

Article

Measurement-Driven Estimates of Above-Ground Biomass Change in the Eastern Canadian Boreal Forests from Permanent Sample Plots and Landsat Time Series

Hadi Mahmoudi Meimand ¹, Jiaxin Chen ² , Daniel Kneeshaw ¹ and Changhui Peng ^{1,*} 

¹ Department of Biological Sciences, Institute of Environmental Sciences, University of Quebec at Montreal (UQAM), Montreal, QC H2X 3Y7, Canada; mahmoudi_meimand.hadi@courrier.uqam.ca (H.M.M.); kneeshaw.daniel@uqam.ca (D.K.)

² Ontario Forest Research Institute, Ministry of Natural Resources, 1235 Queen Street East, Sault Ste. Marie, ON P6A 2E5, Canada; jiaxin.chen@ontario.ca

* Correspondence: peng.changhui@uqam.ca

Abstract

Monitoring boreal above-ground biomass (AGB) change requires approaches that are both measurement-based and spatially explicit. We integrated permanent sample plots from Quebec and Ontario with Landsat-7 spectral trajectories (1999–2023) to quantify non-fire-related AGB change after excluding wildfire-affected intervals and to evaluate whether annualized AGB change can be predicted from spectral change at the plot-interval scale. Tree height was estimated using a multilayer perceptron model ($R^2 = 0.83$) and combined with species-specific allometry to derive plot-level AGB and interval Δ AGB. These estimates were aggregated to ecodistricts using effective sample sizes and confidence intervals. Across well-sampled ecodistricts, mean annualized Δ AGB ranged from -0.82 to $+3.54 \text{ t ha}^{-1} \text{ yr}^{-1}$, with lower or negative changes mainly occurring in eastern regions. Spectral indices derived from NIR–SWIR bands showed relatively stronger associations with Δ AGB than greenness-based indices, consistent with the sensitivity of moisture- and disturbance-related metrics to canopy stress, including defoliation. An XGBoost ensemble correctly predicted the direction of change in 77% of intervals. These results provide a measurement-constrained and scalable framework for monitoring non-fire-related biomass change and supporting greenhouse-gas reporting across boreal forest landscapes.

Keywords: above-ground biomass change (Δ AGB); boreal forests; sample plots; Landsat time series; spectral vegetation indices (VIs); ecodistricts; XGBoost ensemble; Google Earth Engine (GEE); machine learning (ML); defoliation



Academic Editors: Giovanni D'Amico, Davide Travaglini and Walter Mattioli

Received: 25 March 2026

Revised: 25 April 2026

Accepted: 5 May 2026

Published: 8 May 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Boreal forests are among the world's largest terrestrial carbon reservoirs and strongly influence the climate system through coupled exchanges of carbon, water, and energy [1–3]. Canada contains roughly one quarter of the world's boreal forest, making changes in forest biomass particularly consequential for national greenhouse-gas inventories and global carbon budget assessments [4–7]. Accurate estimation of above-ground biomass (AGB) in forests is therefore crucial for quantifying carbon storage and assessing forest dynamics. Recent studies have shown that integrating ground-plot inventories with multi-source satellite observations and machine or deep learning models substantially improves AGB

estimation by linking field-based biomass measurements to complex spectral and structural signals in remote-sensing data [8–12].

At the same time, the magnitude, the drivers, and the spatial patterns of recent AGB change in boreal forests remain uncertain. In Canada's boreal forests, drought-induced increases in tree mortality and associated reductions in carbon-sink strength, together with biotic disturbances such as insect outbreaks, have already produced measurable losses of live biomass [13–18]. Repeated inventory analyses and satellite-based vegetation indices reveal widespread greening and biomass gains in many regions, but also climate-related slowdowns and local declines, with different methods often disagreeing on the relative roles of climate versus disturbance processes [19–24]. Near-term (5–30-year) [25] forecasts of Δ AGB are particularly challenging: stand-level growth trajectories are comparatively predictable, whereas partial mortality and other non-stand-replacing processes are multi-factorial and stochastic [26,27]. Recent machine learning work combining large permanent sample plot networks, climate covariates, and multi-decadal Landsat time series has begun to map where near-term biomass gains and losses are likely to occur, but mainly at continental or broad ecozone scales rather than at decision-relevant units such as eastern Canadian ecodistricts [25].

Biomass monitoring in Canada rests on two complementary observational pillars. The plot-based pillar consists of the National Forest Inventory (NFI) and aligned provincial remeasurement programmes (e.g., Ontario Ministry of Natural Resources and Forestry; Ministère des Ressources naturelles du Québec), which provide detailed repeated tree and stand measurements from sample plots (including permanent sample plots (PSP), Permanent Inventory Plots (PIP), and Permanent Growth Plots (PGP) networks) that underpins standardized national reporting and model calibration [28]. The remote-sensing pillar encompasses a suite of model-based mapping approaches that integrate plot measurements with optical, radar, and lidar observations to generate wall-to-wall estimates of biomass and related forest attributes. A range of statistical and machine learning approaches has been used to map AGB from optical time series calibrated with inventory plots (e.g., k-nearest neighbour (kNN), regression/ensemble learning, and deep learning) [9,11,29–34], yet uncertainties remain in translating plot-level change to spatially explicit, temporally consistent estimates across administrative/ecological units.

Despite major advances, three important gaps remain for the eastern Canadian boreal forest. First, to our knowledge, most carbon budget and biomass studies report trends at national or broad ecozone scales and do not explicitly isolate wildfire-masked, multi-decadal changes in live AGB at the ecodistrict scale, which represents a primary ecological reporting unit used in regional planning and greenhouse-gas inventories; consequently, background growth and non-wildfire biomass losses remain poorly quantified at this scale [4,25,35]. Second, although model-assisted and model-based frameworks that combine forest inventory and remote sensing are well developed, they have rarely been applied to estimate annualized changes in AGB (Δ AGB; $\text{t ha}^{-1} \text{ yr}^{-1}$) derived from multi-year remeasurement intervals at the ecodistrict level using long Landsat time series with design-consistent uncertainty, as many operational products prioritize single-epoch biomass or disturbance status [9,12,36]. Third, because near-term biomass change is only intermittently predictable at the plot scale [25], ecodistrict-scale Δ AGB remains difficult to predict, and the sources of disagreement between plot-based and Landsat-derived Δ AGB remain poorly understood, particularly with respect to sampling intensity, disturbance history, and spatial representativeness [37–40].

The primary objective of this study was to quantify non-wildfire, annualized change in live above-ground biomass (Δ AGB; $\text{t ha}^{-1} \text{ yr}^{-1}$) at the ecodistrict scale across boreal Ontario and Québec by integrating sample-plot remeasurements with Landsat-7 time series

and quantifying uncertainty consistent with the sampling and remeasurement structure. To isolate background dynamics relevant to management and greenhouse-gas accounting, we exclude stand-replacing wildfire [41] intervals using national burned-area composites and an independent burn-detection map, such that ΔAGB reflects growth, partial harvesting, insects, and other non-wildfire processes. We then quantify how well plot-interval ΔAGB and its components can be predicted from Landsat spectral trajectories, and we diagnose both disagreements and similarities between plot-based and satellite-derived indicators of biomass change as a function of sampling intensity, disturbance history, and spatial representativeness. Finally, we used these relationships to support spatially explicit characterization of biomass gain and loss across ecodistricts, including areas with sparse plot coverage.

2. Materials and Methods

2.1. Study Area

The study domain spans the eastern Canadian boreal forest across Québec and Ontario, covering approximately 193 million ha. The site-level inventory comprises 12,366 unique plots, acquired between 1999 and 2024 from permanent sample plots (PSPs), Permanent Growth Plots (PGPs), and Permanent Inventory Plots (PIPs) (Table A1), with sampling primarily concentrated in the southern boreal portion of the region (Figure 1a,b). Following standard PSP protocols, each census represents a complete field measurement for a given plot at a specific time, and successive censuses from the same plot were paired to define remeasurement intervals used to characterize stand dynamics [42]. Plot areas follow provincial design specifications and are overwhelmingly dominated by 0.04 ha plots ($n = 12,258$), whereas alternative plot sizes (0.01, 0.02, 0.05, 0.06, 0.09, and 0.1 ha) occur only rarely ($n \leq 38$ for each class). The dataset also provides repeated measurements over time, with 7641 plots remeasured twice, 1324 three times, and 178 four times, while 3223 plots were measured only once. For each census, we compiled plot identifiers and georeferenced coordinates, measurement year, species-resolved tree records, diameter at breast height (DBH), total height, and tree status from the Ontario Ministry of Natural Resources and Forestry Growth and Yield Program and Québec's Données Québec repository (Ministère des Ressources naturelles et des Forêts).

The study area was stratified at the ecodistrict level of the Canadian National Ecological Framework. Ecodistricts are fourth-level units nested within ecozones, ecoprovinces, and ecoregions, delineated as relatively homogeneous areas in terms of climate, landforms, surficial deposits, soils, water and potential vegetation [43]. Within the study area, a total of 85 ecodistricts were considered; 66 ecodistricts contained sample plots whereas 19 ecodistricts lacked plots. Figure 1b shows pronounced spatial variation in sampling intensity, with plot density ranging from 0.02 to 9.83 plots per 10,000 ha and highest densities in the southern and central managed forest belt. Figure 1c summarizes ecodistrict-level sampling effort and temporal coverage, showing that both plot counts and remeasurement intervals vary by more than an order of magnitude across ecodistricts, and that observation periods typically span approximately two to three decades. Collectively, this ecodistrict-based stratification provides a policy-relevant reporting framework for summarizing plot-derived AGB dynamics and for mapping spatial patterns of ΔAGB using wall-to-wall Landsat coverage, while recognizing higher uncertainty in ecodistricts lacking direct plot support [44].

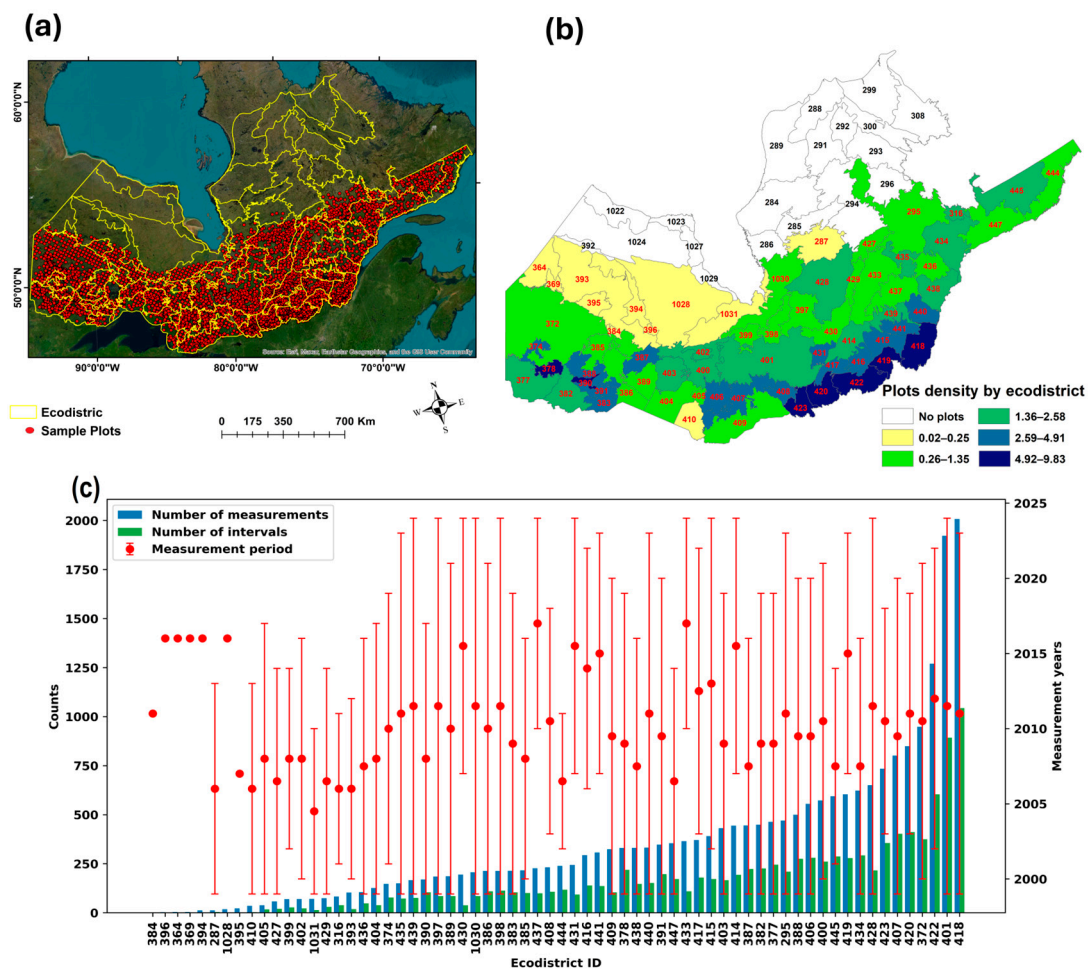


Figure 1. (a) Location of sample plots within the study area in the eastern boreal forest of Ontario and Québec. (b) Spatial distribution of sample plot density by ecodistrict. Colours indicate plot density (plots per 10,000 ha) for ecodistricts with at least one plot, whereas white polygons (“No plots”) denote ecodistricts without plot measurements (not used for model calibration/validation). Sampling is concentrated within the southern and central managed/commercial forest belt. Red numbers in panel (b) indicate ecodistrict identification codes. (c) Sampling effort and temporal coverage by ecodistrict. Bars show, for each ecodistrict (x-axis), the number of plots (blue) and the number of remeasurement intervals used in the analysis (green). Red symbols represent the measurement period in each ecodistrict, from the first to the last remeasurement year. Ecodistrict boundaries and numeric codes follow the Canadian National Ecological Framework, and ecodistricts in panel (c) are ordered from left to right by increasing number of plots.

2.2. Ground Sample Plots and AGB Estimation

Field data collection in the provincial plot inventories followed standard boreal forest inventory protocols, with diameter at breast height (DBH; 1.3 m) and total height and species identity recorded for each enumerated stem [11]. However, tree inclusion thresholds and associated biomass estimation conventions differ among inventories. For example, in the United States stems are typically classified as “trees” when $\text{DBH} \geq 12.7$ cm (USDA Forest Service), whereas the Canadian National Forest Inventory follows a lower threshold of $\text{DBH} \geq 9.1$ cm [45]. Consistent with Canadian National Forest Inventory conventions and provincial PSP protocols, AGB in this study was quantified for all live trees with $\text{DBH} \geq 9.1$ cm measured at 1.3 m. Across 1,272,299 DBH records (Ontario: 780,392; Quebec: 491,907), the distribution is strongly dominated by smaller diameter classes. The 9.1–15 cm class accounts for 64.8% of stems, followed by 15–20 cm (22.4%) and 20–30 cm (11.4%). Larger-diameter classes are comparatively rare, with 30–50 cm representing 1.3% of stems

and >50 cm only 0.1% (Figure A1). To examine the structural characteristics of the sampled trees, all individuals with both recorded DBH and measured height were compiled across the study area. This subset comprised 341,771 paired DBH–height observations. Within this dataset, DBH averaged 18.08 cm with a range from 9.10 to 107.40 cm. Correspondingly, tree height averaged 14.54 m, ranging from 1.30 to 43.24 m. The distributions of both DBH and tree height are illustrated in (Figure A2), showing that DBH exhibits a right-skewed distribution with most observations concentrated in smaller diameter classes, whereas tree height is primarily distributed within intermediate ranges.

Species identity is a key determinant of AGB because allometric relationships between DBH, height, and biomass vary systematically among taxa owing to differences in wood density, crown architecture, and stem form. Broad allometric compilations for North America have shown that applying generic or mis-specified species groups can introduce stand-level biomass biases on the order of 10%–30%, particularly along hardwood–softwood gradients [46,47]. To reduce such bias and ensure consistency with Canadian forest carbon reporting, we used species-specific or genus-level equations, together with hardwood and softwood group models from the Canadian national biomass equation sets [48,49] with the general functional form described in Equations (1)–(4) to convert tree measurements to individual-tree AGB prior to plot-level aggregation.

Given that about 27% of trees with DBH measurements had observed heights and DBH–height allometry is nonlinear and species-dependent, we developed and validated a multilayer perceptron (MLP) model to predict tree height from DBH and species, thereby obtaining complete height information for AGB estimation. As shown in Figure 2, the height model was trained only on trees with both DBH and height measurements. Height data are available for 71 of the 76 species in the study area; the remaining 5 species have DBH records only. Given this uneven coverage and the strong imbalance in sample sizes among species, we grouped taxa into allometric classes to stabilize model training while retaining all observations. As shown in Table 1, the grouping follows established biomass–height allometry families, with softwood groups 1–3 (spruce/fir/hemlock, pine, cedar/larch) and hardwood groups 4–7 (oak/hickory/hard maples, soft maple/birch/cherries, aspen/poplar/willow, and mixed hardwoods), adapted from [47,48,50]. Height prediction is therefore conditioned on allometric group rather than individual species, improving model robustness and species coverage. This MLP approach was selected given its capacity to approximate complex non-linear functions [51]. The dataset was partitioned by stratified random sampling across allometric classes into training (70%), validation (15%), and test (15%) subsets, preserving class proportions. To assess potential data leakage associated with tree-level random partitioning, an additional independent validation was conducted using a plot-level split, where all trees from a given plot were assigned to a single subset. This approach ensured that no information was shared between training and evaluation datasets at the plot level. Predictors comprised DBH (cm) as a continuous variable and species identity as a categorical variable (encoded), with total tree height (m) as the response. Hyperparameters were tuned using RandomizedSearchCV over a predefined search space (Table A2) [52] to balance model flexibility and simplicity. Model performance was assessed using regression metrics for continuous outcomes, including the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE), providing complementary measures of model accuracy and error magnitude under species-specific allometry and a skewed DBH distribution [53]. The final tuned MLP was then applied to all trees lacking measured height, yielding a spatially and taxonomically complete height dataset for subsequent AGB computation.

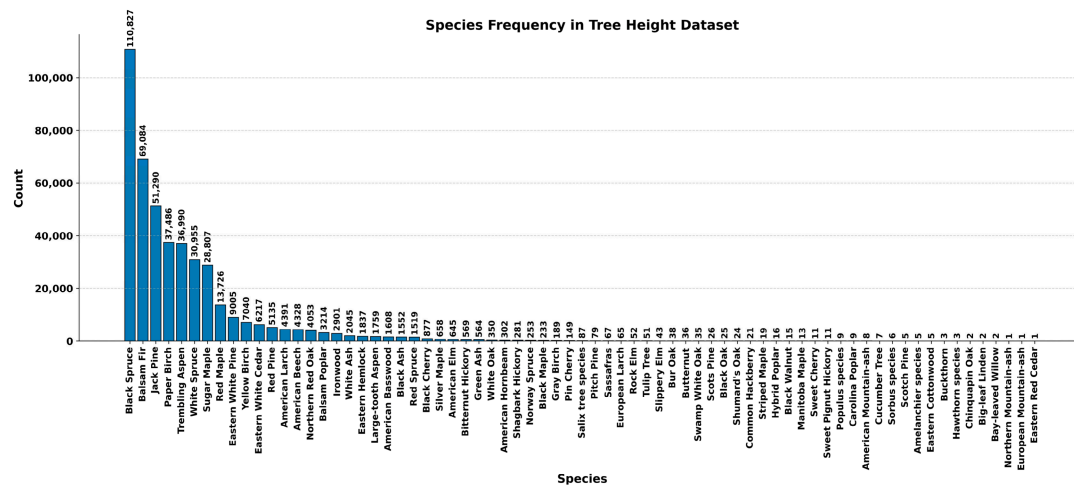


Figure 2. Frequency distribution of tree species included in the tree height dataset (71 species).

Table 1. Classification of tree groups and their associated allometric subgroups.

Group Name	Allometric Subgroup	Number of Tree Records
Softwood Group 1	Spruce/Fir/Hemlock	214,475
Softwood Group 2	Pine	65,540
Hardwood Group 5	Soft Maple/Birch/Cherries	59,131
Hardwood Group 6	Aspen/Poplar/Willow	42,075
Hardwood Group 4	Oak/Hickory/Hard Maples	41,975
Softwood Group 3	Cedar/Larch	10,674
Hardwood Group 7	Mixed Hardwoods	7772

After predicting tree height with the MLP, AGB was estimated at the component, tree, and plot levels and then expressed per unit area. For each biomass component c (wood, bark, branches, foliage), component-level biomass (kg) for tree j was computed using species-specific power-law allometric equations [48,49]:

$$B_{c,j} = \alpha_{c,s} DBH_j^{\beta_{c,s}} H_j^{\theta_{c,s}} \quad (1)$$

where $B_{c,j}$ is the biomass (kg) of component c for tree j ; DBH_j is diameter at breast height (cm); H_j is total height (m); and $\alpha_{c,s}$, $\beta_{c,s}$, $\theta_{c,s}$ are species- and component-specific coefficients (See Table A3). Total tree biomass (kg) was then obtained as the sum of component-level biomass across all components:

$$AGB_j = \sum_{c \in \{\text{wood, bark, branches, foliage}\}} B_{c,j} \quad (2)$$

Plot-level biomass (kg) for plot p was obtained by summing over all trees j within that plot:

$$AGB_p^{(\text{kg})} = \sum_{j \in p} AGB_j \quad (3)$$

finally, biomass per unit area ($AGB_p^{(\text{t ha}^{-1})}$) for plot p is computed by converting kilograms to tonnes and normalizing by plot area ($Area_p^{(\text{ha})}$):

$$AGB_p^{(\text{t ha}^{-1})} = \frac{AGB_p^{(\text{kg})} / 1000}{Area_p^{(\text{ha})}} \quad (4)$$

Figure 3 summarizes the workflow used to convert ground-plot measurements to plot-level AGB.

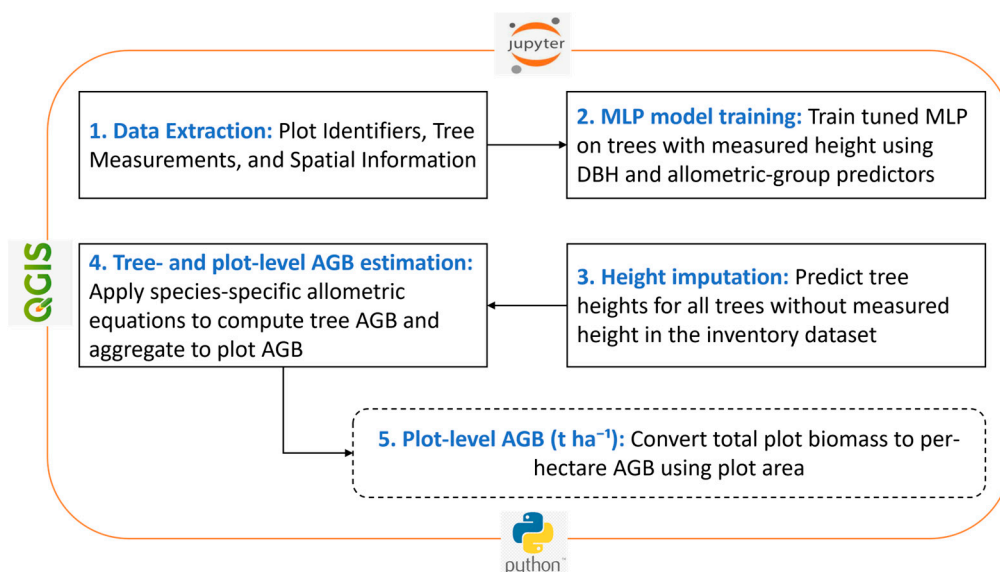


Figure 3. Workflow for estimating plot-level above-ground biomass (AGB, t ha^{-1}) from ground plots (2000–2024).

2.3. Plot-Level AGB Change and Ecodistrict Aggregation

Annualized change in live AGB for each remeasurement interval i ($\Delta \text{AGB}_{\text{yr},i}$ ($\text{t ha}^{-1} \text{yr}^{-1}$)) was computed using plot-level biomass expressed per unit area (t ha^{-1}), as defined in Equation (4), as:

$$\Delta \text{AGB}_{\text{yr},i} = \left(\text{AGB}_{p, \text{end}, i}^{(\text{t ha}^{-1})} - \text{AGB}_{p, \text{start}, i}^{(\text{t ha}^{-1})} \right) / \Delta t_i \quad (5)$$

where $\text{AGB}_{p, \text{start}, i}^{(\text{t ha}^{-1})}$ and $\text{AGB}_{p, \text{end}, i}^{(\text{t ha}^{-1})}$ denote the plot-level AGB per unit area at the beginning and end of remeasurement interval i , respectively and Δt_i is the length (years) of interval i . After initial screening, interval-level quality control (QC) was applied to a set of 10,147 remeasurement intervals prior to aggregation [25,54]. Intervals with missing required variables (e.g., plot-level attributes or interval metrics) were excluded, although no records were removed at this stage. Intervals with short remeasurement periods ($\Delta t_i < 4$ years) were excluded to avoid unstable annualized estimates, resulting in the removal of 28 intervals. In addition, implausible values of $\Delta \text{AGB}_{\text{yr},i}$ were filtered using a conservative, data-informed threshold of $|\Delta \text{AGB}_{\text{yr},i}| > 10 \text{ t ha}^{-1} \text{yr}^{-1}$, removing a further 83 intervals. This threshold was selected to exclude extreme outliers inconsistent with observed boreal biomass dynamics while retaining realistic variability. After QC, 10,036 intervals were retained for analysis. To assess sensitivity to the plausibility threshold, alternative cutoffs of ± 8 and $\pm 12 \text{ t ha}^{-1} \text{yr}^{-1}$ were also evaluated. Resulting ecodistrict-level estimates and classifications were largely unchanged, supporting the robustness of the selected cutoff.

The plot to ecodistrict aggregation and uncertainty-estimation workflow is summarized in Figure 4. Aggregating interval-level changes rather than static stocks provides a more direct measure of contemporary carbon fluxes and disturbance–recovery dynamics in boreal forests [25,35,40,55,56].

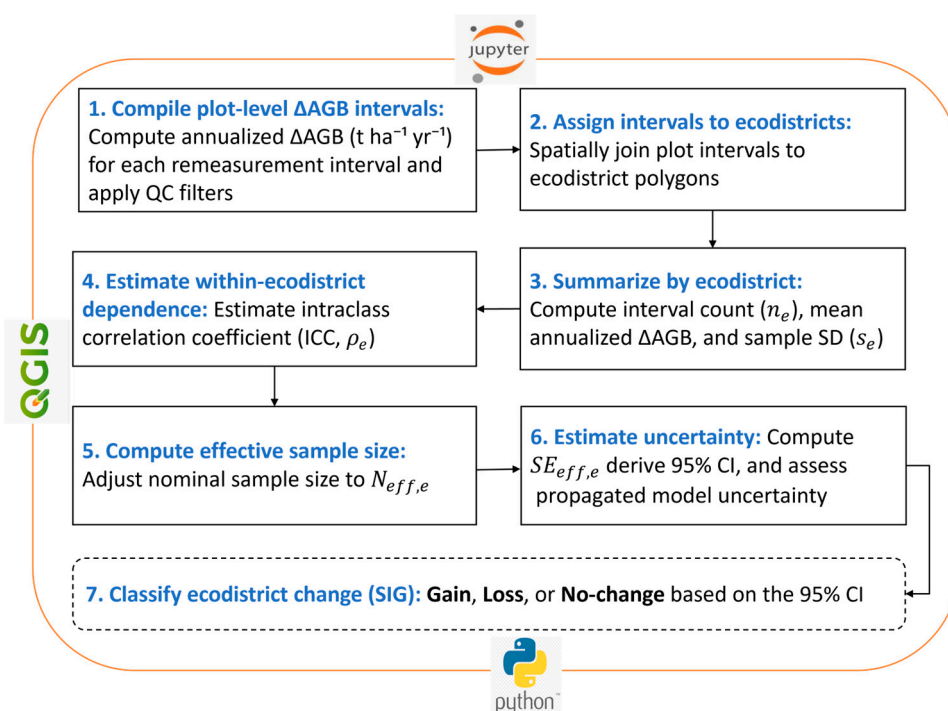


Figure 4. Workflow for estimating ecodistrict-level annualized above-ground biomass change (ΔAGB) from plot remeasurement intervals, including aggregation, effective sample size adjustment, uncertainty estimation, and significance classification (2000–2024).

To obtain spatially explicit statistics suitable for linking with remote sensing predictors, we grouped intervals by ecodistrict and summarized them using domain-level summary estimators. For each ecodistrict e , we computed the arithmetic mean annual AGB change ($\Delta\text{AGB}_{\text{yr},e}$ ($\text{t ha}^{-1} \text{yr}^{-1}$)), as:

$$\Delta\text{AGB}_{\text{yr},e} = \frac{1}{n_e} \sum_{i \in e} \Delta\text{AGB}_{\text{yr},i} \quad (6)$$

where n_e is the number of intervals in ecodistrict e . We also computed the sample standard deviation s_e of interval-level $\Delta\text{AGB}_{\text{yr},i}$ and the number of contributing plot intervals. Because multiple intervals within an ecodistrict are not fully independent, we adjusted the nominal sample size using an effective sample size $N_{\text{eff},e}$ derived from the intraclass correlation coefficient (ICC) and calculated an effective standard error as

$$SE_{\text{eff},e} = s_e / \sqrt{N_{\text{eff},e}} \quad (7)$$

A 95% confidence interval ($\text{CI}_{95,e}$) for the $\Delta\text{AGB}_{\text{yr},e}$ was obtained using a t -critical value with $N_{\text{eff},e} - 1$ degrees of freedom. We do not assume that interval-level $\Delta\text{AGB}_{\text{yr},i}$ itself follows a symmetric Student's t distribution; rather, the t -based CI is used as an approximation for the ecodistrict-level mean $\Delta\text{AGB}_{\text{yr},e}$ and aggregation across intervals reduces the effect of skewness. Ecodistricts were classified as “Gain” if the lower CI bound of $\text{CI}_{95,e}$ was >0 , “Loss” if the upper CI bound was <0 , and “No change” otherwise, and a separate “Insufficient” class was assigned to domains with very few intervals or undefined variance. Full formulas for $N_{\text{eff},e}$, $SE_{\text{eff},e}$, the t -critical values, and the 95% confidence bounds, as well as the significance (SIG) classification rule, are provided in Appendix B. This combination of interval-level QC, domain-level effective sample sizes, and CI-based significance labelling is consistent with commonly used approaches for large-area biomass-change estimation from NFI-type data [25,40,55,57].

Monte Carlo simulation is commonly used to propagate uncertainty in environmental models by repeatedly perturbing uncertain input variables and propagating these uncertainties through the full modelling chain [58,59]. In this study, uncertainty in tree height estimates was propagated through biomass calculations and Δ AGB estimation using repeated stochastic simulations ($n = 100$ runs). For trees with imputed height, perturbations were applied assuming approximately Gaussian prediction errors:

$$H_j^{(r)} = H_j + \varepsilon_j^{(r)}, \varepsilon_j^{(r)} \sim N(0, \sigma_H^2) \quad (8)$$

where $H_j^{(r)}$ is the perturbed height of tree j in simulation r , H_j is the original (predicted) height of tree j , and $\varepsilon_j^{(r)}$ is a random error term drawn independently for each tree and simulation. The parameter σ_H represents the standard deviation of the height prediction error, approximated here by the RMSE of the height prediction model. The perturbed heights were then propagated through allometric equations, plot-level aggregation, interval-level Δ AGB estimation, and ecodistrict-level confidence interval and classification (SIG) calculations. In addition, a sensitivity analysis was performed to evaluate the influence of allometric uncertainty by applying a uniform $\pm 5\%$ scaling factor to AGB estimates. Together, these analyses provide an approximate quantitative assessment of the sensitivity of Δ AGB estimates and ecodistrict classifications to uncertainties in tree height prediction and allometric scaling.

2.4. Remote Sensing Predictors and Δ AGB Modelling

In Google Earth Engine (GEE), we generated Landsat-7 Collection 2, Level-2 surface reflectance composites for July–August over 1999–2023 to represent peak growing-season conditions in the eastern Canadian boreal forest [25]. For each calendar year Y , we constructed a two-year moving window ($Y - 1$ to Y), applied the standard USGS radiometric scaling, and removed cloud, cloud shadow, snow, and cirrus using the pixel-quality band, following national Landsat compositing protocols [8,30]. To mitigate the Landsat-7 SLC-off (scan line corrector failure) striping effects after 2003, we relied on multi-temporal compositing (two-year window) combined with median aggregation to reduce data gaps, together with a 3×3 focal median filter to suppress residual striping artefacts. In addition, the use of a single-sensor (Landsat-7) time series ensured spectral consistency across the study period, avoiding inter-sensor radiometric differences that could affect spectral indices and Δ AGB modelling. Following these steps, we computed the median surface reflectance in the BLUE, GREEN, RED, NIR, SWIR1, and SWIR2 bands, as well as the per-pixel number of clear observations. Plot-level predictors were obtained by sampling the start- and end-year composites (30 m resolution) at plot centroids. At the ecodistrict scale, the same July–August composites were spatially aggregated within each ecodistrict polygon after masking permanent water bodies and stand-replacing burned areas [41]. Band medians and mean clear-observation counts were then derived for an “early” period (2000–2012) and a “late” period (2012–2023), consistent with recent Landsat-based biomass-change studies [25].

From these July–August Landsat-7 composites, we derived a suite of spectral vegetation indices characterizing canopy greenness, moisture status, structure, and senescence at both plot and ecodistrict scales. For each plot interval, we calculated Normalized Difference Vegetation Index (NDVI), Normalized Burn Ratio (NBR), Normalized Difference Infrared Index (NDII), Normalized Difference Water Index (NDWI), Moisture Stress Index (MSI), Plant Senescence Reflectance Index (PSRI), Soil-Adjusted Total Vegetation Index (SATVI), Enhanced Vegetation Index (EVI), and Near-Infrared Reflectance of Vegetation (NIRv) at interval start and end (Table A4). Annualized change metrics ($\Delta \text{VI}_{\text{yr}}$, index units yr^{-1})

were obtained by differencing end and start interval values and dividing by the remeasurement interval length. At the ecodistrict level, the same indices and annualized changes were computed for the early (2000–2012) and late (2012–2023) periods, yielding spatially aggregated spectral trajectories analogous to the plot-level $\Delta\text{AGB}_{\text{yr}}$ estimates and used as predictors in subsequent XGBoost-based ΔAGB modelling [25].

To model annualized plot-level AGB change, we used an Extreme Gradient Boosting (XGBoost) regression ensemble [60]. XGBoost was selected because it can efficiently model nonlinear relationships and interactions in large remote-sensing datasets and has been widely used in biomass modelling studies [25]. To benchmark the added value of XGBoost, two simpler models were also evaluated using the same predictors and grouped cross-validation framework: Ridge regression [61] and Random Forest [62]. Additional configuration details for the primary and benchmark models are summarized in Table A7. Predictors included start-of-interval values and annualized changes in spectral indices, along with interval duration, interval mid-year, a binary indicator identifying intervals spanning both the early (2000–2012) and late (2012–2023) periods, and a clear-sky observation metric. We used five-fold cross-validation grouped by plot identifier [63] to avoid leakage among intervals from the same plot, tuning a compact grid of XGBoost hyperparameters with early stopping and winsorizing ΔVI features to limit the influence of outliers. The XGBoost model was fit for absolute ΔAGB ($\text{t ha}^{-1} \text{ yr}^{-1}$) and model performance was summarized using MAE, RMSE, R^2 , and directional accuracy, defined as the proportion of intervals for which the model correctly predicts the sign of ΔAGB . Additional binary classification metrics (precision, recall, F1-score, balanced accuracy, and false negative rate (FNR)) were computed from out-of-fold predictions to assess model performance for loss detection [25]. A final seed ensemble was then refit using all QC intervals, after which permutation-based feature importance was applied to rank predictors.

In the final step, we used the trained XGBoost ensemble of plot-level ΔAGB predictions to generate ecodistrict-level estimates of $\Delta\text{AGB}_{\text{yr},e}$ ($\text{t ha}^{-1} \text{ yr}^{-1}$). For each ecodistrict, we assembled a feature table of Landsat-7 spectral indices for the early (2000–2012) and late (2012–2023) periods, together with their annualized changes, and passed these predictors through each member of the absolute ΔAGB ensemble, averaging across seeds to obtain a single model-based ΔAGB estimate per ecodistrict. We then combined these predictions with ecodistrict-level sampling statistics from the plot network (number of intervals, within ecodistrict variance, effective sample size) to compute standard errors and 95% CI, following model-assisted small area estimation practice in large area forest inventories [36,64,65]. Ecodistricts whose CIs were entirely above zero were classified as biomass “Gain”, those entirely below zero as “Loss”, and all others as “No change”, yielding a directionally explicit map of ΔAGB consistent with the gain/loss summaries in [25]. The overall remote-sensing workflow is summarized in Figure 5.

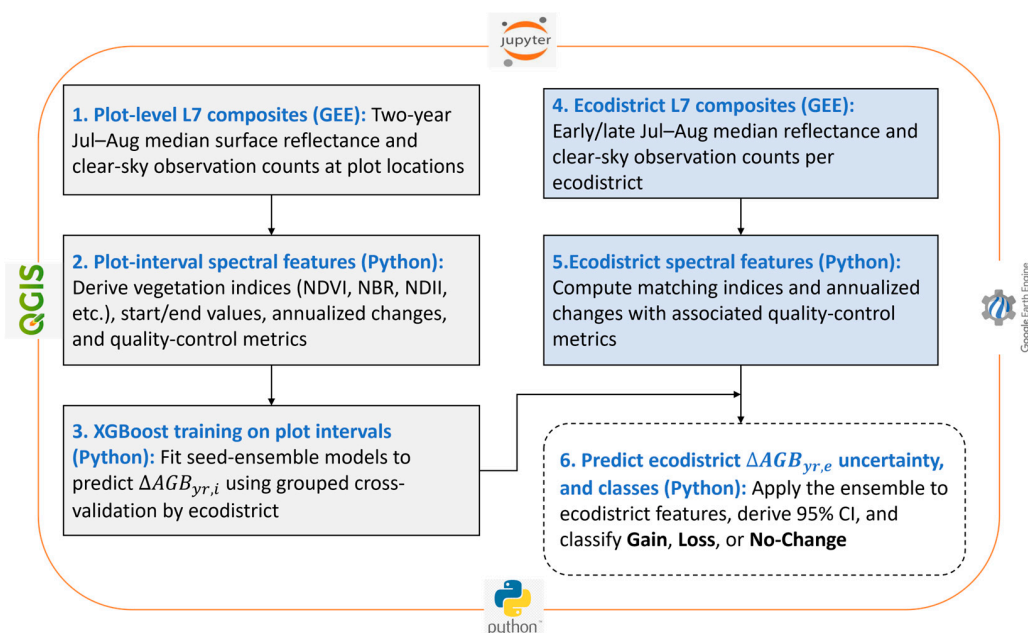


Figure 5. Workflow for Landsat-7 predictor generation, plot-level XGBoost model training, and ecodistrict-scale annualized above-ground biomass change (ΔAGB) prediction with uncertainty assessment (2000–2023).

3. Results

3.1. Plot-Level AGB Estimates from MLP-Based Height Prediction and Allometric Equations

The study area contains 76 tree species in total, including 18 softwoods and 58 hardwoods. Softwoods account for 81% of all records, while hardwoods comprise 19% of records. Species composition was dominated by three boreal softwoods, including black spruce (35.42% of stems), jack pine (19.25%), and balsam fir (18.59%), while the leading hardwoods are paper birch (8.27%) and trembling aspen (7.37%). Together, these five species represent about 89% of all records. The sampled dataset was dominated by spruce, fir, and pine, whereas birch and aspen represented the most common broadleaf associates (see Table A5).

The final MLP configuration selected by grid search comprised three hidden layers with 128, 64, and 32 neurons, ReLU activation, stochastic gradient descent (SGD), and modest L2 regularization ($\alpha = 0.0001$). This architecture showed consistent performance across the validation and independent test datasets, with $R^2 \approx 0.83$ in both cases. Corresponding error metrics were also nearly identical, with RMSE values of 2.32 and 2.31 m and MAE values of 1.76 and 1.75 m for the validation and test sets, respectively. The similarity between validation and test metrics suggests good generalization with limited evidence of overfitting.

Scatterplots of observed versus MLP-predicted tree height for both the validation and test datasets (Figure 6a,b) show a dense, approximately linear distribution of points aligned with the 1:1 line. The model captures the primary gradient of tree height across the full range of observations, with no strong evidence of systematic curvature or severe heteroscedasticity. These patterns suggest that a single global non-linear model was adequate across species groups and DBH classes. A slight tendency toward underestimation is observable for taller trees (>30–35 m), likely reflecting the limited representation of large stems in the training data; however, these deviations are confined to the upper tail of the distribution and do not materially affect overall model performance. An additional independent validation using a plot-level split yielded nearly identical results to those obtained from the original

partitioning ($R^2 \approx 0.83$; $RMSE \approx 2.30$ m; Figure A3), indicating that model performance is robust and not materially influenced by sample dependence.

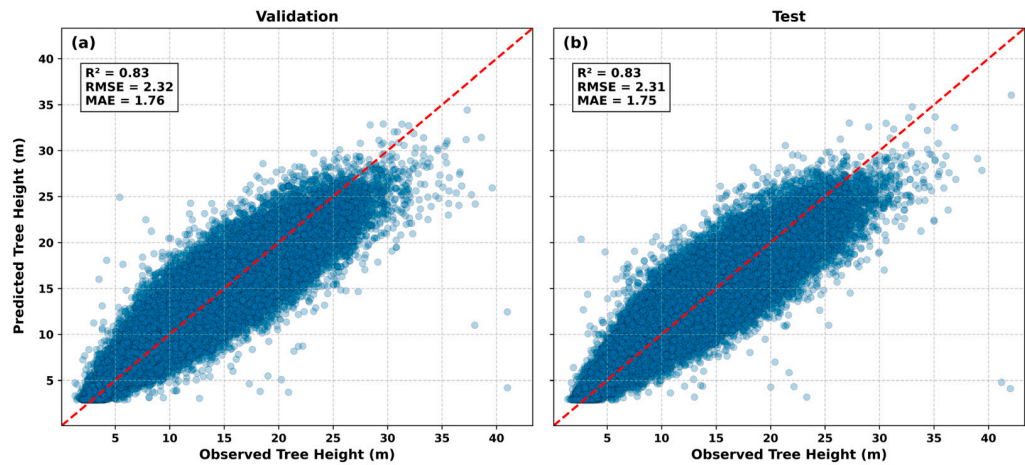


Figure 6. Observed versus MLP-predicted tree height for (a) validation and (b) test datasets. The dashed line indicates the 1:1 relationship, showing strong agreement ($R^2 \approx 0.83$) and consistent errors across datasets.

Model errors were relatively consistent across the seven allometric groups, with MAE on the test set ranging from 1.52 to 2.20 m (Table 2). RMSE values showed a similar pattern (2.02–2.80 m), while R^2 ranged from 0.64 to 0.86, indicating broadly consistent predictive performance across groups. The lowest errors occurred for the dominant softwood group comprising spruce, fir, and hemlock (MAE = 1.52 m; $R^2 = 0.82$), whereas higher errors were observed for the pine group (MAE = 2.20 m) and the cedar–larch group (MAE = 2.08 m), possibly reflecting greater variability in height–diameter relationships and the relatively small sample size of the cedar–larch group ($n = 1601$). Hardwood groups showed intermediate performance. Overall, these results indicate that the MLP provides sufficiently consistent height predictions across the main species groupings required for subsequent biomass estimation.

Table 2. Test-set performance metrics (MAE, RMSE, and R^2) of the MLP tree height model across seven allometric groups.

Group Name	Allometric Subgroup	Count	MAE	RMSE	R^2
Softwood Group 1	Spruce/Fir/Hemlock	32,172	1.52	2.02	0.82
Softwood Group 2	Pine	9831	2.20	2.80	0.79
Hardwood Group 5	Soft	8870	1.78	2.32	0.76
Hardwood Group 6	Maple/Birch/Cherries	6311	1.87	2.43	0.85
Hardwood Group 4	Aspen/Poplar/Willow	6296	1.93	2.54	0.81
Softwood Group 3	Oak/Hickory/Hard Maples	1601	2.08	2.78	0.64
Hardwood Group 7	Cedar/Larch	1166	1.89	2.48	0.86
	Mixed Hardwoods				

After tree height prediction with the MLP, measured and predicted heights were used in the species-specific allometric equations to compute tree-level AGB. The resulting biomass distribution is summarized below. Most stems fell within low to intermediate biomass classes (10–50 kg), with progressively fewer trees in higher AGB classes. Detailed counts by biomass class and tree type, together with summary statistics, are provided in Figure A4. Plot-level AGB ranged from approximately 0.1 to 380 t ha^{−1}, with a mean of 71.2 t ha^{−1} and a standard deviation of 46.6 t ha^{−1}, indicating considerable variability among plots. The boxplot (Figure A5) shows a right-skewed distribution with several

high-AGB outliers above approximately 200 t ha^{-1} , likely representing a relatively small number of very dense stands.

Mean plot-derived AGB at the ecodistrict scale ranged from about 23 to 148 t ha^{-1} (Figure 7). The lowest mean values ($23\text{--}43 \text{ t ha}^{-1}$) occurred mainly in northern and interior ecodistricts with colder climates and a higher prevalence of open or low-productivity forest types, whereas intermediate values ($44\text{--}77 \text{ t ha}^{-1}$) dominated much of central Québec and Ontario. The highest mean AGB classes ($78\text{--}148 \text{ t ha}^{-1}$) were concentrated in southern and southwestern ecodistricts, where longer growing seasons and the predominance of productive mixedwood and hardwood stands likely contributed to higher biomass, consistent with previous findings for Canadian boreal forests [60–63]. Ecodistricts lacking plots are shown as “No plots” in this plot-based summary and were not included in the calculation of mean plot-derived AGB.

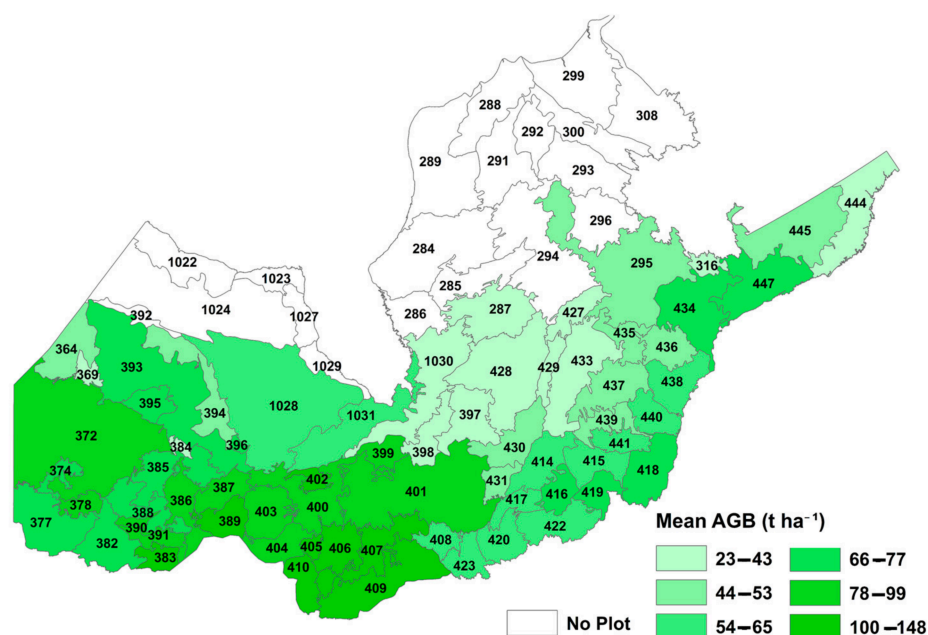


Figure 7. Spatial distribution of mean plot-derived above-ground biomass (AGB, t ha^{-1}) by ecodistrict in the eastern commercial boreal forests of Québec and Ontario (1999–2024). Coloured polygons show ecodistricts classified into AGB classes based on the mean of all plots within each unit; white polygons (“No plots”) denote ecodistricts without plot measurements, which were excluded from model calibration and evaluation.

3.2. Ecodistrict-Level Mean Annualized AGB Change

Having characterized static AGB patterns, we next examined annualized biomass change and associated uncertainty. Effective sample sizes ($N_{\text{eff},e}$) varied from 3 to more than 1000 remeasurement intervals (median ≈ 130), producing 95% confidence-interval widths (CI widths) between 0.27 and $5.44 \text{ t ha}^{-1} \text{ yr}^{-1}$ (median ≈ 0.72). Figure 8 illustrates the inverse relationship between data density and precision: CI width declines steeply as the number of remeasurement intervals increases and gradually levels off at higher sampling frequencies. This pattern indicates that ecodistricts with relatively few remeasurement intervals yield more uncertain estimates of ΔAGB , whereas ecodistricts with larger numbers of intervals provide substantially narrower confidence intervals for mean biomass change. The observed trend is broadly consistent with the expected reduction in sampling uncertainty with increasing sample size. Variance decomposition further showed that most variability occurred within ecodistricts rather than between them ($\sigma_{\text{within}}^2 = 8.60$ vs. $\sigma_{\text{between}}^2 = 1.18$), corresponding to an ICC of $\rho_d = 0.12$. This moderate within-ecodistrict correlation supports the use of an effective sample size adjustment.

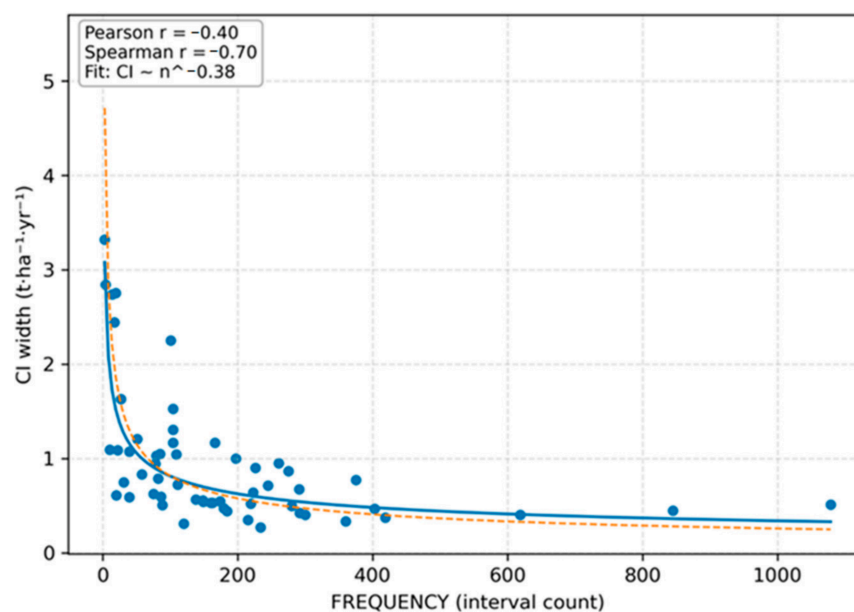


Figure 8. Relationship between sampling density and precision of ecodistrict-level ΔAGB estimates. Each point represents one ecodistrict. The x-axis shows the number of remeasurement intervals, and the y-axis shows the 95% confidence-interval width for mean annualized ΔAGB ($\text{t ha}^{-1} \text{ yr}^{-1}$). The downward trend indicates tighter estimates with increasing data density. The solid curve shows the fitted power-law relationship, whereas the dashed curve indicates a $1/\sqrt{n}$ reference decline.

Across the 59 ecodistricts with sufficient data, $\Delta\text{AGB}_{\text{yr},e}$ ranged from -0.82 to $+3.54 \text{ t ha}^{-1} \text{ yr}^{-1}$, with a median of 0.78 and an overall mean of $0.97 \text{ t ha}^{-1} \text{ yr}^{-1}$, indicating that modest net biomass gains predominated overall, although two ecodistricts experienced substantial net losses. The strongest gains occurred mainly in southern ecodistricts, including 403, 404, 406, and 389, whereas pronounced losses were concentrated in eastern ecodistricts including 447 and 445. Based on the CIs, 37 ecodistricts were classified as significant biomass gains, 2 as losses, and 20 as no change; the remaining 7 ecodistricts containing plots were classified as “Insufficient data” (Figure 9a,b). Sensitivity analyses confirmed that these patterns were robust to the choice of plausibility threshold. Only two ecodistricts (393 and 419) changed classification under the stricter $\pm 8 \text{ t ha}^{-1} \text{ yr}^{-1}$ threshold, shifting from no change to gain due to minor shifts in their confidence intervals around zero. No classification changes were observed between the ± 10 and ± 12 thresholds, indicating that the main conclusions were insensitive to the selected cutoff.

Uncertainty propagation analysis indicated that ecodistrict-level ΔAGB estimates were robust to both tree height prediction error and allometric scaling uncertainty. Monte Carlo simulations ($n = 100$) showed that, on average, fewer than one ecodistrict changed classification in a typical run (mean = 0.16), and only four out of 59 ecodistricts exhibited any classification change across all simulations, with only one ecodistrict showing a classification-change probability greater than 10%. The mean absolute change in interval-level annualized ΔAGB was $0.17 \text{ t ha}^{-1} \text{ yr}^{-1}$, while changes at the ecodistrict level were minimal, with a mean absolute difference of approximately $0.025 \text{ t ha}^{-1} \text{ yr}^{-1}$ and a mean change in confidence interval width of approximately $0.04 \text{ t ha}^{-1} \text{ yr}^{-1}$ (Table A6). These variations are small relative to the overall range and distribution of ΔAGB values (Figure 9a). Similarly, a $\pm 5\%$ sensitivity analysis of allometric scaling resulted in no changes in ecodistrict classification (SIG), with only minor proportional effects on ecodistrict mean ΔAGB ($\approx 0.025 \text{ t ha}^{-1} \text{ yr}^{-1}$) and confidence interval width ($\approx 0.04 \text{ t ha}^{-1} \text{ yr}^{-1}$). Together, these results demonstrate that the spatial patterns of biomass gain, loss, and no

change are insensitive to uncertainties in tree height estimation and moderate variations in allometric equations.

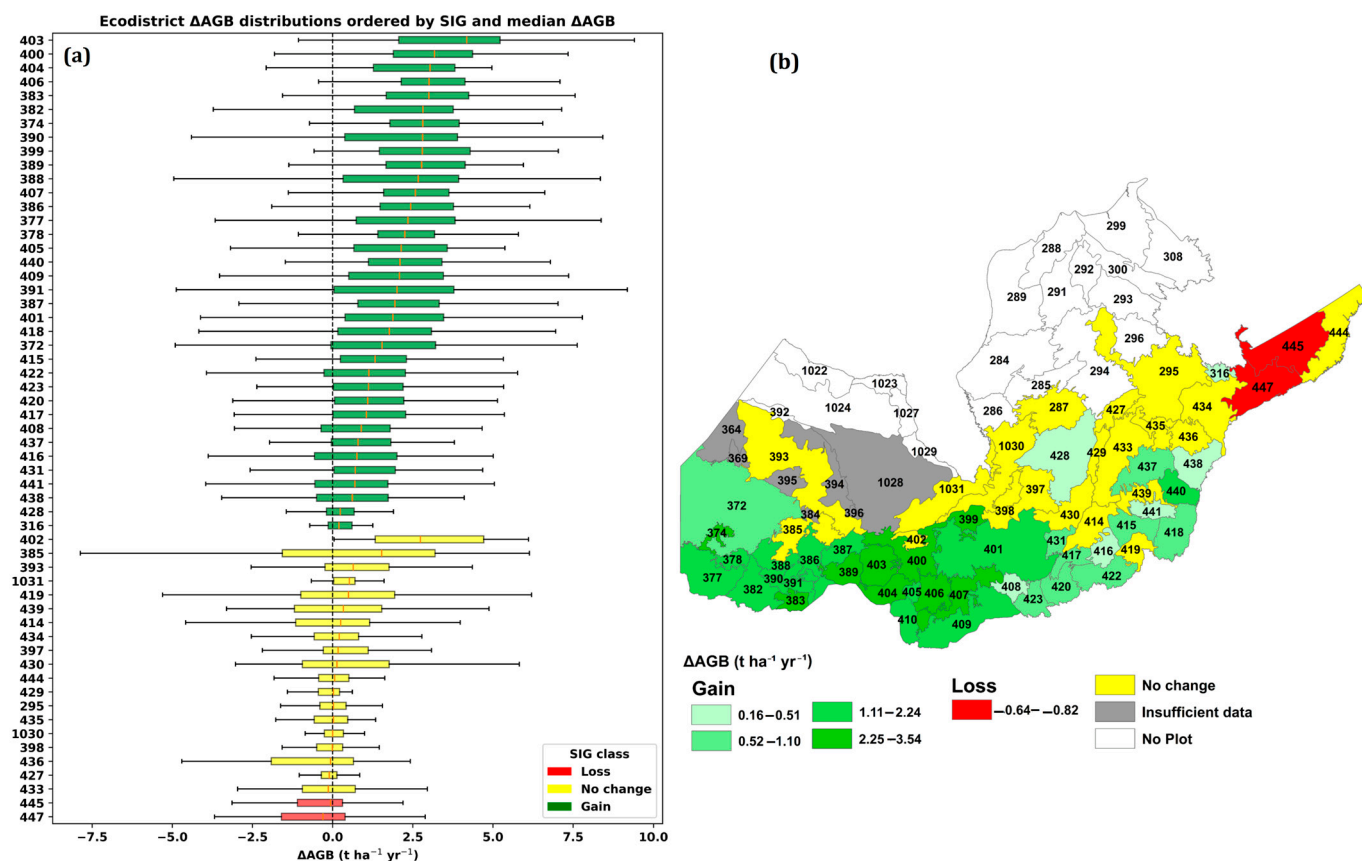


Figure 9. Plot-based ecodistrict classifications of mean annual above-ground biomass change ($\Delta\text{AGB}_{\text{yr},e}$ ($\text{t ha}^{-1} \text{yr}^{-1}$)) across the eastern boreal forest, 1999–2024. (a) Distributions of interval level $\Delta\text{AGB}_{\text{yr}}$ for each ecodistrict shown as horizontal boxplots, ordered by significance class (SIG) and median $\Delta\text{AGB}_{\text{yr}}$. Boxes are coloured by SIG class (gain, no change, loss) based on 95% confidence intervals (CI) around the mean. (b) Corresponding map of ecodistrict mean ΔAGB , with green tones indicating significant biomass gain, red tones significant biomass loss, and yellow no significant change. Dark grey polygons denote ecodistricts with insufficient interval data and light grey polygons denote ecodistricts without plots. Numbers label ecodistrict IDs used in the text.

3.3. Remote-Sensing Predictors and ΔAGB Model Performance

Before fitting the XGBoost models, we assessed whether July–August Landsat spectral indices contained a detectable signal of annualized ΔAGB change. Across 8123 quality-controlled (QC) plot intervals, annual changes in all indices were generally small in magnitude; however, several moisture and disturbance-sensitive indices exhibited moderate but consistent associations with ΔAGB (Figure 10). In particular, NDII and NBR showed relatively higher absolute Pearson correlation coefficients with ΔAGB . These relationships are consistent with the physical interpretation of these indices. NBR, which contrasts near-infrared and shortwave infrared reflectance, is widely used to detect canopy disturbance and vegetation recovery and can therefore respond to structural changes in vegetation [66]. Similarly, NDII is sensitive to variations in canopy water content and vegetation moisture stress, conditions that are frequently associated with changes in forest productivity and biomass [67]. Consequently, intervals characterized by increasing canopy moisture and structural recovery (higher NDII and NBR and lower MSI) tended to correspond to biomass gains, whereas intervals exhibiting the opposite spectral responses were more likely to show biomass losses. In contrast, greenness indices such as NDVI and EVI exhibited com-

paratively lower correlations with ΔAGB , suggesting that simple changes in photosynthetic activity alone are a relatively poor proxy for structural biomass change over 5–10-year periods. Hexagonally binned scatterplots of ΔNBR and ΔNDVI (Figure A6) further illustrate these patterns. The highest densities of points occur near zero change in both ΔAGB and ΔVI , reflecting the predominance of low-magnitude gains and losses across intervals, while larger biomass decreases tend to be associated with modest declines in indices such as NBR and NDVI. However, the clouds of points are broad and approximately elliptical, with substantial overlap between gain and loss intervals, emphasizing that individual spectral indices explain only a modest fraction of the variance in ΔAGB and that the underlying relationships are non-linear and noisy. Together, these diagnostics motivate the use of a multi-index, non-linear ensemble model rather than reliance on any single spectral index or threshold to infer biomass change.

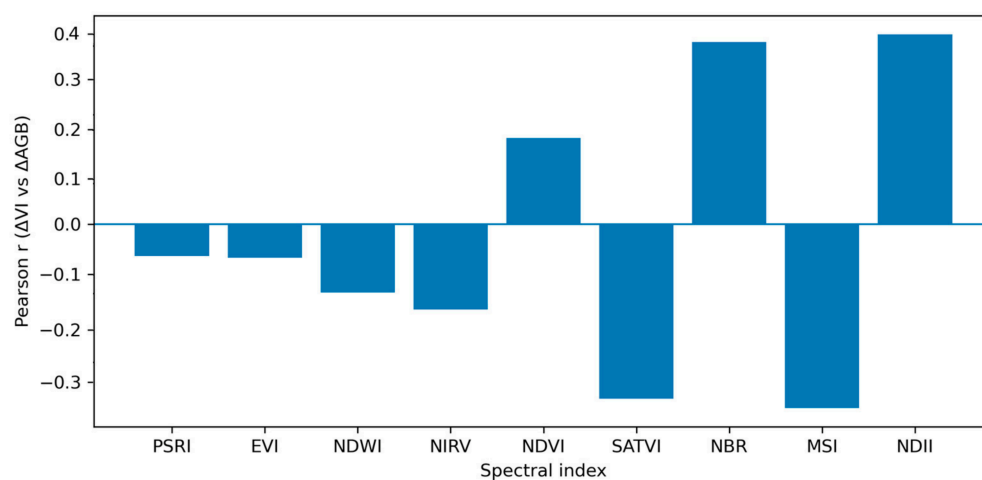


Figure 10. Pearson correlations between annualized changes in spectral vegetation indices (ΔVI) and annualized above-ground biomass change (ΔAGB) across 8123 quality-controlled plot intervals. Indices derived from NIR–SWIR bands (e.g., NDII, NBR) show relatively stronger associations with ΔAGB compared to greenness-based indices.

At the plot level, the XGBoost ensemble for predicting $\Delta\text{AGB}_{\text{yr}}$ achieved an MAE of $1.54 \text{ t ha}^{-1}\text{yr}^{-1}$, an RMSE of $2.42 \text{ t ha}^{-1}\text{yr}^{-1}$, and $R^2 = 0.40$ across QC remeasurement intervals. The model correctly predicted the direction (sign) of biomass change for 77% of intervals, with directional accuracy of 0.88 for gains and 0.48 for losses. Additional classification metrics further quantified this asymmetry. Loss detection showed a precision of 0.57, a recall of 0.48, and an F1-score of 0.52, with a balanced accuracy of 0.68 and a false negative rate of 0.52. The confusion matrix indicated that 1363 observed loss intervals were misclassified as gains, whereas 1265 loss intervals were correctly identified, highlighting that omission of loss events remains the primary limitation. This asymmetry suggests that the ensemble more reliably identifies biomass gains than losses. Loss events are captured less consistently, which may reflect their lower frequency, and greater variability among disturbance processes. In addition, this limitation likely reflects the reduced sensitivity of optical predictors to abrupt disturbance-driven biomass losses compared to more gradual gain processes [40]. Residual diagnostics (Figure 11) indicate that errors are centred near zero for moderate ΔAGB values, with a tendency to underestimate the magnitude of large losses and slightly overestimate strong gains (Figure 11a), consistent with regression toward the mean. Residuals show no systematic bias with census interval length over 5–20 years (Figure 11b), indicating that the weighting scheme and explicit inclusion of interval duration effectively control duration-related biases and support subsequent aggregation of plot-level predictions to the ecodistrict scale.

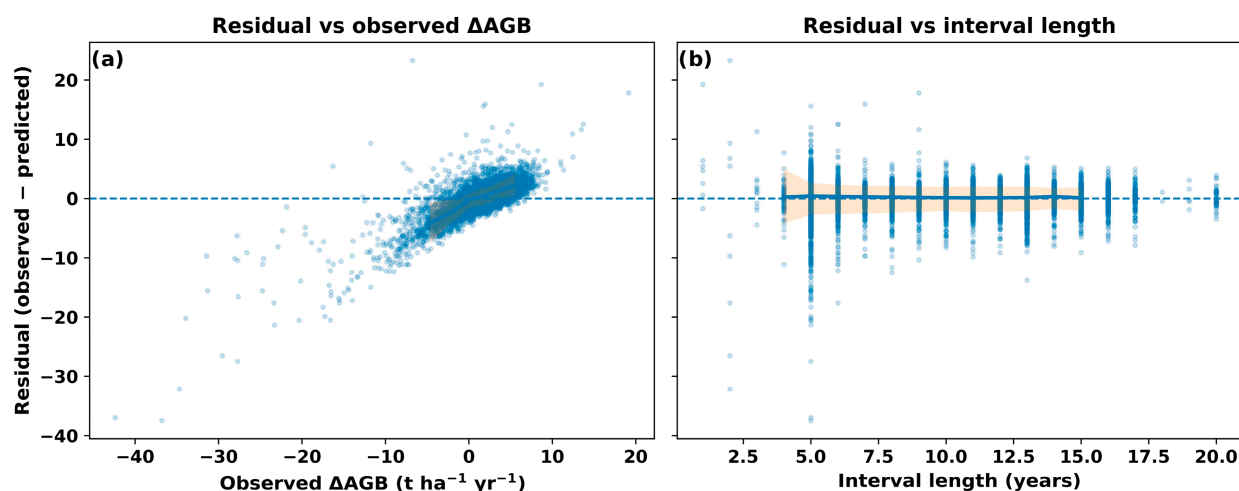


Figure 11. Residual diagnostics for the plot-level XGBoost ΔAGB model. (a) Residuals (observed – predicted) versus observed annual biomass change (ΔAGB , $\text{t ha}^{-1} \text{ yr}^{-1}$), with binned median and 10th–90th percentile envelopes indicating underestimation of the largest losses and slight overestimation of strong gains. (b) Residuals versus census interval length, showing no clear systematic relationship across 5–20-year remeasurement periods.

Comparative model performance is summarized in Table 3. The comparison shows that the linear Ridge model performs substantially worse than the tree-based models, whereas both Random Forest and XGBoost outperformed Ridge regression, highlighting the importance of nonlinear relationships. XGBoost provided the best overall performance, with modest improvements over Random Forest in R^2 , RMSE, and loss-detection accuracy.

Table 3. Comparison of models for plot-level ΔAGB prediction; Overall Direction Accuracy (ODA).

Model	R^2	RMSE	MAE	ODA	Accuracy	
					Gain	Loss
Ridge regression	0.29	2.63	1.70	0.75	0.86	0.44
RF	0.38	2.45	1.54	0.77	0.88	0.44
XGBoost	0.40	2.42	1.54	0.77	0.88	0.48

Across the 85 ecodistricts with valid remote-sensing-derived predictors, the XGBoost model predicts that biomass dynamics between the early (2000–2012) and late (2012–2023) periods are dominated by net gains (Figure 12). Using a conservative threshold of $|\Delta\text{AGB}_{\text{yr}}| > 0.05 \text{ t ha}^{-1} \text{ yr}^{-1}$ to avoid over-interpreting near-zero model outputs, 61 ecodistricts are classified as experiencing biomass gains, with a mean $\Delta\text{AGB}_{\text{yr}}$ of $0.44 \text{ t ha}^{-1} \text{ yr}^{-1}$, and a maximum of $1.10 \text{ t ha}^{-1} \text{ yr}^{-1}$. Eleven ecodistricts exhibit net losses, with mean $\Delta\text{AGB}_{\text{yr}} = -0.17 \text{ t ha}^{-1} \text{ yr}^{-1}$ and a range from -0.29 to $-0.06 \text{ t ha}^{-1} \text{ yr}^{-1}$, while a further 13 ecodistricts fall within the no change band ($|\Delta\text{AGB}_{\text{yr}}| \leq 0.05 \text{ t ha}^{-1} \text{ yr}^{-1}$) and have near-zero mean change ($\sim 0.01 \text{ t ha}^{-1} \text{ yr}^{-1}$). Spatially, positive $\Delta\text{AGB}_{\text{yr}}$ is widespread across most of Ontario, except for some western ecodistricts, and is also prevalent in southern Québec. In contrast, biomass losses are more spatially localized and occur in a limited number of ecodistricts, particularly in parts of western Ontario (e.g., 372, 384, 395) and eastern Québec (e.g., 436, 444), as well as in northern Québec (e.g., 288, 289, 291, 292, 308). These localized losses may be associated with non-fire disturbances such as harvesting, insect outbreaks, windthrow, or stand-level mortality. Predictions for ecodistricts with sparse or no plot coverage should be interpreted with greater caution, as they rely more heavily on model extrapolation.

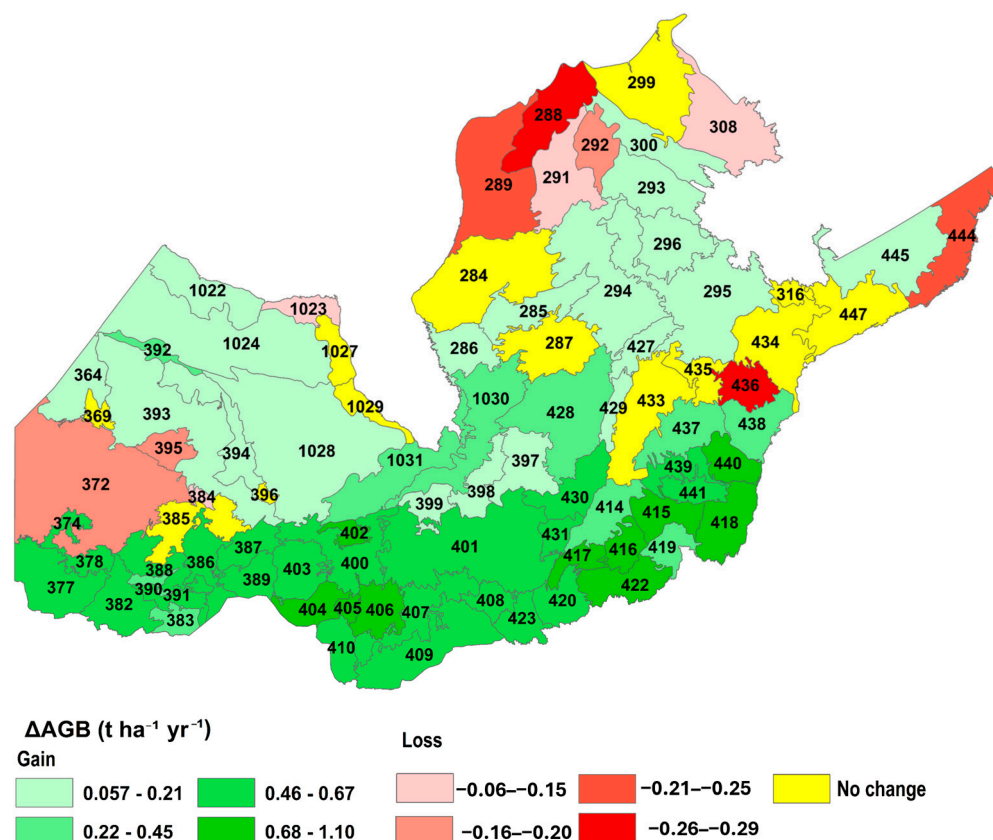


Figure 12. Model-predicted mean annualized above-ground biomass change ($\Delta AGB_{yr,e}$ (t ha⁻¹ yr⁻¹)) at the ecodistrict level across the eastern boreal forests of Ontario and Quebec between the early (2000–2012) and late (2012–2023) periods.

4. Discussion

4.1. Disturbance Regimes and Interpretation of Biomass Gain and Loss

Disturbance-composition analysis, based on ecodistrict-level aggregation of plot-derived ΔAGB intervals and Landsat-based disturbance classifications [68], indicates that non-fire biomass dynamics are strongly dominated by defoliation, with comparatively smaller contributions from harvest and windthrow (Figures 13a,b and A7). Defoliation accounts for 76.8%, 74.4%, and 92.3% of total non-fire disturbance proportion in the gain, no change, and loss classes, respectively, in the plot-based framework (Figure 13a), increasing to 87.4%, 93.2%, and 94.4% in the Landsat-based framework (Figure 13b). The consistently high contribution of defoliation in the loss class across both frameworks suggests that biomass decline in our study area is primarily associated with defoliation-driven disturbance regimes. This interpretation is consistent with previous Canadian forest-health studies showing that insect defoliation can cause cumulative canopy stress, growth reduction, mortality, and changes in forest carbon dynamics [69]. In contrast, harvest shows its highest contribution in the gain class and declines toward the loss class, indicating that harvesting was not consistently associated with long-term biomass decline at the ecodistrict scale, likely due to regeneration and spatial averaging effects. Windthrow remains a minor component across all classes. These results further suggest that Landsat time series may preferentially capture spatially coherent, multi-year canopy stress signals, whereas plot-based observations remain more sensitive to local heterogeneity and stand-level management effects.

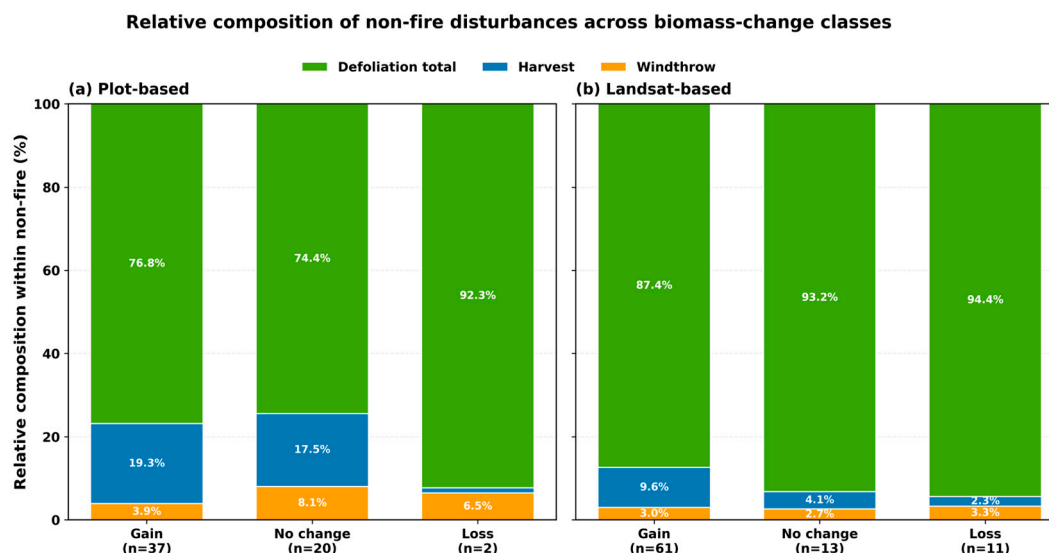


Figure 13. Relative composition of non-fire disturbances across biomass-change classes (gain, no change, loss) at the ecodistrict scale, based on (a) plot-derived Δ AGB intervals and (b) Landsat-based disturbance classifications. Defoliation represents the dominant disturbance type across all classes in both frameworks, with comparatively smaller contributions from harvest and windthrow.

In contrast, harvest shows its highest contribution in the gain class and declines toward the loss class, indicating that harvesting does not necessarily lead to long-term biomass decline at the ecodistrict scale, likely due to regeneration and spatial averaging effects. Windthrow remains a minor component across all classes. These results further suggest that Landsat time series preferentially capture spatially coherent, multi-year canopy stress, whereas plot-based observations retain sensitivity to local heterogeneity and management effects [68,69]. This is ecologically plausible for Quebec and Ontario, where spruce budworm and other defoliators drive widespread canopy stress.

This disturbance-driven interpretation also helps explain the asymmetric model performance observed between biomass gains and losses. While gains are detected more reliably, losses are more difficult to predict, likely due to their spatial heterogeneity, episodic nature, and weaker or less consistent expression in optical spectral signals [25,40]. This asymmetry is consistent with previous studies showing that AGB loss is inherently more difficult to predict than gain because biomass losses are often driven by stochastic and multifactorial processes, including mortality, partial disturbance, insect outbreaks, and climate stress [25]. In contrast, biomass gains tend to follow more gradual and predictable growth trajectories, resulting in higher model skill.

It is also important to note that not all ecodistricts classified as losses in the Landsat-based results are equally supported by field observations. A comparison among the plot-based map (Figure 9), the Landsat-based prediction map (Figure 12), and the disturbance composition results (Figure A7) indicates that several northern ecodistricts (e.g., 288, 289, 291, 292, and 308), where plot coverage is limited or absent, were classified as losses primarily through model extrapolation rather than direct plot support. Consistent with this, Figure A7 shows that some of these ecodistricts exhibit little or no detectable non-fire disturbance signal, suggesting a potential mismatch between predicted biomass decline and observed disturbance patterns. In contrast, ecodistricts such as 445–447, which were also classified as losses in the plot-based results, are better supported by field data and therefore provide more robust evidence of biomass decline. These discrepancies highlight the need for caution when interpreting loss patterns in sparsely sampled regions and emphasize the importance of improved field coverage and validation in northern ecodistricts.

These patterns indicate that disagreement between plot-based and Landsat-based Δ AGB estimates arises from the interaction of sampling intensity, disturbance history, and spatial representativeness.

4.2. Temporal, Spatial, and Methodological Factors Affecting AGB Trend Interpretation

Discrepancies between plot-based and satellite-derived AGB trends are often driven by temporal mismatches in their observation windows. Ground plots typically record biomass change over discrete remeasurement intervals of varying length, whereas Landsat-based analyses summarize dynamics across fixed and temporally consistent epochs. For example, ecodistricts 445 and 447 exhibited statistically significant AGB losses based on plot-based results (Figure 9) during approximately 2000–2014. However, Landsat-based estimates (Figure 12) aggregated over the early (2000–2012) and late (2012–2023) periods instead indicated a small gain and no change, respectively. Such differences can arise because plot networks may capture disturbance phases (e.g., harvest or insect outbreaks) without fully representing subsequent recovery, whereas satellite composites integrate both decline and regrowth processes over broader time windows. Previous studies have similarly shown that asynchronous comparison periods can introduce systematic bias and generate apparent inconsistencies in biomass trend assessments [34,70].

Spatial support differences between field plots and satellite observations provide a second source of discrepancy. Permanent sample plots typically represent areas of approximately 400 m², whereas a single Landsat pixel covers about 900 m². Consequently, localized disturbances or recovery processes detected within individual plots may be diluted when averaged to the pixel scale and further smoothed during aggregation to the ecodistrict level. This mismatch is particularly relevant in heterogeneous forest landscapes, where fine-scale mortality, partial disturbance, or patchy regeneration may strongly influence plot measurements but remain muted in coarser satellite summaries. As a result, plot-based and remotely sensed estimates may differ even when both accurately characterize biomass dynamics at their respective spatial supports [38,39].

A further source of discrepancy may arise from uneven plot representation within large ecodistricts. Where permanent sample plots are clustered in accessible, productive, or less disturbed portions of an ecodistrict, plot-based summaries may not fully capture the broader spatial mosaic of disturbance and recovery conditions. This issue may help explain pronounced inconsistencies in ecodistrict 372, where plot-derived estimates indicated biomass gain, whereas Landsat-based results suggested biomass decline. In such cases, field plots may overrepresent intact or recovering stands while underrepresenting disturbed areas occurring elsewhere within the ecodistrict. Consequently, discrepancies between plot-based and remotely sensed estimates may reflect sampling representativeness as much as true ecological disagreement.

A further source of mismatch may arise from preprocessing choices applied to the Landsat predictors. In our workflow, plot-level spectral values were derived from two-year July–August composites and a 3 × 3 focal-median window to reduce noise, cloud contamination, and data gaps. While this preprocessing improves radiometric stability and reduces the influence of anomalous individual pixels, it may also attenuate highly localized disturbance or recovery signals at individual plot locations by incorporating information from neighbouring pixels and adjacent years. As a result, abrupt stand-level changes detected in field measurements may appear weaker in the remotely sensed predictors. This scale-dependent smoothing effect is consistent with known challenges in linking plot observations to landscape-level AGB estimates [38,39].

In addition to temporal and spatial mismatches, spectral limitations inherent to optical remote sensing further constrain the detection of Δ AGB. In high-biomass forests, vegetation

indices often exhibit saturation effects, showing limited sensitivity to continued biomass accumulation despite measurable increases at the plot level [71,72]. Similarly, disturbance- and moisture-sensitive indices may capture initial biomass loss but may fail to fully represent gradual post-disturbance recovery [73,74]. These limitations highlight the need for multi-index and non-linear modelling approaches when interpreting biomass dynamics from satellite data.

Taken together, these sources of discrepancy do not imply that either framework is inherently flawed; rather, they reflect expected differences in temporal coverage, spatial support, sampling representativeness, preprocessing choices, and spectral sensitivity between field and satellite observations.

4.3. Spectral Vegetation Indices (VIs) as Imperfect but Complementary Proxies of AGB Change

Landsat-derived vegetation indices provide indirect proxies of Δ AGB, but their capacity to capture structural dynamics is limited. As shown in Section 3.3, greenness indices such as NDVI and EVI exhibited relatively lower correlations with Δ AGB, consistent with previous findings that photosynthetic proxies often fail to reflect changes in woody biomass over decadal periods [71,75]. In contrast, indices derived from the NIR and SWIR regions, including NDII, NBR, MSI, and SATVI, showed more pronounced relationships with Δ AGB than greenness-based indices [76], which likely makes them more informative for capturing multi-year biomass dynamics at the ecodistrict scale. However, index performance varied among ecodistricts, echoing broader findings that no single spectral index provides universally robust predictions across forest types [77,78].

Hexagonal binned scatterplots (Figure A6) reveal dense elliptical point clouds centred near zero change, with substantial overlap among gain, loss, and no change classes, confirming that spectral–biomass relationships are nonlinear and confounded by multiple drivers [79]. To illustrate these relationships more directly at the plot level, Figure 14a,b presents two representative examples, NDVI (Figure 14a) and NDII (Figure 14b), based on 100 randomly selected non-fire plot intervals for visual illustration only. For each sampled interval, Landsat July–August annual medians were used to calculate Theil–Sen slopes of spectral indices over the interval period, which were then compared against observed annual Δ AGB. These illustrative plots are not intended to replace the full correlation analysis, but rather to clarify how spectral trajectories align with plot-level biomass change classes. NDVI shows substantial overlap among gain, loss, and no change classes, indicating that greenness trends alone have limited ability to distinguish the direction of biomass change. By contrast, NDII provides clearer directional separation, with positive slopes more frequently associated with biomass gains and near-zero or negative slopes more often associated with losses.

However, these patterns should not be interpreted deterministically. This does not imply that NDII uniquely identifies gain or loss conditions; rather, it indicates that indices sensitive to canopy moisture and NIR–SWIR contrast are generally more informative for representing the direction of Δ AGB change than greenness-only metrics. More broadly, the additional charts for NBR, MSI, and SATVI show the same general tendency, although with varying degrees of overlap and dispersion. These results indicate that spectral–biomass relationships are nonlinear and influenced by multiple interacting factors, including stand structure, recovery stage, disturbance legacy, and compositional change [79]. Consequently, spectral indices should be interpreted as complementary indicators of different ecological processes rather than as direct one-to-one measures of biomass change. Their greatest value emerges when they are integrated within a multi-index, non-linear modelling framework. This interpretation is consistent with [20], who showed that spectral greenness trends in boreal forests are often influenced by disturbance and post-disturbance recovery dynamics.

Together, these results suggest that integrating plot-based measurements with remote sensing is essential to disentangle structural biomass change from spectral signals, which may otherwise be confounded by disturbance–recovery processes.

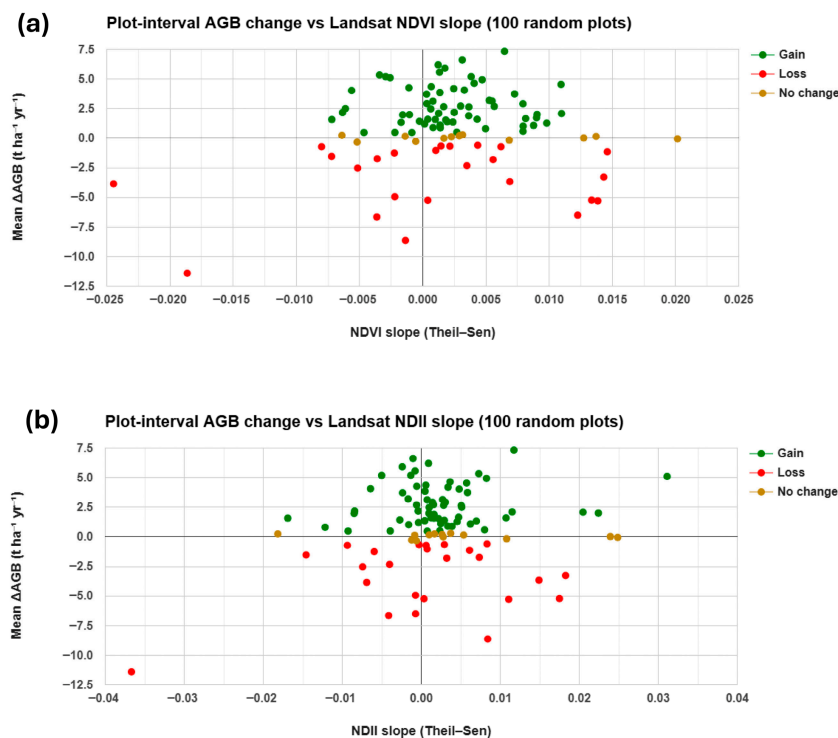


Figure 14. Illustrative plot-level relationships between annual Δ AGB and Landsat spectral trends for two representative indices, (a) NDVI and (b) NDII, based on 100 randomly selected non-fire plot intervals. For each interval, Theil–Sen slopes were calculated from July–August annual median Landsat composites over the interval period. Points are coloured by observed Δ AGB class (gain, loss, or no change).

4.4. Study Limitations and Priorities for Future Work

Our analysis is constrained by 7 ecodistricts (e.g., 364, 369, 384, 394, 395, 396, 1028) that currently have only a single remeasurement period, corresponding to regions classified as “insufficient data” (grey) in Figure 9b, which provides a temporal snapshot rather than robust multi-year Δ AGB estimates. Continued remeasurement of existing plots in these ecodistricts is therefore required before they can fully support long-term trend analyses. In addition, some ecodistricts that met our minimum data filter remain weakly sampled, resulting in wider CI and reduced precision. Increasing plot numbers and remeasurement frequency in these sparsely sampled units should be a clear priority for future work.

An additional limitation is that our remote-sensing predictors were derived from a single-sensor Landsat-7 time series. As described in Section 2.4, this choice was made to ensure radiometric consistency across the study period and to mitigate Landsat-7 SLC-off artefacts using multi-temporal compositing and spatial filtering. While reliance on a single sensor may reduce temporal data density compared to harmonized multi-sensor approaches, it avoids potential biases introduced by cross-sensor differences in spectral response and calibration [25]. Importantly, because our analysis focuses on relative spectral changes and their relationship with Δ AGB, rather than absolute reflectance values, this choice is unlikely to affect the main conclusions. Nevertheless, future research could extend this framework by incorporating harmonized Landsat 5/7/8/9 time series [20,80–82] and additional data sources such as LiDAR (e.g., GEDI) and SAR-based biomass products to further improve temporal coverage and cross-validation at broader spatial scales [83–85].

5. Conclusions

This study demonstrates that integrating forest permanent sample plots with Landsat time series can provide a measurement-constrained framework for assessing non-fire-related AGB change at the ecodistrict scale in the eastern Canadian boreal forest. Across 59 plot-supported ecodistricts (63% of the study area), biomass dynamics are dominated by gains: 37 ecodistricts, accounting for 55% of the plot-supported area, exhibit significant increases, while 20 ecodistricts (39% of this area) show no significant change and only 2 ecodistricts exhibit losses. These plot-based results indicate that, in the absence of wildfire, a substantial portion of the boreal forest functions as a net biomass sink, with additional potential for carbon uptake in ecodistricts classified as no change, where targeted management of non-fire disturbance processes (e.g., harvesting, insect outbreaks, and windthrow), could shift these systems toward net biomass gains. At the full spatial extent, the Landsat-based wall-to-wall assessment suggests a predominance of biomass gains. A total of 61 ecodistricts (71.8%) are classified as gains, while 11 ecodistricts (12.9%) show losses, and 13 ecodistricts (15.3%) exhibit no detectable change. Disturbance composition analyses further indicated that biomass declines were more commonly associated with defoliation-dominated disturbance regimes, whereas harvested regions often exhibited neutral or recovering longer-term biomass trajectories. However, predicted losses in sparsely sampled northern ecodistricts should be interpreted cautiously because they rely more heavily on model extrapolation than direct plot support.

Strong data-density effects on confidence intervals and systematic differences between plot-based and satellite-derived indicators underscore the importance of sampling design, spatial representativeness, and scale when extrapolating from plots to landscapes. Moisture- and disturbance-sensitive indices carried the clearest spectral signal of Δ AGB, and the XGBoost ensemble captured biomass gains more consistently than losses, but severe losses remained under-predicted, reflecting persistent challenges in modelling low-frequency, high-impact disturbance processes.

Overall, the convergence between plot-based and Landsat-derived evidence indicates that much of the eastern Canadian boreal forest exhibits either no significant change or increasing biomass trajectories. The resulting gain–loss–no change framework provides a transparent, spatially explicit basis for greenhouse-gas reporting and for prioritizing monitoring and management in ecodistricts with persistent departures from stability. Future work should extend this framework to harmonized multi-sensor time series, improve representation in data-sparse regions, and explicitly attribute Δ AGB to specific non-wildfire disturbance agents, while complementing these analyses with wildfire-driven dynamics.

Author Contributions: Conceptualization, H.M.M. and C.P.; methodology, H.M.M. and C.P.; software, H.M.M.; formal analysis, H.M.M., C.P., J.C. and D.K.; data curation, H.M.M.; writing—original draft preparation, H.M.M.; writing—review and editing, H.M.M., J.C., D.K. and C.P.; supervision, C.P.; project administration, H.M.M. and C.P.; funding acquisition, C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fonds de recherche du Québec—Nature et technologies (FRQNT), doctoral research scholarship, grant number B2X-336235, and by the Natural Sciences and Engineering Research Council of Canada (NSERC), grant number 371706.

Data Availability Statement: The data used in this study originate from multiple sources with different access conditions. Burned area data were obtained from the National Burned Area Composite (NBAC) dataset, which is publicly available from the Canadian Wildland Fire Information System (CWFIS) DataMart: <https://cwfis.cfs.nrcan.gc.ca/datamart/metadata/nbac> (accessed on 4 May 2026). Permanent sample plot (PSP) data for Québec are publicly available from the Données Québec open data portal: <https://www.donneesquebec.ca/recherche/fr/dataset/placettes-echantillons->

[permanentes-1970-a-aujourd-hui](#) (accessed on 4 May 2026). PSP data for Ontario were obtained from the Ontario Ministry of Natural Resources and Forestry (MNRF), Growth and Yield Program, under a data-sharing agreement. These data are not publicly available and cannot be redistributed by the authors. Access may be granted upon reasonable request and with permission from the Ontario MNRF. Remote sensing datasets, including Landsat-derived vegetation indices and the North American Land Cover dataset (NALCMS), were accessed via the Google Earth Engine (GEE) platform (<https://earthengine.google.com/>), where they are freely available. Derived datasets and analysis code supporting the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to thank Compute Canada for providing access to high-performance computing resources, including GPU-accelerated systems, which facilitated the training and evaluation of models on large-scale datasets. The authors also acknowledge the Ontario Ministry of Natural Resources and Forestry Growth and Yield Program for providing access to permanent sample plot data used in this study. During the preparation of this manuscript, the first author used ChatGPT (OpenAI, GPT-4) to assist with language editing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GEE	Google Earth Engine
MLP	Multilayer Perceptron
NALCMS	North American Land Cover Monitoring System
PSPs	Permanent Sample Plots
PGPs	Permanent Growth Plots
PIPs	Permanent Inventory Plots
XGBoost	eXtreme Gradient Boosting
Δ AGB	Above-ground biomass change
CI	Confidence Interval
QC	Quality Control
NFI	National Forest Inventory

Appendix A

Table A1. Summary of the permanent plot inventory for the eastern boreal Québec–Ontario study area (1999–2024), reporting total plot-level census observations by plot network (PGP, PSP, PIP), the distribution of sampled plot areas (ha), and the frequency of plot remeasurement intervals for each province and the combined domain.

		Quebec	Ontario			Total
			PGP	PSP	PIP	
The total number of unique sample plots		7585	3234	655	892	12,366
Sample plot area (ha)	0.01	0		1		1
	0.02	0		4		4
	0.04	7585		4673		12,258
	0.05	0		14		14
	0.06	0		38		38
	0.09	0		31		31
	0.1	0		20		20

Table A1. Cont.

		Quebec	Ontario	Total
Frequency of sample plot remeasurement intervals	1 time	1248	1975	3223
	2 times	6090	1551	7641
	3 times	216	1108	1324
	4 times	31	147	178

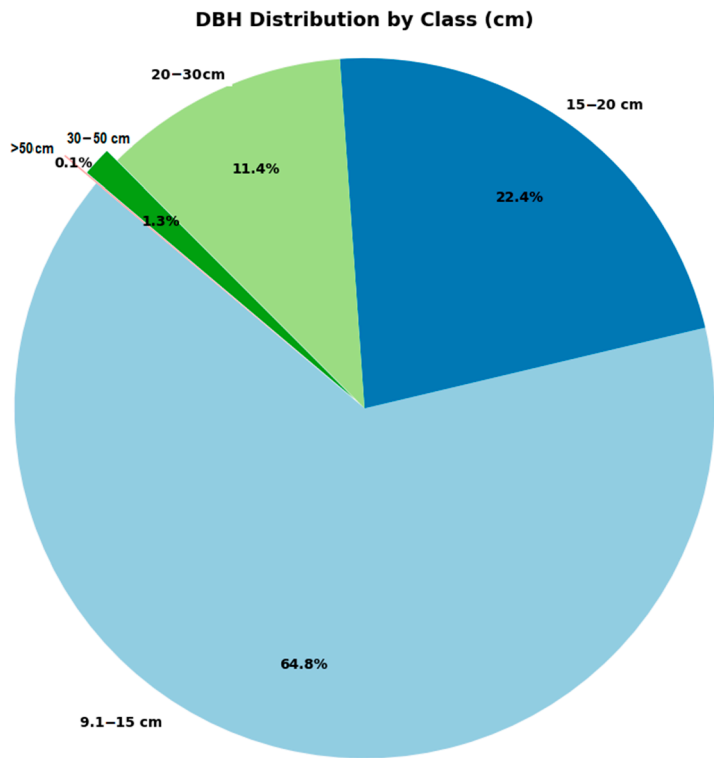


Figure A1. Diameter at breast height (DBH) distribution by class (cm) in the study area.

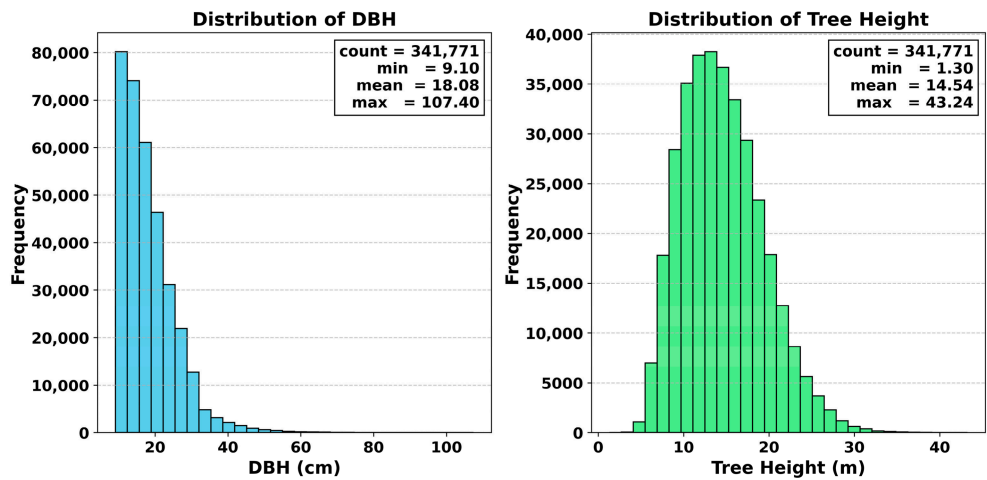


Figure A2. Frequency distribution of DBH and tree height for the period of 1980–2021 based on the Ontario dataset.

Table A2. Summary of the MLP model configuration, training strategy, and selected hyperparameters.

Component	Specification
Model type	Multi-layer Perceptron (MLP)
Objective	predict tree height (m)
Input variables	DBH (cm) and species (allometric group)
Output	Continuous height (m)
Architecture	Fully connected feedforward network
Hidden Layer Sizes	[(256, 128, 64), (128, 64, 32), (64, 32)]
Alpha (Regularization Strength)	[0.0001, 0.001, 0.01]
Activation Functions	['ReLU']
Solvers (Optimization Algorithms)	['Adam', 'SGD']
Loss function	Mean Squared Error (MSE)
Train/validation/Test split	70%/15%/15%
Performance	R ² ; MAE; RMSE

Table A3. Species- and component-specific coefficients ($\alpha_{c,s}$, $\beta_{c,s}$, $\theta_{c,s}$) of the allometric power-law model used to estimate above-ground biomass (AGB) components, where c denotes biomass components: wood (wo), bark (ba), branches (br), and foliage (fo) [48,49]. Rows with a white background represent softwood species, whereas rows with a grey background represent hardwood species.

Species	$\alpha_{wo,s}$	$\beta_{wo,s}$	$\theta_{wo,s}$	$\alpha_{ba,s}$	$\beta_{ba,s}$	$\theta_{ba,s}$	$\alpha_{br,s}$	$\beta_{br,s}$	$\theta_{br,s}$	$\alpha_{fo,s}$	$\beta_{fo,s}$	$\theta_{fo,s}$
Black Spruce	0.0335	1.7389	0.9835	0.0132	1.7657	0.5775	0.0405	3.1917	−1.3674	0.2078	2.5517	−1.3453
Balsam Fir	0.0294	1.8357	0.8640	0.0053	2.0876	0.5842	0.0117	3.5097	−1.3006	0.1245	2.5230	−1.1230
Eastern Hemlock	0.0257	1.9277	0.8576	0.0118	1.9893	0.4700	0.0215	2.6553	−0.4682	0.1471	2.0108	−0.6080
Eastern Red Cedar	0.0520	1.7731	0.7054	0.0283	1.7079	0.0000	0.0219	2.3585	0.0000	0.2575	2.5136	−1.5565
Eastern White Cedar	0.0295	1.7026	0.9428	0.0076	1.7861	0.6132	0.0501	2.5165	−0.8774	0.0813	2.2180	−0.7907
Eastern White Pine	0.0170	1.7779	1.1370	0.0069	1.6589	0.9582	0.0184	3.1968	−1.0876	0.0584	2.2389	−0.5968
Jack Pine	0.0199	1.6883	1.2456	0.0141	1.5994	0.5957	0.0185	3.0584	−0.9816	0.0325	1.7879	0.0000
Red Pine	0.0106	1.7725	1.3285	0.0277	1.5192	0.4645	0.0125	3.3865	−1.1939	0.0731	2.3439	−0.7378
Red Spruce	0.0143	1.6441	1.4065	0.0274	2.0188	0.0000	0.0005	3.3136	0.0000	0.0106	2.2709	0.0000
White Spruce	0.0265	1.7952	0.9733	0.0124	1.6962	0.6489	0.0325	2.8573	−0.9127	0.2020	2.3802	−1.1103
Other softwood species	0.0276	1.6868	1.0953	0.0101	1.8486	0.5525	0.0313	2.9974	−1.0383	0.1379	2.3981	−1.0418
Eastern Cottonwood	0.0051	1.0697	2.2748	0.0009	1.3061	2.0109	0.0131	2.5760	0.0000	0.0224	1.8368	0.0000
Balsam Poplar	0.0117	1.7757	1.2555	0.0180	1.8131	0.5144	0.0112	3.0861	−0.7164	0.0617	1.8615	−0.5375
American Basswood	0.0168	1.9844	0.8989	0.0057	1.5881	1.1472	0.0039	2.0084	0.8588	0.0147	1.8300	0.0000
American Beech	0.0432	2.0378	0.7000	0.0049	1.9057	0.6770	0.0355	2.3749	0.0000	0.0452	1.5567	0.0000
Black Ash	0.0306	2.1836	0.5740	0.0897	2.2634	−0.567	0.0994	2.1630	−0.4809	0.0124	1.0325	0.8747
Black Cherry	0.0181	1.7013	1.3057	0.0101	1.5956	0.9190	0.0005	2.8004	0.8603	0.1976	1.4421	−0.5264
Gray Birch	0.0295	1.9064	0.9139	0.0148	1.8433	0.5021	0.0150	3.0347	−0.7629	0.0455	2.6447	−1.4955
Bitternut Hickory	0.0139	1.5913	1.5080	0.0081	1.4943	1.1324	0.0050	3.0463	0.0000	0.0121	2.0865	0.0000
Large-tooth Aspen	0.0128	2.0633	0.9516	0.0240	2.3055	0.0000	0.0131	3.1274	−0.8379	0.0382	2.1673	−0.6842
Red Maple	0.0315	2.0342	0.7485	0.0283	2.0907	0.0000	0.0225	2.4106	0.0000	0.0571	1.4898	0.0000
Silver Maple	0.0274	1.7126	1.1086	0.0123	1.8250	0.5010	0.0543	3.7343	−1.6497	6.6808	2.1092	−2.1697

Table A3. Cont.

Species	$\alpha_{wo,s}$	$\beta_{wo,s}$	$\theta_{wo,s}$	$\alpha_{ba,s}$	$\beta_{ba,s}$	$\theta_{ba,s}$	$\alpha_{br,s}$	$\beta_{br,s}$	$\theta_{br,s}$	$\alpha_{fo,s}$	$\beta_{fo,s}$	$\theta_{fo,s}$
Sugar Maple	0.0301	2.0313	0.8171	0.0103	1.7111	0.8509	0.0661	2.5940	−0.4933	2.5019	2.4527	−2.3008
Trembling Aspen	0.0142	1.9389	1.0572	0.0063	2.0819	0.6617	0.0137	2.9270	−0.6221	0.0270	1.6183	0.0000
White Ash	0.0224	1.7438	1.1899	0.0126	1.6456	0.7893	0.0354	2.3046	0.0000	0.0195	1.0509	0.7836
American Elm	0.0207	2.2276	0.6488	0.0078	2.4540	0.0000	0.0393	2.1880	0.0000	0.0516	1.4511	0.0000
White Oak	0.0442	1.6818	1.0310	0.0308	1.7479	0.3504	0.0022	2.0165	1.3953	0.0053	1.2822	1.1323
Yellow Birch	0.0259	1.9044	0.9715	0.0069	2.0834	0.5371	0.0325	2.3851	0.0000	0.1683	1.2764	0.0000
Black Maple	0.0315	2.0342	0.7485	0.0283	2.0907	0.0000	0.0225	2.4106	0.0000	0.0571	1.4898	0.0000
Bur Oak	0.0285	1.8501	1.0204	0.0326	1.8100	0.4153	0.0013	3.0637	0.3153	0.0582	1.5438	0.0000
Shagbark Hickory	0.0139	1.5913	1.5080	0.0081	1.4943	1.1324	0.0050	3.0463	0.0000	0.0121	2.0865	0.0000
Swamp White Oak	0.0442	1.6818	1.0310	0.0308	1.7479	0.3504	0.0022	2.0165	1.3953	0.0053	1.2822	1.1323
Manitoba Maple	0.0315	2.0342	0.7485	0.0283	2.0907	0.0000	0.0225	2.4106	0.0000	0.0571	1.4898	0.0000
Paper Birch	0.0338	2.0702	0.6876	0.0080	1.9754	0.6659	0.0257	3.1754	−0.9417	0.1415	2.3074	−1.1189
Other hardwood species	0.0353	2.0249	0.7048	0.0090	1.8677	0.7144	0.0448	2.6855	−0.5911	0.0869	1.8541	−0.5491

Table A4. Landsat vegetation index (VI) metrics.

Vegetation Index	Formulas	Biophysical Meaning	Reference
Normalized Difference Vegetation Index (NDVI)	$(NIR - R)/(NIR + R)$	photosynthetic activity and canopy density	[86–88]
Enhanced Vegetation Index (EVI)	$2.5 \times (NIR - R)/(NIR + 6R - 7.5B + 1)$		
Normalized Difference Infrared Index (NDII)	$(NIR - SWIR1)/(NIR + SWIR1)$	canopy and soil moisture and fire-related structural change	[89,90]
Normalized Burn Ratio (NBR)	$(NIR - SWIR2)/(NIR + SWIR2)$		
Moisture Stress Index (MSI)	$SWIR1/NIR$	emphasize water stress	[88,91]
Normalized Difference Water Index (NDWI)	$(G - NIR)/(G + NIR)$		
Soil-Adjusted Total Vegetation Index (SATVI)	$((SWIR1 - R)/(SWIR1 + R + 0.5)) \times 1.5 - (SWIR2/2)$	track senescence and soil/background effects	[92,93]
Plant Senescence Reflectance Index (PSRI)	$(R - G)/NIR$		
Near-Infrared Reflectance of Vegetation (NIRV)	$NDVI \times NIR$	provides a proxy for canopy structure and primary productivity	[94]

Table A5. Tree species composition of the plot inventory, reporting relative abundance (%) by species and functional group, and summarizing the overall contributions of softwoods (81%) and hardwoods (19%) across the study area.

Softwood: 81%			Hardwood: 19%		
Species	Percentage	Type	Species	Percentage < 0.01	Type
Black spruce	35.42	Softwood	Black maple		Hardwood
Balsam fir	19.25		Shagbark hickory		
Jack pine	18.59		Scots pine		Softwood
Paper birch	8.27	Hardwood	European larch		
Trembling aspen	7.37		American hornbeam		
White spruce	4.67	Softwood	Speckled alder		
Eastern white cedar	1.22				
Red maple	0.89	Hardwood	Bur oak		Hardwood
American larch	0.72	Softwood	Rock elm		
Sugar maple	0.56	Hardwood	Swamp white oak		
Red pine	0.46	Softwood	Hybrid poplar		
Yellow birch	0.35	Hardwood	Slippery elm		
Balsam poplar	0.28		Cucumber tree		
Eastern white pine	0.20	Softwood	Big-leaf linden		
Red spruce	0.13		Hybrid larch		Softwood
American beech	0.12	Hardwood	Sorbus species		
Eastern hemlock	0.11	Softwood	Chokecherry		
Pin cherry	0.09	Hardwood	Sassafras		Hardwood
Black ash	0.06		Manitoba maple		
Salix tree species	0.04		Carolina poplar		
American mountain-ash	0.03		Sweet pignut hickory		
Large-tooth aspen	0.03		Sweet cherry		
Northern red oak	0.03		Eastern cottonwood		
American basswood	0.03		Bay-leaved willow		
White ash	0.02		Populus species		
Ironwood	0.02		Black walnut		
Striped maple	0.02		Amelanchier species		
Northern mountain-ash			Tulip tree		Hardwood
Black cherry			Buckthorn		
Gray birch		Hardwood	Eastern red cedar		Softwood
American elm	0.01		Unknown tree species		
Silver maple		Hardwood	Sweet chestnut		Hardwood
Green ash			Chinquapin oak		
Mountain maple			Black oak		
Bitternut hickory			Unknown hardwood		
White oak			Black oak		
Pitch pine		Softwood	White willow		
Norway spruce	0.01		Common hackberry		

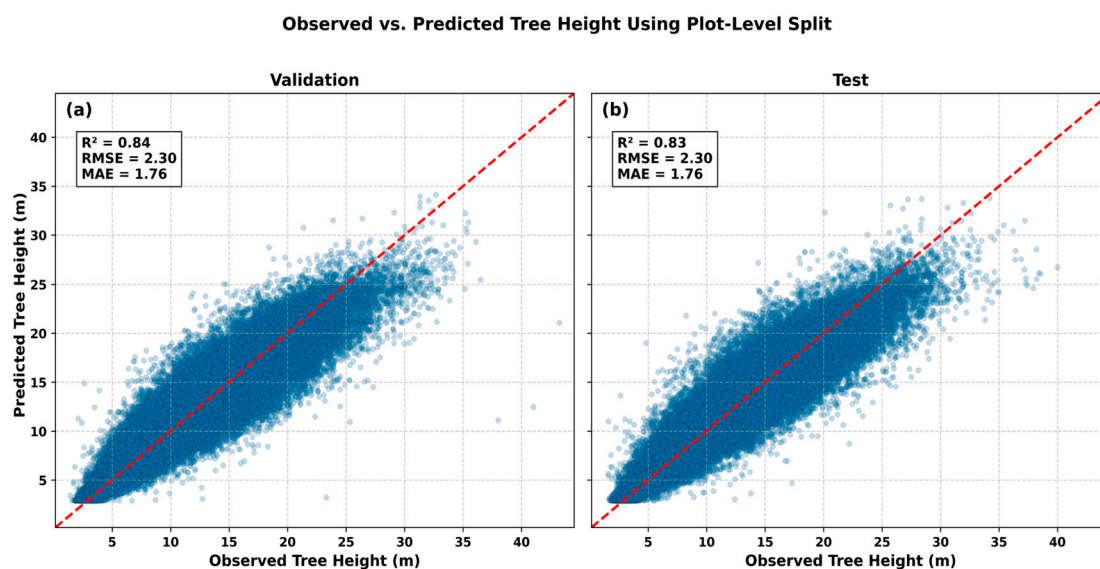


Figure A3. Observed versus predicted tree height for (a) validation and (b) test datasets using a plot-level split, where all trees from the same plot were assigned to a single subset. The results show strong agreement ($R^2 \approx 0.83$; RMSE ≈ 2.30 m) and are nearly identical to those obtained using tree-level partitioning, indicating minimal influence of data leakage.

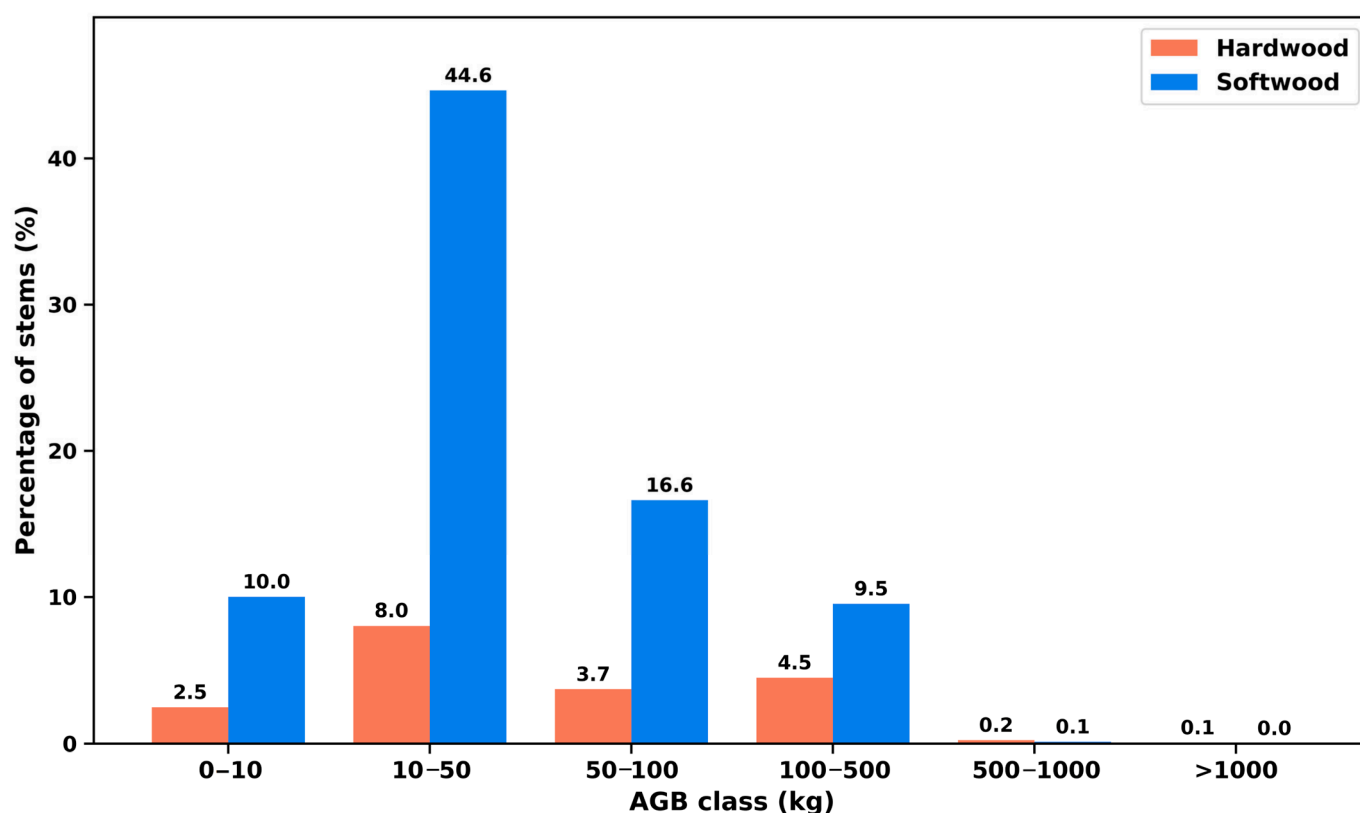


Figure A4. Distribution of individual-tree above-ground biomass (AGB, kg) by biomass class for hardwood and softwood stems across all plots.

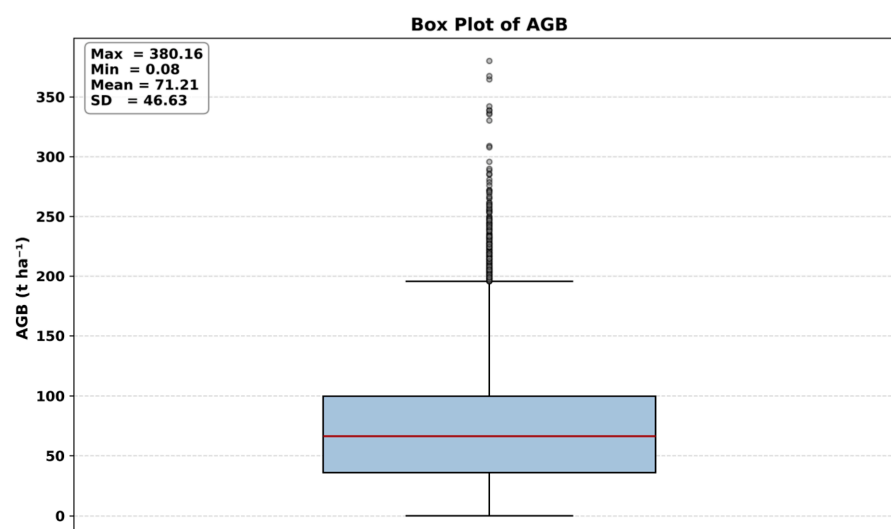


Figure A5. Distribution of plot-level above-ground biomass (AGB; t ha^{-1} ; $N = 32,380$ plots measurements). The boxplot shows the median, interquartile range, non-outlier range, and outliers.

Table A6. Sensitivity of ecodistrict-level ΔAGB estimates to tree height uncertainty, assessed using Monte Carlo simulations ($n = 100$). Reported metrics include the number of ecodistricts with classification changes (SIG), as well as changes in ΔAGB magnitude and confidence interval (CI) width.

Metric	Value
Number of ecodistricts	59
Mean number of SIG changes per run	0.16
Maximum number of SIG changes per run	1
Number of ecodistricts with any SIG change	4
Number with probability of SIG change $\geq 5\%$	1
Number with probability of SIG change $\geq 10\%$	1
Mean absolute interval-level ΔAGB change ($\text{t ha}^{-1} \text{ yr}^{-1}$)	0.17
Mean absolute change in ecodistrict mean ΔAGB ($\text{t ha}^{-1} \text{ yr}^{-1}$)	0.025
Mean absolute change in CI width ($\text{t ha}^{-1} \text{ yr}^{-1}$)	0.04

Table A7. Summary of model configurations, validation strategy, and tuning procedures for XGBoost and benchmark models.

Component	Specification
Primary model	XGBoost (gradient boosting)
Benchmark models	Random Forest regression; Ridge regression
Predictors	Spectral vegetation indices at the start of each interval; annualized changes in spectral indices; interval duration; interval mid-year
Response variable	Annualized AGB change, ΔAGB ($\text{t ha}^{-1} \text{ yr}^{-1}$)
Validation design	Five-fold grouped cross-validation
Preprocessing	Outlier control, imputation, and Ridge standardization
XGBoost tuning	Tree depth (4–6), learning rate (0.03–0.10), minimum child weight (1–5), row and predictor subsampling (0.70–0.90), L1/L2 regularization, and early stopping

Table A7. Cont.

Component	Specification
RF tuning	Number of trees (300–500), maximum tree depth (10–unlimited), minimum terminal node size (1–2), and predictors considered per split (square-root rule)
Ridge tuning	Regularization strength, α (0.01–100)
Hyperparameter tuning	Grid search over a compact parameter set
Loss function	Mean Squared Error (MSE)
Performance metrics	R^2 ; MAE; RMSE; Directional Accuracy, precision, recall, F1-score, false negative rate

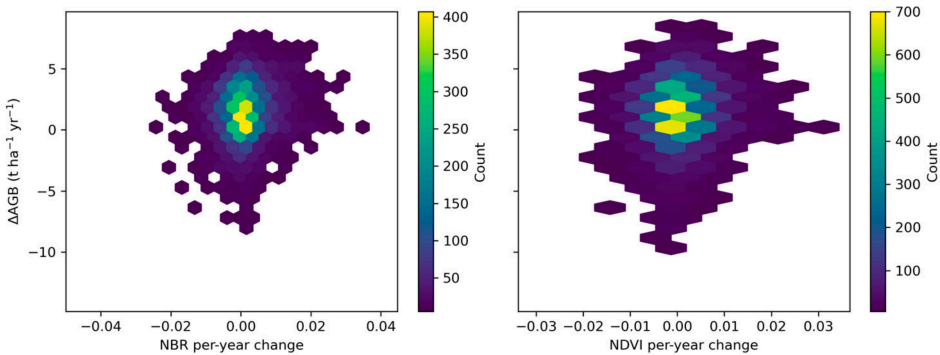


Figure A6. Relationships between annual biomass change and per-year spectral index change at plots.

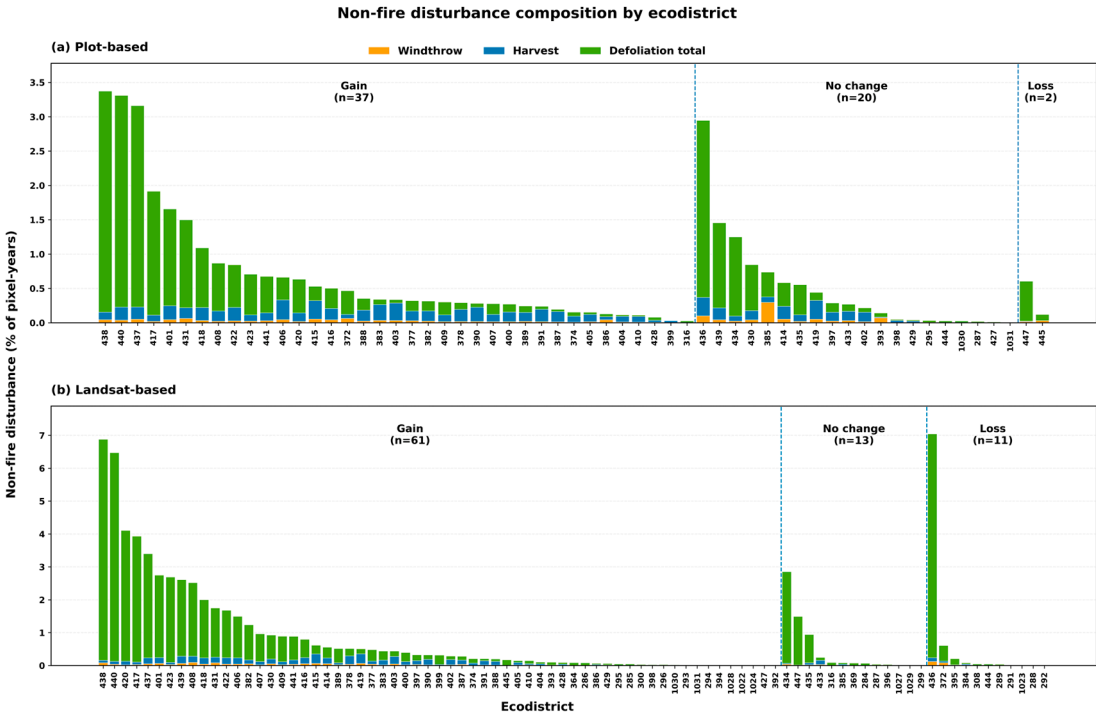


Figure A7. Ecodistrict-level composition of non-fire disturbances (defoliation, harvest, windthrow) for biomass-change classes (gain, no change, loss), derived from (a) plot-based Δ AGB intervals and (b) Landsat-based disturbance classifications. Blue dashed vertical lines separate the gain, no-change, and loss biomass-change classes. Ecodistricts are ordered by disturbance magnitude, highlighting strong spatial heterogeneity and the dominance of defoliation across most regions.

Appendix B

For each remeasurement interval i within ecodistrict e , the annualized change in AGB is defined as:

$$\Delta AGB_{yr,i} = \left(AGB_{p,end,i}^{(t\ ha^{-1})} - AGB_{p,start,i}^{(t\ ha^{-1})} \right) / \Delta t_i \quad (A1)$$

where Δt_i is the length (years) of interval i .

For an ecodistrict e with n_e contributing intervals, the mean annual AGB change is:

$$\Delta AGB_{yr,e} = \frac{1}{n_e} \sum_{i \in e} \Delta AGB_{yr,i} \quad (A2)$$

and the sample variance is

$$S_e^2 = \frac{1}{n_e - 1} \sum_{i \in e} (\Delta AGB_{yr,i} - \Delta AGB_{yr,e})^2. \quad (A3)$$

Because interval-level observations within an ecodistrict are not fully independent, we account for spatial and temporal correlation using an effective sample size:

$$N_{\text{eff},e} = \frac{n_e}{1 + (n_e - 1)\rho_e} \quad (A4)$$

where ρ_e is an intraclass correlation coefficient (ICC). The ICC is estimated by decomposing total variance into between-ecodistrict ($\sigma_{\text{between}}^2$) and within-ecodistrict (σ_{within}^2) components:

$$\rho_e = \frac{\sigma_{\text{between}}^2}{\sigma_{\text{between}}^2 + \sigma_{\text{within}}^2} \quad (A5)$$

the effective standard error of the mean is then:

$$SE_{\text{eff},e} = \frac{s_e}{\sqrt{N_{\text{eff},e}}} \quad (A6)$$

Let $t_{0.975,\nu_e}$ be the two-sided 97.5% t-quantile with $\nu_e = \max(1, N_{\text{eff},e} - 1)$ degrees of freedom. The 95% confidence interval (CI) for the $\Delta AGB_{yr,e}$ is:

$$CI_{95,e} = [\Delta AGB_{yr,e} - t_{0.975,\nu_e} \cdot SE_{\text{eff},e}, \Delta AGB_{yr,e} + t_{0.975,\nu_e} \cdot SE_{\text{eff},e}]. \quad (A7)$$

Ecodistricts are classified based on the confidence interval as:

- “Gain” if the lower bound of $CI_{95,e} > 0$;
- “Loss” if the upper bound of $CI_{95,e} < 0$;
- “No change” otherwise (i.e., if the CI includes zero);
- “Insufficient” if $n_e < 2$, $S_e^2 = 0$, or $SE_{\text{eff},e}$ is undefined.

These estimators follow standard design-based inference for clustered ecological data and explicitly account for within-ecodistrict dependence through the effective sample size adjustment.

References

1. Bonan, G.B. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science* **2008**, *320*, 1444–1449. [CrossRef] [PubMed]
2. Bradshaw, C.J.A.; Warkentin, I.G. Global Estimates of Boreal Forest Carbon Stocks and Flux. *Glob. Planet. Change* **2015**, *128*, 24–30. [CrossRef]
3. Gauthier, S.; Bernier, P.; Kuuluvainen, T.; Shvidenko, A.Z.; Schepaschenko, D.G. Boreal Forest Health and Global Change. *Science* **2015**, *349*, 819–822. [CrossRef]

4. Pan, Y.; Birdsey, R.A.; Fang, J.; Houghton, R.; Kauppi, P.E.; Kurz, W.A.; Phillips, O.L.; Shvidenko, A.; Lewis, S.L.; Canadell, J.G.; et al. A Large and Persistent Carbon Sink in the World's Forests. *Science* **2011**, *333*, 988–993. [[CrossRef](#)] [[PubMed](#)]
5. Kurz, W.A.; Dymond, C.C.; White, T.M.; Stinson, G.; Shaw, C.H.; Rampley, G.J.; Smyth, C.; Simpson, B.N.; Neilson, E.T.; Trofymow, J.A.; et al. CBM-CFS3: A Model of Carbon-Dynamics in Forestry and Land-Use Change Implementing IPCC Standards. *Ecol. Modell.* **2009**, *220*, 480–504. [[CrossRef](#)]
6. Stinson, G.; Kurz, W.A.; Smyth, C.E.; Neilson, E.T.; Dymond, C.C.; Metsaranta, J.M.; Boisvenue, C.; Rampley, G.J.; Li, Q.; White, T.M.; et al. An Inventory-Based Analysis of Canada's Managed Forest Carbon Dynamics, 1990 to 2008. *Glob. Change Biol.* **2011**, *17*, 2227–2244. [[CrossRef](#)]
7. Brandt, J.P. The Extent of the North American Boreal Zone. *Environ. Rev.* **2009**, *17*, 101–161. [[CrossRef](#)]
8. Matasci, G.; Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W.; Zald, H.S.J. Large-Area Mapping of Canadian Boreal Forest Cover, Height, Biomass and Other Structural Attributes Using Landsat Composites and Lidar Plots. *Remote Sens. Environ.* **2018**, *209*, 90–106. [[CrossRef](#)]
9. Beaudoin, A.; Bernier, P.Y.; Guindon, L.; Villemaire, P.; Guo, X.J.; Stinson, G.; Bergeron, T.; Magnussen, S.; Hall, R.J. Mapping Attributes of Canada's Forests at Moderate Resolution through KNN and MODIS Imagery. *Can. J. For. Res.* **2014**, *44*, 521–532. [[CrossRef](#)]
10. Nguyen, T.H.; Jones, S.; Soto-Berelov, M.; Haywood, A.; Hislop, S. Landsat Time-Series for Estimating Forest Aboveground Biomass and Its Dynamics across Space and Time: A Review. *Remote Sens.* **2020**, *12*, 98. [[CrossRef](#)]
11. Zhang, J.; Zhou, C.; Zhang, G.; Yang, Z.; Pang, Z.; Luo, Y. A Novel Framework for Forest Above-Ground Biomass Inversion Using Multi-Source Remote Sensing and Deep Learning. *Forests* **2024**, *15*, 456. [[CrossRef](#)]
12. Guindon, L.; Manka, F.; Correia, D.L.P.; Villemaire, P.; Smiley, B.; Bernier, P.; Gauthier, S.; Beaudoin, A.; Boucher, J.; Boulanger, Y. A New Approach for Spatializing the Canadian National Forest Inventory (SCANFI) Using Landsat Dense Time Series. *Can. J. For. Res.* **2024**, *54*, 793–815. [[CrossRef](#)]
13. Peng, C.; Ma, Z.; Lei, X.; Zhu, Q.; Chen, H.; Wang, W.; Liu, S.; Li, W.; Fang, X.; Zhou, X. A Drought-Induced Pervasive Increase in Tree Mortality across Canada's Boreal Forests. *Nat. Clim. Change* **2011**, *1*, 467–471. [[CrossRef](#)]
14. Ma, Z.; Peng, C.; Zhu, Q.; Chen, H.; Yu, G.; Li, W.; Zhou, X.; Wang, W.; Zhang, W. Regional Drought-Induced Reduction in the Biomass Carbon Sink of Canada's Boreal Forests. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2423–2427. [[CrossRef](#)] [[PubMed](#)]
15. Zhang, X.; Lei, Y.; Ma, Z.; Kneeshaw, D.; Peng, C. Insect-Induced Tree Mortality of Boreal Forests in Eastern Canada under a Changing Climate. *Ecol. Evol.* **2014**, *4*, 2384–2394. [[CrossRef](#)] [[PubMed](#)]
16. Liu, Q.; Peng, C.; Schneider, R.; Cyr, D.; McDowell, N.G.; Kneeshaw, D. Drought-Induced Increase in Tree Mortality and Corresponding Decrease in the Carbon Sink Capacity of Canada's Boreal Forests from 1970 to 2020. *Glob. Change Biol.* **2023**, *29*, 2274–2285. [[CrossRef](#)]
17. Ameray, A.; Cavard, X.; Cyr, D.; Valeria, O.; Girona, M.M.; Bergeron, Y. One Century of Carbon Dynamics in the Eastern Canadian Boreal Forest under Various Management Strategies and Climate Change Projections. *Ecol. Modell.* **2024**, *498*, 110894. [[CrossRef](#)]
18. Smyth, C.E.; Metsaranta, J.; Tompalski, P.; Hararuk, O.; Le Noble, S. 10-Year Progress on Forest Carbon Research in Canada. *Environ. Rev.* **2024**, *32*, 611–637. [[CrossRef](#)]
19. Wang, J.A.; Baccini, A.; Farina, M.; Randerson, J.T.; Friedl, M.A. Disturbance Suppresses the Aboveground Carbon Sink in North American Boreal Forests. *Nat. Clim. Change* **2021**, *11*, 435–441. [[CrossRef](#)]
20. Sulla-Menashe, D.; Woodcock, C.E.; Friedl, M.A. Canadian Boreal Forest Greening and Browning Trends: An Analysis of Biogeographic Patterns and the Relative Roles of Disturbance versus Climate Drivers. *Environ. Res. Lett.* **2018**, *13*, 014007. [[CrossRef](#)]
21. Hember, R.A.; Kurz, W.A.; Coops, N.C. Increasing Net Ecosystem Biomass Production of Canada's Boreal and Temperate Forests despite Decline in Dry Climates. *Glob. Biogeochem. Cycles* **2017**, *31*, 134–158. [[CrossRef](#)]
22. D'Orangeville, L.; Houle, D.; Duchesne, L.; Phillips, R.P.; Bergeron, Y.; Kneeshaw, D. Beneficial Effects of Climate Warming on Boreal Tree Growth May Be Transitory. *Nat. Commun.* **2018**, *9*, 3213. [[CrossRef](#)]
23. Girardin, M.P.; Bouriaud, O.; Hogg, E.H.; Kurz, W.; Zimmermann, N.E.; Metsaranta, J.M.; De Jong, R.; Frank, D.C.; Esper, J.; Büntgen, U.; et al. No Growth Stimulation of Canada's Boreal Forest under Half-Century of Combined Warming and CO₂ Fertilization. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E8406–E8414. [[CrossRef](#)] [[PubMed](#)]
24. Chen, H.Y.H.; Luo, Y.; Reich, P.B.; Searle, E.B.; Biswas, S.R. Climate Change-Associated Trends in Net Biomass Change Are Age Dependent in Western Boreal Forests of Canada. *Ecol. Lett.* **2016**, *19*, 1150–1158. [[CrossRef](#)]
25. Burrell, A.L.; Cooperdock, S.; Potter, S.; Berner, L.T.; Hember, R.; Macander, M.J.; Walker, X.J.; Massey, R.; Foster, A.C.; Mack, M.C.; et al. The Predictability of Near-Term Forest Biomass Change in Boreal North America. *Ecosphere* **2024**, *15*, e4737. [[CrossRef](#)]
26. Itter, M.S.; D'Orangeville, L.; Dawson, A.; Kneeshaw, D.; Duchesne, L.; Finley, A.O. Boreal Tree Growth Exhibits Decadal-Scale Ecological Memory to Drought and Insect Defoliation, but No Negative Response to Their Interaction. *J. Ecol.* **2019**, *107*, 1288–1301. [[CrossRef](#)]

27. Hember, R.A.; Kurz, W.A.; Coops, N.C. Relationships between Individual-Tree Mortality and Water-Balance Variables Indicate Positive Trends in Water Stress-Induced Tree Mortality across North America. *Glob. Change Biol.* **2017**, *23*, 1691–1710. [\[CrossRef\]](#)
28. Gillis, M.D.; Omule, A.Y.; Brierley, T. Monitoring Canada's Forests: The National Forest Inventory. *For. Chron.* **2005**, *81*, 214–221. [\[CrossRef\]](#)
29. Beaudoin, A.; Bernier, P.Y.; Villemaire, P.; Guindon, L.; Guo, X.J. Tracking Forest Attributes across Canada between 2001 and 2011 Using a k Nearest Neighbors Mapping Approach Applied to MODIS Imagery. *Can. J. For. Res.* **2018**, *48*, 85–93. [\[CrossRef\]](#)
30. White, J.C.; Wulder, M.A.; Hermosilla, T.; Coops, N.C.; Hobart, G.W. A Nationwide Annual Characterization of 25 Years of Forest Disturbance and Recovery for Canada Using Landsat Time Series. *Remote Sens. Environ.* **2017**, *194*, 303–321. [\[CrossRef\]](#)
31. Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W. An Integrated Landsat Time Series Protocol for Change Detection and Generation of Annual Gap-Free Surface Reflectance Composites. *Remote Sens. Environ.* **2015**, *158*, 220–234. [\[CrossRef\]](#)
32. Zald, H.S.J.; Wulder, M.A.; White, J.C.; Hilker, T.; Hermosilla, T.; Hobart, G.W.; Coops, N.C. Integrating Landsat Pixel Composites and Change Metrics with Lidar Plots to Predictively Map Forest Structure and Aboveground Biomass in Saskatchewan, Canada. *Remote Sens. Environ.* **2016**, *176*, 188–201. [\[CrossRef\]](#)
33. Matasci, G.; Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W.; Bolton, D.K.; Tompalski, P.; Bater, C.W. Three Decades of Forest Structural Dynamics over Canada's Forested Ecosystems Using Landsat Time-Series and Lidar Plots. *Remote Sens. Environ.* **2018**, *216*, 697–714. [\[CrossRef\]](#)
34. Nguyen, T.D.; Kappas, M. Estimating the Aboveground Biomass of an Evergreen Broadleaf Forest in Xuan Lien Nature Reserve, Thanh Hoa, Vietnam, Using SPOT-6 Data and the Random Forest Algorithm. *Int. J. For. Res.* **2020**, *2020*, 4216160. [\[CrossRef\]](#)
35. Luo, Y.; Chen, H.Y.H.; McIntire, E.J.B.; Andison, D.W. Divergent Temporal Trends of Net Biomass Change in Western Canadian Boreal Forests. *J. Ecol.* **2019**, *107*, 69–78. [\[CrossRef\]](#)
36. Ståhl, G.; Saarela, S.; Schnell, S.; Holm, S.; Breidenbach, J.; Healey, S.P.; Patterson, P.L.; Magnussen, S.; Næsset, E.; McRoberts, R.E.; et al. Use of Models in Large-Area Forest Surveys: Comparing Model-Assisted, Model-Based and Hybrid Estimation. *For. Ecosyst.* **2016**, *3*, 5. [\[CrossRef\]](#)
37. Metsaranta, J.M.; Shaw, C.H.; Kurz, W.A.; Boisvenue, C.; Morken, S. Uncertainty of Inventory-Based Estimates of the Carbon Dynamics of Canada's Managed Forest (1990–2014). *Can. J. For. Res.* **2017**, *47*, 1082–1094. [\[CrossRef\]](#)
38. Peereman, J.; Hogan, J.A.; Lin, T.C. Landscape Representation by a Permanent Forest Plot and Alternative Plot Designs in a Typhoon Hotspot, Fushan, Taiwan. *Remote Sens.* **2020**, *12*, 660. [\[CrossRef\]](#)
39. Wang, Y.; Wang, X.; Ji, P.; Li, H.; Wei, S.; Peng, D. Evaluating Forest Aboveground Biomass Products by Incorporating Spatial Representativeness Analysis. *Remote Sens.* **2025**, *17*, 2898. [\[CrossRef\]](#)
40. Zhang, Y.; Zou, Y.; Wang, Y. Remote Sensing of Forest Above-Ground Biomass Dynamics: A Review. *Forests* **2025**, *16*, 821. [\[CrossRef\]](#)
41. Meimand, H.M.; Chen, J.; Kneeshaw, D.; Bakhtyari, M.; Peng, C. Burned Area Detection in the Eastern Canadian Boreal Forest Using a Multi-Layer Perceptron and MODIS-Derived Features. *Remote Sens.* **2025**, *17*, 2162. [\[CrossRef\]](#)
42. Talbot, J.; Lewis, S.L.; Lopez-Gonzalez, G.; Brien, R.J.W.; Monteagudo, A.; Baker, T.R.; Feldpausch, T.R.; Malhi, Y.; Vanderwel, M.; Araujo Murakami, A.; et al. Methods to Estimate Aboveground Wood Productivity from Long-Term Forest Inventory Plots. *For. Ecol. Manag.* **2014**, *320*, 30–38. [\[CrossRef\]](#)
43. Ecological Stratification Working Group. *A National Ecological Framework for Canada*; Agriculture and Agri-Food Canada; Environment Canada: Ottawa, ON, Canada; Hull, QC, Canada, 1996; Cat. No. A42-65/1996E; ISBN 0-662-24107-X. Available online: <https://sis.agr.gc.ca/cansis/publications/manuals/1996/index.html> (accessed on 4 May 2026).
44. Wester, M.C.; Henson, B.L.; Crins, W.J.; Uhlig, P.W.C.; Gray, P.A. *The Ecosystems of Ontario, Part 2: Ecodistricts*; Ontario Ministry of Natural Resources and Forestry, Science and Research Branch: Peterborough, ON, Canada, 2018.
45. Searle, E.B.; Chen, H.Y.H. Tree Size Thresholds Produce Biased Estimates of Forest Biomass Dynamics. *For. Ecol. Manag.* **2017**, *400*, 468–474. [\[CrossRef\]](#)
46. Chojnacky, D.C.; Heath, L.S.; Jenkins, J.C. Updated Generalized Biomass Equations for North American Tree Species. *Forestry* **2014**, *87*, 129–151. [\[CrossRef\]](#)
47. Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. National-Scale Biomass Estimators for United States Tree Species. *For. Sci.* **2003**, *49*, 12–35. [\[CrossRef\]](#)
48. Lambert, M.C.; Ung, C.H.; Raulier, F. Canadian National Tree Aboveground Biomass Equations. *Can. J. For. Res.* **2005**, *35*, 1996–2018. [\[CrossRef\]](#)
49. Ung, C.H.; Bernier, P.; Guo, X.J. Canadian National Biomass Equations: New Parameter Estimates That Include British Columbia Data. *Can. J. For. Res.* **2008**, *38*, 1123–1132. [\[CrossRef\]](#)
50. Clough, B.J.; Russell, M.B.; Domke, G.M.; Woodall, C.W.; Radtke, P.J. Comparing Tree Foliage Biomass Models Fitted to a Multispecies, Felled-Tree Biomass Dataset for the United States. *Ecol. Modell.* **2016**, *333*, 79–91. [\[CrossRef\]](#)
51. Murat, H.S. A Brief Review of Feed-Forward Neural Networks. *Commun. Fac. Sci. Univ. Ank.* **2006**, *50*, 11–17. [\[CrossRef\]](#)

52. Al Shalabi, L.; Shaaban, Z.; Kasasbeh, B. Data Mining: A Preprocessing Engine. *J. Comput. Sci.* **2006**, *2*, 735–739. [\[CrossRef\]](#)
53. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [\[CrossRef\]](#)
54. Vangi, E.; D'Amico, G.; Francini, S.; Borghi, C.; Giannetti, F.; Corona, P.; Marchetti, M.; Travaglini, D.; Pellis, G.; Vitullo, M.; et al. LARGE-SCALE High-Resolution Yearly Modeling of Forest Growing Stock Volume and above-Ground Carbon Pool. *Environ. Model. Softw.* **2023**, *159*, 105580. [\[CrossRef\]](#)
55. McRoberts, R.E.; Næsset, E.; Gobakken, T.; Bollandsås, O.M. Indirect and Direct Estimation of Forest Biomass Change Using Forest Inventory and Airborne Laser Scanning Data. *Remote Sens. Environ.* **2015**, *164*, 36–42. [\[CrossRef\]](#)
56. Chen, H.Y.H.; Luo, Y. Net Aboveground Biomass Declines of Four Major Forest Types with Forest Ageing and Climate Change in Western Canada's Boreal Forests. *Glob. Change Biol.* **2015**, *21*, 3675–3684. [\[CrossRef\]](#)
57. Puliti, S.; Breidenbach, J.; Schumacher, J.; Hauglin, M.; Klingenberg, T.F.; Astrup, R. Above-Ground Biomass Change Estimation Using National Forest Inventory Data with Sentinel-2 and Landsat. *Remote Sens. Environ.* **2021**, *265*, 112644. [\[CrossRef\]](#)
58. Dega, S.; Dietrich, P.; Schrön, M.; Paasche, H. Probabilistic Prediction by Means of the Propagation of Response Variable Uncertainty through a Monte Carlo Approach in Regression Random Forest: Application to Soil Moisture Regionalization. *Front. Environ. Sci.* **2023**, *11*, 1009191. [\[CrossRef\]](#)
59. Refsgaard, J.C.; van der Sluijs, J.P.; Brown, J.; van der Keur, P. A Framework for Dealing with Uncertainty Due to Model Structure Error. *Adv. Water Resour.* **2006**, *29*, 1586–1597. [\[CrossRef\]](#)
60. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [\[CrossRef\]](#)
61. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [\[CrossRef\]](#)
62. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
63. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
64. Saarela, S.; Grafström, A.; Ståhl, G.; Kangas, A.; Holopainen, M.; Tuominen, S.; Nordkvist, K.; Hyypä, J. Model-Assisted Estimation of Growing Stock Volume Using Different Combinations of LiDAR and Landsat Data as Auxiliary Information. *Remote Sens. Environ.* **2015**, *158*, 431–440. [\[CrossRef\]](#)
65. Esteban, J.; McRoberts, R.E.; Fernández-Landa, A.; Tomé, J.L.; Næsset, E. Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. *Remote Sens.* **2019**, *11*, 1944. [\[CrossRef\]](#)
66. Parks, S.A.; Holsinger, L.M.; Panunto, M.H.; Jolly, W.M.; Dobrowski, S.Z.; Dillon, G.K. High-Severity Fire: Evaluating Its Key Drivers and Mapping Its Probability across Western US Forests. *Environ. Res. Lett.* **2018**, *13*, 044037. [\[CrossRef\]](#)
67. Davidson, A.; Wang, S.; Wilmshurst, J. Remote Sensing of Grassland-Shrubland Vegetation Water Content in the Shortwave Domain. *Int. J. Appl. Earth Obs. Geoinf.* **2006**, *8*, 225–236. [\[CrossRef\]](#)
68. Perbet, P.; Guindon, L.; Correia, D.L.P.; Gahrouei, O.R.; Côté, J.F.; Béland, M. Historical Insect Disturbance Maps from 1985 Onwards for Canadian Forests Derived Using Earth Observation Data. *Sci. Data* **2025**, *12*, 1212. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Hall, R.J.; Castilla, G.; White, J.C.; Cooke, B.J.; Skakun, R.S. Remote Sensing of Forest Pest Damage: A Review and Lessons Learned from a Canadian Perspective. *Can. Entomol.* **2016**, *148*, S296–S356. [\[CrossRef\]](#)
70. Knott, J.A.; Liknes, G.C.; Giebk, C.L.; Oh, S.; Domke, G.M.; McRoberts, R.E.; Quirino, V.F.; Walters, B.F. Effects of Outliers on Remote Sensing-Assisted Forest Biomass Estimation: A Case Study from the United States National Forest Inventory. *Methods Ecol. Evol.* **2023**, *14*, 1587–1602. [\[CrossRef\]](#)
71. Mutanga, O.; Masenyama, A.; Sibanda, M. Spectral Saturation in the Remote Sensing of High-Density Vegetation Traits: A Systematic Review of Progress, Challenges, and Prospects. *ISPRS J. Photogramm. Remote Sens.* **2023**, *198*, 297–309. [\[CrossRef\]](#)
72. Naicker, R.; Mutanga, O.; Peerbhay, K.; Odebiri, O. Estimating High-Density Aboveground Biomass within a Complex Tropical Grassland Using Worldview-3 Imagery. *Environ. Monit. Assess.* **2024**, *196*, 370. [\[CrossRef\]](#)
73. Lewis, S.A.; Robichaud, P.R.; Hudak, A.T.; Strand, E.K.; Eitel, J.U.H.; Brown, R.E. Evaluating the Persistence of Post-Wildfire Ash: A Multi-Platform Spatiotemporal Analysis. *Fire* **2021**, *4*, 68. [\[CrossRef\]](#)
74. Morresi, D.; Vitali, A.; Urbinati, C.; Garbarino, M. Forest Spectral Recovery and Regeneration Dynamics in Stand-Replacing Wildfires of Central Apennines Derived from Landsat Time Series. *Remote Sens.* **2019**, *11*, 308. [\[CrossRef\]](#)
75. Ma, T.; Zhang, C.; Ji, L.; Zuo, Z.; Beckline, M.; Hu, Y.; Li, X.; Xiao, X. Development of Forest Aboveground Biomass Estimation, Its Problems and Future Solutions: A Review. *Ecol. Indic.* **2024**, *159*, 111653. [\[CrossRef\]](#)
76. Kumar, S.; Arya, S.; Jain, K. A SWIR-Based Vegetation Index for Change Detection in Land Cover Using Multi-Temporal Landsat Satellite Dataset. *Int. J. Inf. Technol.* **2022**, *14*, 2035–2048. [\[CrossRef\]](#)
77. White, J.C. Characterizing Forest Recovery Following Stand-Replacing Disturbances in Boreal Forests: Contributions of Optical Time Series and Airborne Laser Scanning Data. *Silva Fenn.* **2024**, *58*, 23076. [\[CrossRef\]](#)

78. Kurbanov, E.; Vorobev, O.; Lezhnin, S.; Sha, J.; Wang, J.; Li, X.; Cole, J.; Dergunov, D.; Wang, Y. Remote Sensing of Forest Burnt Area, Burn Severity, and Post-Fire Recovery: A Review. *Remote Sens.* **2022**, *14*, 4714. [\[CrossRef\]](#)
79. Chang, C.; Chang, Y.; Xiong, Z.; Liu, H.; Bu, R. Estimating the Aboveground Biomass of the Hulunbuir Grassland and Exploring Its Spatial and Temporal Variations over the Past Ten Years. *Ecol. Indic.* **2024**, *161*, 112010. [\[CrossRef\]](#)
80. Berner, L.T.; Massey, R.; Jantz, P.; Forbes, B.C.; Macias-Fauria, M.; Myers-Smith, I.; Kumpula, T.; Gauthier, G.; Andreu-Hayles, L.; Gaglioti, B.V.; et al. Summer Warming Explains Widespread but Not Uniform Greening in the Arctic Tundra Biome. *Nat. Commun.* **2020**, *11*, 4621. [\[CrossRef\]](#)
81. Berner, L.T.; Goetz, S.J. Satellite Observations Document Trends Consistent with a Boreal Forest Biome Shift. *Glob. Change Biol.* **2022**, *28*, 3275–3292. [\[CrossRef\]](#)
82. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 Surface Reflectance Data Set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [\[CrossRef\]](#)
83. Dubayah, R.; Blair, J.B.; Goetz, S.; Fatoyinbo, L.; Hansen, M.; Healey, S.; Hofton, M.; Hurtt, G.; Kellner, J.; Luthcke, S.; et al. The Global Ecosystem Dynamics Investigation: High-Resolution Laser Ranging of the Earth's Forests and Topography. *Sci. Remote Sens.* **2020**, *1*, 100002. [\[CrossRef\]](#)
84. Baccini, A.; Goetz, S.J.; Walker, W.S.; Laporte, N.T.; Sun, M.; Sulla-Menashe, D.; Hackler, J.; Beck, P.S.A.; Dubayah, R.; Friedl, M.A.; et al. Estimated Carbon Dioxide Emissions from Tropical Deforestation Improved by Carbon-Density Maps. *Nat. Clim. Change* **2012**, *2*, 182–185. [\[CrossRef\]](#)
85. Avitabile, V.; Herold, M.; Heuvelink, G.B.M.; Lewis, S.L.; Phillips, O.L.; Asner, G.P.; Armston, J.; Ashton, P.S.; Banin, L.; Bayol, N.; et al. An Integrated Pan-Tropical Biomass Map Using Multiple Reference Datasets. *Glob. Change Biol.* **2016**, *22*, 1406–1420. [\[CrossRef\]](#)
86. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [\[CrossRef\]](#)
87. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [\[CrossRef\]](#)
88. Hunt, E.R.; Rock, B.N. Detection of Changes in Leaf Water Content Using Near- and Middle-Infrared Reflectances. *Remote Sens. Environ.* **1989**, *30*, 43–54. [\[CrossRef\]](#)
89. Hunt, E.R., Jr.; Yilmaz, M.T. Remote Sensing of Vegetation Water Content Using Shortwave Infrared Reflectances. In *Proceedings of the Remote Sensing and Modeling of Ecosystems for Sustainability IV, San Diego, CA, United States, 6 October 2007*; SPIE: Bellingham, WA, USA, 2007; Volume 6679, pp. 15–22. [\[CrossRef\]](#)
90. Key, C.H.; Benson, N.C. Landscape Assessment: Remote Sensing of Severity, the Normalized Burn Ratio and Ground Measure of Severity, the Composite Burn Index. In *FIREMON: Fire Effects Monitoring and Inventory System*; USDA Forest Service, Rocky Mountain Research Station: Ogden, UT, USA, 2005.
91. McFeeters, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [\[CrossRef\]](#)
92. Merzlyak, M.N.; Gitelson, A.A.; Chivkunova, O.B.; Rakitin, V.Y. Non-Destructive Optical Detection of Pigment Changes during Leaf Senescence and Fruit Ripening. *Physiol. Plant.* **1999**, *106*, 135–141. [\[CrossRef\]](#)
93. Marsett, R.C.; Qi, J.; Heilman, P.; Biedenbender, S.H.; Watson, M.C.; Amer, S.; Weltz, M.; Goodrich, D.; Marsett, R. Remote Sensing for Grassland Management in the Arid Southwest. *Rangel. Ecol. Manag.* **2006**, *59*, 530–540. [\[CrossRef\]](#)
94. Badgley, G.; Field, C.B.; Berry, J.A. Canopy Near-Infrared Reflectance and Terrestrial Photosynthesis. *Sci. Adv.* **2017**, *3*, e1602244. [\[CrossRef\]](#)

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