical distribution of the studies, the highest concentration of research was found in Asia, Brazil, and North America.

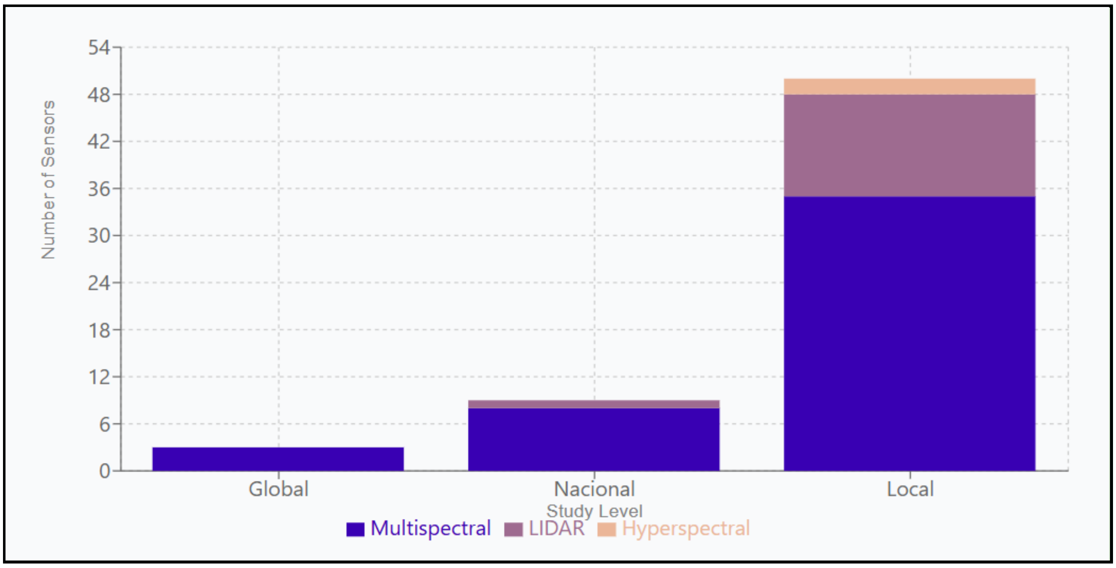
Figure 2. The x-axis highlights the publication year, and the y-axis highlights the frequency of published studies. A trend line represents the number of citations over the years.



Based on Figure 2, it can be observed that during the initial period from 2010 to 2018, few articles were published. In 2019, however, there was a considerable increase in the number of publications. This growth is related to the implementation of the Paris Agreement, which was signed in 2015 but only began to be effectively enforced during this period. From that point on, there was a greater demand for transparency and accountability, which drove the development of more accurate remote sensing methods and artificial intelligence modeling. It also encouraged greater engagement from the scientific community and policymakers. It is important to note that the 2024 study was not included in the graph, as the data were collected before the end of the year, which could lead to inaccurate results.

Therefore, it is necessary to analyze the different scales of study level to understand which type of imagery is most used in the reviewed articles.

Figure 3. Cross-study scale data with sensor type.

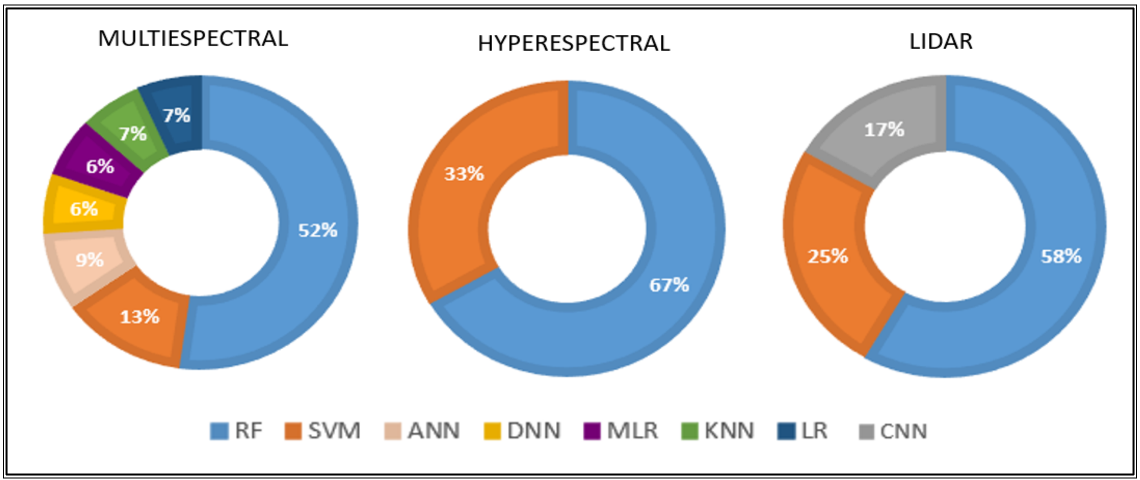


Upon analyzing the articles, it was found that the most frequent combination between the study scale and the type of sensor was the local scale with multispectral sensors. Although multispectral images offer a lower level of detail, they can be considered reliable, as they are often compared with well-established forest inventories, such as those from the Intergovernmental Panel, or with field data collected from small sample plots, which can also serve as reference data.

Additionally, it was noted that global-scale studies were among the least common, mainly due to the greater difficulty in data validation.

Therefore, to deepen the understanding of the algorithms used with each type of sensor, the information from each article was extracted and represented through donut charts.

Figure 4. Survey of the most used methods/algorithms in multispectral, hyperspectral e LIDAR sensor.



When comparing Figures 4, it is observed that the Random Forest algorithm predominates in all charts for the types of sensors—multispectral, hyperspectral, and LiDAR—as also highlighted in the study by Tu (2025). This algorithm is increasingly used due to its non-parametric nature, which allows it to effectively and intrinsically capture information, demonstrating robust and generalizable performance.

Additionally, it is worth noting that in Figures 4, the second most frequently used algorithm is SVM, which performs particularly well with high spatial resolution imagery.

Beyond the data collected and processed according to the study area of each article, few open-access websites were found to provide direct information on carbon monitoring. Among all the reviewed articles, only two websites offered freely available data. The details are listed below (table 1).

Table 1. Identification of free websites for carbon monitoring based on the articles surveyed.

Article	Description	Link	Years Covered	Spatial Resolution	Download Format	Additional Notes
Arevalo (2023)	Global Forest Watch (GFW)	https://data.globalforestwatch.org/datasets/gfw::aboveground-live-woody-biomass-density/explore	Last Update 2022	30 m	CSV, Shapefile, GeoJSON, Kml	-
Bos (2020)	Global Forest Change (GFC)	https://glad.earthengine.app/view/global-forest-chang	2000-2012 and 2000 - 2023	30 m	-	Changing forests (trees above 5 m)

In Table 1, it can be observed that the available data is of high spatial resolution, with information provided in different formats and covering different topics. The first website focuses on forest fires, while the second provides data on the density of aboveground live woody biomass.

In addition to the public websites found, other websites were found outside of the articles raised (Table 2).

Table 2. Identifying free carbon monitoring sites based on Google searches

Dataset name	Provider	Years Covered	Spatial Resolution	Global Coverage	Download Format	Site
ESA Biomass CCI (Biomass_cci)	European Space Agency	2010, 2017, 2018, 2019, 2020	100 m	Yes	GeoTIFF, netCDF	https://gee-community-catalog.org/projects/cci_agb/#citation
GlobBiomass	European Space Agency	2010	100 m	Yes	GeoTIFF	https://globbiomass.org/wp-content/uploads/GB_Maps/Globbiomass_global_dataset.html
GEDI L4B Gridded Biomass Density	NASA/ ORNL DAAC	2019 – 2023 (Version 2.1)	1 km	Partially (51.6° N/S)	GeoTIFF	https://daac.ornl.gov/GEDI/guides/GEDI_L4B_Gridded_Biomass.html

In Table 2, a lower spatial resolution can be observed, meaning it provides less detail compared to the free websites listed in Table 1, which were identified in one of the articles used for the development of this work. After compiling all the articles, the most estimated carbon monitoring unit is above-ground biomass (AGB). Below, studies with this unit in multispectral and hyperspectral analysis will be highlighted, as shown in Table 3.

Sensor Type	Authors
Multiespectral	Araza (2022), Arévalo (2023), Chen (2024), Qian (2021), El Hajj (2019), Emick (2023), Ghosh (2020), Jha (2021), Lv (2021), Padalia (2023), Saad (2023), Singh (2022), Wang (2021), Yang (2023), Zhang (2014),), Zhang (2022), Zhang (2023).
Hyperespectral	Dalponte (2016), Gao (2021), Gao (2022), Yu (2022).

It is observed that, even though hyperspectral data offers a higher level of image detail, multispectral sensors still predominate in practical applications. This is because, in many cases, less precise data may be sufficient depending on the study's objective.

CONCLUSION

Currently, remote sensing has become an essential tool for forest carbon monitoring. Although multispectral imagery is the most commonly used in studies, it still presents limitations in terms of accuracy when compared to hyperspectral and LiDAR data, which offer a higher level of detail and precision.

In the field of machine learning, there is a predominance of shallow, supervised, and efficient algorithms — especially Random Forest, which is widely applied due to its robustness and generalization capability. However, Deep Learning approaches have gained increasing attention for enabling greater automation and accuracy in detecting land-use changes, deforestation, and estimating carbon stocks.

Despite technological advances, significant challenges remain, such as the lack of data transparency and consistency, particularly in regions with dense and heterogeneous vegetation. Furthermore, acquiring up-to-date and recurring data from forest inventories — which are essential for model validation — is often expensive and time-consuming.

Given this context, investing in more advanced technological platforms is crucial to reduce statistical errors and increase the reliability of carbon estimates. With more accurate and accessible data, it will become increasingly feasible to develop free and operational carbon monitoring systems, accelerating the implementation of large-scale carbon credit projects and contributing more effectively to environmental management and climate change mitigation.

ACKNOWLEDGEMENTS

We would like to thank CAPES (Coordination for the Improvement of Higher Education Personnel) for the doctoral scholarship to one of the authors of this work. We would also like to thank CNPq (National Council for Scientific and Technological Development) for the financial support through processes 305814/2023-0 and 403213/2023-1.

REFERÊNCIAS

ARAZA, Arnan et al. A comprehensive framework for assessing the accuracy and uncertainty of global above-ground biomass maps. **Remote Sensing of Environment**, v. 272, p. 112917, 2022.

ARÉVALO, Paulo et al. Continuous mapping of aboveground biomass using Landsat time series. **Remote Sensing of Environment**, v. 288, p. 113483, 2023.

AREVALO. Global Forest Watch. **Aboveground Live Woody Biomass Density.** https://data.globalforestwatch.org/datasets/gfw::aboveground-live-woody-biomass-density/explore**.** 2023.

BOS. **Global Forest Change (GFC)**. https://glad.earthengine.app/view/global-forest-chang. 2020.

CHEN, Zhao et al. Aboveground forest biomass estimation using tent mapping atom search optimized backpropagation neural network with Landsat 8 and Sentinel-1A data. **Remote Sensing**, v. 15, n. 24, p. 5653, 2023.

DALPONTE, Michele; COOMES, David A. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. **Methods in ecology and evolution**, v. 7, n. 10, p. 1236-1245, 2016.

EL HAJJ, Mohammad et al. Mapping of aboveground biomass in Gabon. **Comptes Rendus Geoscience**, v. 351, n. 4, p. 321-331, 2019.

EMICK, Ethan et al. An approach to estimating forest biomass while quantifying estimate uncertainty and correcting bias in machine learning maps. **Remote Sensing of Environment**, v. 295, p. 113678, 2023.

GAO, Linghan; ZHANG, Xiaoli. Above-ground biomass estimation of plantation with complex forest stand structure using multiple features from airborne laser scanning point cloud data. **Forests**, v. 12, n. 12, p. 1713, 2021.

GAO, Linghan; CHAI, Guoqi; ZHANG, Xiaoli. Above-ground biomass estimation of plantation with different tree species using airborne lidar and hyperspectral data. **Remote Sensing**, v. 14, n. 11, p. 2568, 2022.

GHOSH, Sujit Madhab; BEHERA, Mukunda Dev; PARAMANIK, Somnath. Canopy height estimation using sentinel series images through machine learning models in a mangrove forest. **Remote Sensing**, v. 12, n. 9, p. 1519, 2020.

HARRIS, Nancy L. et al. Global maps of twenty-first century forest carbon fluxes. **Nature Climate Change**, v. 11, n. 3, p. 234-240, 2021.

JHA, Nidhi et al. The real potential of current passive satellite data to map aboveground biomass in tropical forests. **Remote Sensing in Ecology and Conservation**, v. 7, n. 3, p. 504-520, 2021.

LV, Guanting et al. Signatures of Wetland Impact: Spatial Distribution of Forest Aboveground Biomass in Tumen River Basin. **Remote Sensing**, v. 13, n. 15, p. 3009, 2021.

MAHESH, Batta et al. Machine learning algorithms-a review. **International Journal of Science and Research (IJSR).[Internet]**, v. 9, n. 1, p. 381-386, 2020.

PADALIA, Hitendra; PRAKASH, Ankit; WATHAM, Taibanganba. Modelling aboveground biomass of a multistage managed forest through synergistic use of Landsat-OLI, ALOS-2 L-band SAR and GEDI metrics. **Ecological Informatics**, v. 77, p. 102234, 2023.

QIAN, Chunhua et al. Estimation of forest aboveground biomass in karst areas using multi-source remote sensing data and the K-DBN algorithm. **Remote Sensing**, v. 13, n. 24, p. 5030, 2021.

SAAD, Siti Nor Maizah et al. Modeling carbon emissions of post-selective logging in the production forests of Ulu Jelai, Pahang, Malaysia. **Remote Sensing**, v. 15, n. 4, p. 1016, 2023.

Sharp R, Chaplin-Kramer R, Wood S A, Guerry A (2018). Guia do Usuário do InVEST.

SINGH, Chandrakant et al. Remote sensing-based biomass estimation of dry deciduous tropical forest using machine learning and ensemble analysis. **Journal of Environmen tal Management**, v. 308, p. 114639, 2022.

Tu, L., Pan, J. Evaluation of carbon storage in a metropolitan area based on long-term time series of Land Use/Cover Change dynamics. **Front. Earth Science**. V.19,n. 54–67 (2025). https://doi.org/10.1007/s11707-024-1127-9

YANG, Qiuli et al. Mapping high-resolution forest aboveground biomass of China using multisource remote sensing data. **GIScience & Remote Sensing**, v. 60, n. 1, p. 2203303, 2023.

YU, Shichuan et al. Effects of driving factors on forest aboveground biomass (AGB) in China's Loess Plateau by using spatial regression models. **Remote Sensing**, v. 14, n. 12, p. 2842, 2022.

WANG, Jingyi et al. Remote sensing estimation of bamboo forest aboveground biomass based on geographically weighted regression**. Remote sensing**, v. 13, n. 15, p. 2962, 2021.

ZHANG, Jian et al. Estimating spatial variation in Alberta forest biomass from a combination of forest inventory and remote sensing data. **Biogeosciences**, v. 11, n. 10, p. 2793-2808, 2014.

ZHANG, Fanyi et al. Estimation of aboveground carbon density of forests using deep learning and multisource remote sensing. **Remote Sensing**, v. 14, n. 13, p. 3022, 2022.

ZHANG, Xiaoyong et al. Simulation of Spatial and Temporal Distribution of Forest Carbon Stocks in Long Time Series—Based on Remote Sensing and Deep Learning. **Forests**, v. 14, n. 3, p. 483, 2023.