Science of Remote Sensing 13 (2026) 100403

Contents lists available at ScienceDirect

Science of Remote Sensing

journal homepage: www.sciencedirect.com/journal/science-of-remote-sensing

Exploring the potential of Harmonized Landsat-Sentinel-2 in predicting boreal forest structure from UAV-LiDAR data in Northwestern America

Léa Enguehard [a,*], Stefan Kruse [a], Ronny Hänsch [b], Ulrike Herzschuh [a], Santosh Panda [c], Birgit Heim [a]

[a] Polar Terrestrial Environmental Systems, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[b] Microwaves and Radar Institute, German Aerospace Center (DLR), Wessling, Germany
[c] Institute of Agriculture, Natural Resources and Extension, University of Alaska Fairbanks, Fairbanks, AK, United States of America

ARTICLE INFO

Keywords:
UAV-LiDAR
HLS
Landsat
Sentinel-2
Forest structure
Boreal forest

ABSTRACT

Boreal forests play a critical role in global carbon dynamics and climate regulation, yet their structural attributes remain poorly characterized, particularly in structurally complex ecosystems such as the northern treeline. Here, we explored the potential of Harmonized Landsat and Sentinel-2 (HLS) multispectral data to predict UAV-LiDAR-derived forest structure across sites in the western North American boreal forest. We extracted spectral features from peak and late summer HLS and used Random Forest models to predict Canopy Height and Crown Cover at 30 m resolution. Our results show strong relationships between spectral and structural metrics, with HLS NDVI and Tasseled Cap Wetness emerging as key predictors. Predictive performance is higher for dense and sparse forests than medium-density forests, and no significant differences are found between peak and late summer models. We compared our UAV-LiDAR Crown Cover estimates to the ABoVE Tree Canopy Cover product and identified overestimation of Crown Cover in the treeline ecotones. These findings highlight the value of fine-scale UAV-LiDAR structural data for algorithm building and assessments of satellite-derived products. Nevertheless, the high proportion of green understory reduces the sensitivity of HLS spectral features to canopy height and crown cover compared with applications in more productive forests. An open-access HLS–forest structure dataset is provided, containing HLS pixel-wise labeled forest structure information. By combining structural reference data with spectral HLS satellite imagery, this study contributes to filling the structural data gap at the boreal forest northern edge and to predicting forest structure in high-latitude ecosystems.

1. Introduction

Boreal forests represent one of the largest terrestrial biomes on Earth, storing over 30% of all carbon present in the terrestrial biome, and play a critical role in global carbon dynamics, climate regulation, and biodiversity (Bonan et al., 1992; Brandt et al., 2013; Gauthier et al., 2015; Kasischke, 2000). Structural attributes of these forests, such as canopy height and crown cover, are essential for quantifying forest aboveground biomass, productivity, and resilience to disturbance (Bolton et al., 2015; Lefsky et al., 2002; Senf et al., 2020; Shugart et al., 2010). These metrics can also inform ecological models and carbon monitoring systems.

Over the past decade, several remote sensing-derived land cover and forest-focused datasets have been made publicly available, including global land cover maps (ESA, 2017; Buchhorn et al., 2020a,b) and forest cover products (Hansen et al., 2013). Focusing on the boreal region, the NASA Arctic-Boreal Vulnerability Experiment (National Aeronautics and Space Administration; ABoVE) Terrestrial Ecology Program conducted airborne overflights in Alaska and western Canada (Miller et al., 2019) and produced satellite-derived land cover for the Arctic and boreal regions (e.g., Wang et al., 2019, 2020). Using these airborne data for calibration and validation, Feng et al. (2022) developed the ABoVE Tree Canopy Cover (TCC) and Stand Age circum-hemispherical boreal product from Landsat with a 30 m spatial resolution, optimized from the University of Maryland's Global Land Cover Facility global TCC product (Sexton et al., 2013). Despite these advances, obtaining accurate, consistent, and scalable measurements of forest structure across boreal forests remains a significant challenge, particularly in the discontinuous tundra-taiga northern transition zone as described by Matasci et al. (2018) and Montesano et al. (2016).

* Corresponding author.
E-mail address: lea.enguehard@awi.de (L. Enguehard).

https://doi.org/10.1016/j.srs.2026.100403
Received 13 October 2025; Received in revised form 12 February 2026; Accepted 26 February 2026
Available online 3 March 2026
2666-0172/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

In the North American boreal forests, ground-based and airborne forest structural data are still sparse near the treeline and taiga-tundra ecotones (Matasci et al., 2018). While spaceborne laser altimeter missions such as ICESat-2 and GEDI (Global Ecosystem Dynamics Investigation) (Dubayah et al., 2020; Markus et al., 2017) have advanced near-global spaceborne forest structure monitoring (e.g., the Canopy height dataset ATL08 from Neuenschwander et al., 2021), their coverage or resolution is often inadequate in high-latitude areas due to orbital constraints, sparse sampling, or low return densities. GEDI is onboard the International Space Station, covering latitudes between 51.7° North and South, hence excluding boreal forests. ICESat-2 was explicitly designed for inland ice and sea ice, and although it can provide useful canopy height estimates, the errors remain significant for areas of sparse canopy cover, such as at the northern forest edge (Feng et al., 2023; Neuenschwander et al., 2020). As a result, the structural composition of the northern edge of boreal forests remains poorly characterized at the scale needed for regional assessments and modeling.

Airborne Light Detection and Ranging (LiDAR), and increasingly in recent decades, Unpiloted Aerial Vehicle-LiDAR (UAV-LiDAR) systems have emerged as valuable tools for high-resolution boreal forest inventories and structure mapping (Guimarães et al., 2020; White et al., 2016), offering precise 3-D measurements of an area of up to a few km^3. Forest structural attributes, such as canopy height, can be derived from LiDAR point clouds and used to characterize forests (Alonzo et al., 2018; Belmonte et al., 2020; Liu et al., 2018) or estimate aboveground biomass (Lu et al., 2020; Maesano et al., 2022), for example. Compared to traditional plot-based forest inventories, UAV and airborne LiDAR datasets provide extensive, high-resolution spatial coverage, making them well-suited for the development and validation of predictive algorithms. As a result, they are particularly effective for upscaling attributes to wall-to-wall satellite-derived data for large-scale forest monitoring (Bolton et al., 2015; Matasci et al., 2018; Villoslada et al., 2024). Forest characterization based on remote sensing is not spatially limited to areas captured in forest inventory programs or the time window of fieldwork (Coops et al., 2021; Wulder et al., 2020). Previous studies have demonstrated the successful upscaling of canopy height or biomass to spaceborne LiDAR using airborne laser scanning in North American boreal systems (Margolis et al., 2015; Potapov et al., 2021). While UAVs cover smaller spatial extents than airborne systems, they offer significantly higher spatial resolution to derive forest structure. However, the potential of UAV-LiDAR data for upscaling boreal forest structure has yet to be systematically explored, specifically at the northern edge of boreal forests.

Landsat-1 was the first spaceborne multispectral sensor launched in the 1970s by the NASA/United States Geological Survey for land-surface monitoring. Landsat satellites have since played a critical role in monitoring Earth's surface at regional to global scales (Wulder et al., 2022). Landsat data, characterized by a 30 m pixel resolution for the spectral bands since Landsat-4 (launched 1982), have been intensively used to study North American boreal forests from assessing fire severity, forest recovery trends, to the stability of boreal forest stands within the context of climate change (French et al., 2008; Masek, 2001; Pickell et al., 2016). The Sentinel-2 satellite mission is the European Space Agency's (ESA) equivalent to Landsat, offering a higher spatial resolution (10 to 20 m) but with a more recent start date in 2015. Sentinel-2 has been used for global to local-scale monitoring and mapping of the Earth's surface (Phiri et al., 2020) and in boreal forest studies (Delcourt et al., 2021; Enguehard et al., 2024; Grabska et al., 2019; Majasalmi and Rautiainen, 2016). Both Landsat and Sentinel-2 capture spectral reflectance in the visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) wavelength regions.

As vegetation reflectance varies with the optical properties, phenological stage, and spatial arrangement of canopy elements (Baret et al., 2000; Goel, 1988; Goel and Thompson, 2000; Verhoef, 1984), exploration of the mapping of biophysical forest variables using reflectance from Landsat and other multispectral sensors has been the focus of many studies for decades (Cohen and Spies, 1992; Hall et al., 2006; Powell et al., 2010; Hudak et al., 2002; Hansen et al., 2002). Bolton et al. (2020) successfully trained Random Forest models using Landsat time series metrics to upscale airborne LiDAR-derived forest structure, including tree height, across study sites in the southern Canadian boreal domain. Because the Sentinel-2 satellite missions span a shorter time period than Landsat, relatively few studies have so far used Sentinel-2 data to upscale forest structure. Astola et al. (2019) investigated the potential of Sentinel-2 and Landsat-8 using multivariable models to predict forest structure calibrated with in-situ plot data in boreal forest in Finland. Li et al. (2020) compared the performance of deep learning and Random Forest models using Sentinel-1 (SAR), Sentinel-2, and Landsat-8 data to upscale canopy height for deciduous broad-leaf and deciduous needle-leaf forest types in China. Fakhri et al. (2025) developed an optimization approach to estimate forest fractional cover, canopy height, and height difference by iteratively optimizing vegetation indices derived from standard vegetation indices and the four 10-m Sentinel-2 bands, calibrated for semi-arid broadleaf forest types in Iran and transferred to a mixed temperate forest in Southern Germany. Using Sentinel-2, Lang et al. (2019) developed a deep convolutional neural network to upscale airborne LiDAR and Structure from Motion-derived forest structure to country-wide tree height estimates for tropical forests in Gabon and predominantly temperate mixed forests in Switzerland. These studies have employed increasingly complex methodologies, including machine learning and deep learning approaches. Because they focused on areas with available in situ forest plots or airborne-derived forest structure data, they primarily addressed productive forests, often dominated by commercially managed species. To the best of our knowledge, forests assessed in Sentinel-2 and Landsat-based studies rarely include unmanaged, low-productivity forests, such as the low-stature, mostly unmanaged forest types at the northern boreal forest edge examined in this study using our UAV–LiDAR–surveyed sites.

Therefore, we aim to use our UAV field data to explore the spectral relationships between a low-productivity but highly sensitive forest ecosystem and Landsat and Sentinel-2 spectral information. A key component in producing land cover products suitable for monitoring is the availability of freely accessible satellite data at a high processing level (Fassnacht et al., 2024; Bolton et al., 2020; Hermosilla et al., 2016). High-quality, processed Landsat and Sentinel-2 surface reflectance time series are operationally provided by the respective agencies, and the NASA Harmonized Landsat–Sentinel (HLS) product dataset (Ju et al., 2025; Masek et al., 2021) further combines both Landsat-8,9 and Sentinel-2A, 2B surface reflectance data into a geometrically co-registered and spectrally-consistent 30-m product. This fusion effectively increases temporal multispectral product availability compared to using either sensor alone, with observations of the land surface every 2 to 3 days. Common issues for optical remote sensing in the boreal forest regions are large data gaps caused by persistent cloud cover, forest fire smoke in the summers, and low illumination, sun angle, or snow cover from autumn to spring. Therefore, due to its higher temporal resolution, the HLS dataset provides a robust and rich optical dataset for studying boreal forests and offers a promising avenue for upscaling forest structural information derived from higher spatial resolution datasets such as UAV-LiDAR. While HLS data have been recently used for land-cover monitoring (Chen et al., 2021; Mulverhill et al., 2023), its potential in predicting forest structure from UAV-LiDAR, particularly in the northern boreal edge, remains unexplored. Moreover, seasonal variation of spectral surface reflectance from forest land cover, such as peak vs. late summer time series (Van Geffen et al., 2025), may influence how well spectral signals in HLS correlate with forest structural metrics. Understanding which HLS spectral band or indices correlate the most with forest structure is critical for identifying optimal seasonal windows for predicting forest structure and for developing models needed at a scale that can inform science, management, and policy decisions (Radeloff

et al., 2024).

In this study, we evaluate the potential of peak summer and late summer HLS multispectral data to predict key UAV-LiDAR-derived forest structural metrics and specifically, Canopy Height and Crown Cover, across 89 UAV-LiDAR-surveyed sites distributed throughout the northern edge of the western North American boreal forest. Our approach involves: (1) exploring the linkage of HLS spectral bands and indices for peak and late summer to characterize boreal forest structure, (2) exploring regression models to link HLS peak and late summer spectral and UAV-LiDAR structural data, and (3) exploring how forest structure is related to the existing ABoVE TCC product (Feng et al., 2022) using our UAV-LiDAR dataset, focusing on the northern edge of boreal forests. We aim to address the following questions:

What HLS spectral properties in peak and late summer are most predictive of UAV-LiDAR-derived forest structure? How does predictive performance vary across forest density gradients, from open to closed-canopy stands? How does seasonality (peak vs. late summer) influence model accuracy and spectral–structural relationships?

To support this investigation, we introduce a new, openly available, training dataset comprising 30 × 30 m HLS pixels containing peal and late summer spectral features, each labeled with the corresponding pixel-wise UAV-LiDAR-derived Canopy Height and Crown Cover measurement for 89 sites in the boreal forest region of Alaska and Northwest Canada (NW-Canada), also specifically covering sites in the treeline ecotone.

2. Data and methods

2.1. Study region

The study area spans the western North American boreal forest between the latitudes 58°N–68.5°N and longitudes 121°W–165°W. This region is characterized by low temperature and precipitation, low diversity of dominant tree species, generally low canopy height (Appendix B3), and a sparse human population (Chapin et al., 2006). During the 2022 to 2024 summer field campaigns, we conducted UAV-LiDAR surveys at 89 sites across the boreal forests of Alaska and northwestern Canada (Fig. 1).

Our study sites cover the northern forest biome edge (treeline ecotone), with different forest structures (sparse/open and abrupt/uniform transitions) in Canada and Alaska (Mackenzie Delta, CA, mountain slopes of the Brooks Range, AK) (Montesano et al., 2020). In 2024, we extended our field campaigns to the western treeline on the Seward Peninsula. Our sites also cover dense coniferous forests (generally Spruce: *Picea mariana, Picea glauca*) and mixed deciduous-coniferous stands (e.g., *Betula neoalaskana, Populus tremuloides, Populus balsamifera* and Spruce) in Interior Alaska and Canada, as well as mountainous landscapes in Alaska and NW Canada. The study sites are located in multiple ecoregions (Omernik and Griffith, 2014; Ruaro et al., 2024), namely Brooks Range Tundra, Alaska Tundra, Alaska Boreal Interior, Boreal and Taiga Cordillera, Taiga Plain, and a few sites on the Marine West Coast forest (Fig. 1). This wide gradient captures the boreal forests' full range of structural and compositional variability characteristics.

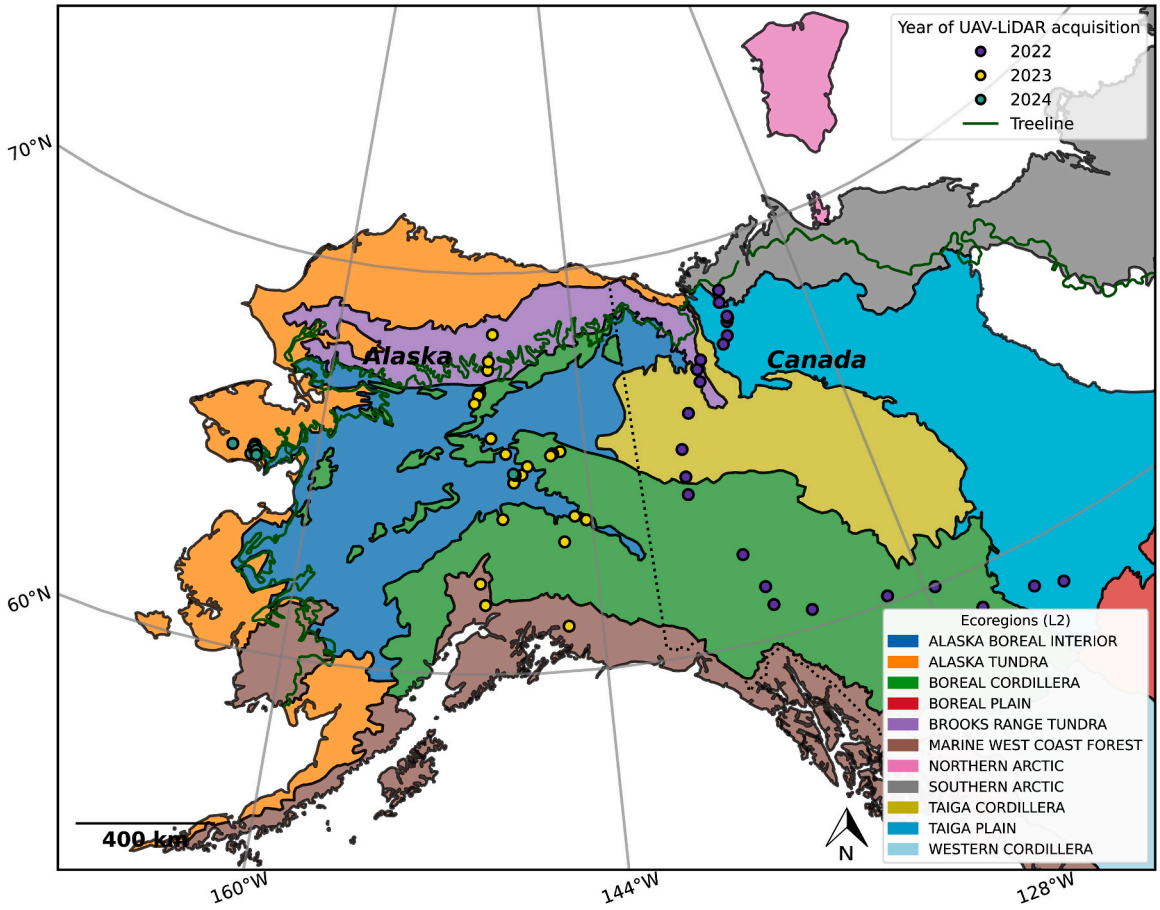

Fig. 1. Location of field sites and unpiloted aerial vehicle (UAV) transects and year of data collection. We established 23 sites during the Canada 2022 expedition (purple), 47 sites during the Alaska expedition 2023 (yellow), and 19 during the Alaska 2024 expedition (green). The green line represents the Circum-Arctic Vegetation Map (CAVM) treeline (Raynolds et al., 2019), the approximate border between the tundra and the boreal forest. The ecoregions represented are the United States Environmental Protection Agency (EPA) Level II ecoregions (Omernik and Griffith, 2014; Ruaro et al., 2024). Some sites overlap because they are closely located. See the associated published dataset for exact site location (Enguehard et al., 2025).

2.2. UAV-LiDAR data collection and processing

2.2.1. Collection of UAV-LiDAR data

Each field site encompassed a minimum 500 × 50 m UAV-LiDAR surveyed area, within which we established a circular 30 m-diameter forest inventory plot. The LiDAR data were acquired with a YellowScan Mapper LiDAR sensor mounted on a DJI M300 drone. The Yellowscan mapper has a Livox Horizon scanner, featuring a long detection distance (260 m), high precision and accuracy (2 and 3 cm), and high field-of-view (81.7° (horizontal) × 25.1° (vertical)), with a laser wavelength of 905 nm. The laser has a point rate of 240,000 to 480,000 points/s, and beam divergence of 0.28° (horizontal) × 25.1° (vertical). Survey paths were flown at a speed of 5 m s^{-1} along parallel lines spaced 20 m apart, at an altitude of 70 m above ground. With this configuration, the point cloud data has a ground resolution of a minimum of 400 points m^{-2}.

2.2.2. LiDAR-derived canopy height and canopy cover

We processed the UAV-LiDAR data from each site using a series of steps to produce high-quality, scalable point clouds. The raw LiDAR data processing steps involved correcting the UAV-flight trajectories in POSPAC PPTX and flight strip alignment using the software YellowScan CloudStation.

Following the PC2RCHIVE method (Farkas and Kruse, 2024), we derived the digital terrain model, Canopy Height Model (CHM), and segmented individual trees at each site (Brieger et al., 2019; Kruse et al., 2025a, 2025b; Kruse et al., 2025). The digital terrain model, representing the ground level without vegetation, was computed using a local minimum moving filter and smoothed to mitigate the effects of micro-topography. The CHM was computed as the difference between the digital top-of-canopy surface model and the digital terrain model, and smoothed (3 × 3 pixel mean) to prevent over-segmentation of tree crowns. Both CHM and digital terrain models have a resolution of 2-10 cm. Individual trees were detected by calculating the position of tree tops and the extent of crowns using a dynamic circular moving window using the R package ForestTools (Plowright, 2023).

The crown cover (or coverage) represents the proportion of the ground surface that is covered by tree crowns when viewed from above. Here, we generated a binary mask from the CHM to compute crown cover for all trees taller than 1 m. By excluding the low-stature vegetation below this threshold, such as low shrubs, seedlings, or herbaceous cover, we minimize noise and misclassification, ensuring that only established tree crowns contribute to the crown cover. Finally, we stacked the CHM and crown cover layers into single raster files, each maintained at the original centimeter-scale resolution.

2.3. HLS data collection and pre-processing

2.3.1. HLS data collection

The Harmonized Landsat Sentinel-2 (HLS) project provides cross-standardized surface reflectance data from the Operational Land Imager aboard the Landsat 8 satellite and the Multi-Spectral Instrument aboard Europe's Copernicus Sentinel-2A and 2B satellites Ju et al. (2025); Masek et al. (2021). In the HLS processing, Landsat and Sentinel-2 are atmospherically corrected using the Land Surface Reflectance Code, developed by NASA (Vermote et al., 2016) and adapted by the United States Geological Survey. Both sensor data are normalized to bi-directional reflectance view angle effects and adjusted for sensor band differences with the Operational Land Imager as the reference and provided as a 30-m product in a common Universal Transverse Mercator -based tile grid from 2015 onwards (Claverie et al., 2018).

We acquired HLS surface reflectance scenes from the cloud computing platform Google Earth Engine (Gorelick et al., 2017) using the geemap Python API package (Wu, 2020) at each of the 89 UAV-LiDAR surveyed areas. The scenes were selected for two phenological seasons: peak summer and late summer of the field campaign years 2022, 2023, and 2024. We defined peak summer as the full leaf-on period from June 20th to August 10th and late summer as the onset of deciduous defoliation from August 20th to September 30th. These windows were carefully chosen to prevent early-season snowfall in the latter period and to ensure the two intervals do not overlap. We selected scenes with less than 30% cloud cover and shadow over the transect area. For our study, we selected the following six main spectral HLS bands: blue, green, red, NIR, SWIR1, and SWIR2 which are all spectral bands acquired by both Landsat and Sentinel-2 sensors.

2.3.2. HLS data processing

After the HLS data collection, we first applied a filter with the Function of mask in Google Earth Engine (Fmask v4.9) optimized for Landsat and Sentinel-2 to eliminate the remaining clouds, water, and snow (Qiu et al., 2019). Next, we added a normalized difference vegetation index (NDVI) band (NIR-Red)/(NIR + Red) following Rouse et al. (1974). To catch any remaining water bodies or road surfaces not flagged by Fmask, we masked pixels with any negative surface reflectance, NIR surface reflectance below 0.02, or NDVI below 0.3 – thereby excluding non-vegetated or water-dominated areas. Finally, we stacked all the masked images and generated a median surface reflectance composite at 30 m HLS pixel spatial scale for each UAV transect for both peak summer and late summer. Raster files were produced for each field site covering the entire UAV transects, with all seven bands' median surface reflectance pixel values and NDVI for peak summer and late summer.

2.4. Data fusion processing

To combine the HLS with the UAV-LiDAR data, we resampled and stacked the LiDAR raster-derivatives with the HLS pixel size (30 by 30 m). We derived canopy height variables for each HLS pixel, namely, the median, minimum, maximum, and standard deviation of the UAV-LiDAR-derived canopy height, which were each produced as single raster layers. We computed the crown cover percentage per HLS pixel from the binary mask representing a single raster layer.

We deleted pixel values where the crown cover percentage was below 5% to remove non-forested areas or very sparse forested landscapes. The final merged dataset, herein referred to as the HLS forest structure dataset (Enguehard et al., 2025), includes one Cloud Optimized GeoTIFF per field site containing 19 stacked bands: seven peak summer bands, seven late summer bands, four canopy height derived bands, and crown cover percentage (Table 1).

2.5. Key HLS spectral information of UAV transects

We applied Principal Component (PC) and Tasseled Cap (TC) transformations (Kauth and Thomas, 1976) to extract the key HLS-derived spectral information characterizing the UAV transects across the boreal region. The PC rotation with fixed axes, i.e., the TC transformation, is a widely used method in forest remote sensing (Dymond et al., 2002; Frazier et al., 2014; Healey et al., 2005), to condense the key biophysical information from multispectral satellite imagery. It transforms the original spectral bands into a new set of orthogonal indices that represent meaningful physical characteristics of the landscape, commonly Brightness, Greenness, and Wetness in the case of a vegetated land surface.

We analyzed the indices to evaluate seasonal differences and land surface dynamics with reduced data dimensionality but enhanced thematic relevance. Here, we used the scikit-learn package (Pedregosa et al., 2011) to compute the PCA for the set of HLS features (seven features for each season) to transform the seven peak summer and seven late summer HLS spectral features into uncorrelated components. We applied the TC transformation using the updated coefficients provided by Zhai et al. (2022) for Landsat 8. We used the same coefficients for both Landsat 8 and Sentinel-2 because the atmospheric and bidirectional

Table 1
Overview of the feature variables included in the HLS forest structure dataset computed per pixel.

Band N°	Feature	Computed as per HLS pixel	Band HLSL30 Landsat	Band HLSS30 Sentinel-2	Unit
Band 1	Blue	Median surface reflectance peak summer	B02	B02	Unitless (0 -1)
Band 2	Green	Median surface reflectance peak summer	B03	B03	Unitless (0 -1)
Band 3	Red	Median surface reflectance peak summer	B04	B04	Unitless (0 -1)
Band 4	NIR	Median surface reflectance peak summer	B05	B08	Unitless (0 -1)
Band 5	SWIR1	Median surface reflectance peak summer	B06	B11	Unitless (0 -1)
Band 6	SWIR2	Median surface reflectance peak summer	B07	B12	Unitless (0 -1)
Band 7	NDVI	(NIR- Red)/ (NIR + Red) peak summer	B04, B05	B04, B08	Unitless ratio (−1 to 1)
Bands 8 to 13	Same band order	Median surface reflectance late summer	B02, B03, B04, B05, B06, B07	B02, B03, B04, B08, B11, B12	Unitless (0 -1)
Band 14	NDVI	(NIR- Red)/ (NIR + Red) late summer	-	-	Unitless ratio (−1 to 1)
Band 15	Median Canopy Height	Median LiDAR-derived canopy height per HSL pixel value	-	-	Meter
Band 16	Standard Deviation Canopy Height	Standard Deviation LiDAR-derived CHM per HSL pixel value	-	-	Meter
Band 17	Minimum Canopy Height	Minimum LiDAR-derived CHM per HLS pixel value	-	-	Meter
Band 18	Maximum Canopy Height	Maximum LiDAR-derived CHM per HLS pixel value	-	-	Meter
Band 19	Crown cover percentage	LiDAR-derived Canopy Cover of trees above 1m	-	-	Percentage

reflectance distribution function correction was similarly applied for all HLS products. The TC transformation was applied to peak and late summer features to derive the three principal indices (Brightness, Greenness, and Wetness) for each temporal snapshot. The coefficients from the referenced table were linearly combined with the respective surface reflectance bands and NDVI to compute the transformed outputs.

2.6. Forest structure upscaling

2.6.1. Upscaling of canopy height and crown cover

We assessed how well median Canopy Height and crown-cover percentage per HLS pixel can be estimated with distinct regression models trained for two seasons: peak summer and late summer. For each season, we fitted a Random Forest regressor (Breiman, 2001) to all 22, 724 pixels using the seven HLS reflectance bands (blue, green, red, NIR, SWIR1, SWIR2, and NDVI) and the three Tasseled Cap variables (Brightness, Greenness, and Wetness).

Hyperparameters (i.e., parameters that can be tuned within the model) were optimized by heuristics to achieve the best overall accuracy without overfitting the models (Appendix A1). Model performance was assessed with 20-fold cross-validation, yielding out-of-sample predictions for every observation so that the reported R^2, root mean square error (RMSE), and mean absolute error (MAE) metrics represent truly independent estimates of accuracy. Each fold had an independent training and validation set, making sure that a distance of at least 100 m between sampled pixels was achieved (Appendix A2). Finally, we inspected the feature importance of both models to evaluate the relative influence of each HLS band on canopy-height and crown-cover predictions.

2.6.2. Upscaling of forest structure types

To evaluate the potential for upscaling specific forest types (e.g., sparse or dense), we classified discrete, balanced forest structure categories identified from the Crown Cover and Canopy Height distributions with Random Forest supervised classification. The categories were derived from the 30 m aggregated UAV-LiDAR-derived Crown Cover distribution: 5-20% (sparse forest), 20-50% (medium density forest), 50-80% below 5 m (dense, shorter forest), 50-80% above 5 m (dense and tall forest), 80-100% below 5 m (very dense and short forest, and 80-100% above 5 m (very dense and tall forest). The classifiers were trained with a shuffled 20-fold cross-validation on the peak and late summer stacks using the same seven HLS features (blue, green, red, NIR, SWIR1, SWIR2, and NDVI), and three Tasseled Cap variables (Brightness, Greenness, and Wetness) as used for the regression in section 2.6.1. Using the sklearn library in Python (Pedregosa et al., 2011). Each fold had an independent training and validation set, ensuring that a distance of at least 100 m between sampled pixels was reached. Hyperparameters were optimized by heuristics to achieve the best overall accuracy (Appendix A1). We evaluated the classifiers with four standard evaluation metrics: precision, recall, F1-score, and overall accuracy, and inspected the feature importance. We plotted One-vs-Rest rate--of-change (ROC) curves for each class and reported the corresponding area under the curve (AUC) scores to assess class separability across all thresholds.

2.7. Assessment of NASA ABoVE upscaled forest structure product

To assess currently available Crown Cover datasets with our data, we compared the HLS forest structure Crown Cover data with the ABoVE TCC dataset (Feng et al., 2022). The TCC dataset contains Landsat-derived and locally calibrated TCC, covering most of our study sites (13 sites not covered) for the year 2020. Here, we matched the pixels from both Crown Cover datasets and compiled the tree canopy cover percentage pixel values. We made a pixel-wise comparison and computed the bias between the TCC and HLS forest structure pixels to identify regions of low accuracy. We assessed the robustness of the TCC dataset by computing Pearson correlation coefficients and RMSE.

3. Results

3.1. HLS forest structure dataset properties

3.1.1. Structural and spectral properties

The HLS forest structure dataset shows that the 30 m median Canopy Height across all pixels varied by up to 25 m, with a right-skewed frequency distribution with a long tail extending towards the higher trees and higher frequency around 5 m (Appendix B1). The 30 m Crown Cover distribution varied between 5 and 100% and is bimodal, with one frequency peak around 15% and one around 90%. The lowest Crown Cover frequency was between 60 and 70% cover. Based on the Crown Cover and Canopy Height distribution and our field knowledge, we identified six forest type categories.

The HLS forest structure dataset was compiled for 648 tiles for peak summer and 258 tiles for late summer, which were then composited as the median per pixel. The spectral range of each band, considering the median surface reflectance per pixel, varied for both seasons (Appendix B2 & Fig. 2). Peak summer surface reflectance showed higher NIR and green and lower red surface reflectance in the VIS than late summer. In contrast, late summer showed higher red surface reflectance in the VIS and lower NIR than peak summer. Specifically, the dense and high forest structure type (80-100% $>$ 5 m) presented the highest NIR surface reflectance in peak summer and late summer (Fig. 2, green). The lowest cover category (5-20%, red) had the second highest NIR surface reflectance in late summer, but is close to the other categories. The medium cover category (50-80% $>$ 5 m) presented the lowest NIR, SWIR1, and SWIR2 surface reflectance for both peak and late summer.

3.1.2. Spatial distribution of forest structure classes

Across the UAV transects within the study region, the most common forest types fell within the 5–20% (orange) and 20–50% (purple) canopy cover categories (Fig. 3), indicating a widespread dominance of sparse to moderately dense forests. Denser forest types were more prevalent in Alaska Boreal Interior and Boreal Cordillera, as well as in the southern portion of the Canadian study area. In western Alaska (Seward Peninsula), relatively dense but short forests were observed in our UAV transects, located at the border of the Alaska Tundra and Boreal Interior ecoregions (Fig. 1). In contrast, many transects along the northern forest edge of our study region, especially those within the treeline ecotone, were dominated by sparse forest cover, with a limited presence of high canopy density. This distribution reflects a clear geographical gradient in forest density across the region and reflects forest structural heterogeneity. For example, we observed lower heterogeneity closer to the treeline, where the 5–20% category predominates. In contrast, forests in Seward Peninsula and southern sites showed greater heterogeneity, with a dominating type but alongside a substantial representation of other categories.

3.1.3. Essential HLS spectral information of UAV transects

We applied PC and TC transformations to identify the key spectral information contained in the UAV-Lidar transects. The PCA of both peak summer and late summer pixels showed that the first two PCs explain more than 90% of the variance together (Appendix C1). Most of the variability (68.3% variance peak summer, 72.5% late summer) was captured by the first principal component in both seasons, suggesting a strong underlying spectral gradient for all HLS features in the visible wavelength range and for SWIR-2. NIR and NDVI are strongly aligned with the second principal component (26.4% variance peak summer, 19% late summer) in both seasons.

Forest structure categories showed distinct clustering but with substantial overlap; the 80-100% $>$5m category (green) was the most tightly clustered and distinct in both seasons. The spread of points was larger in late summer than in peak summer, with more compact ellipses and better separability. In peak summer, we observed one cluster of the 80-100% $<$5m category at the top of principal component 2 (Fig. 4, yellow), suggesting high NIR and NDVI values. HLS VIS and SWIR spectral features dominated the first principal component and separated the majority of the forest structure categories, whereas the vegetation productivity-related HLS NDVI and NIR spectral features dominated the second principal component (Appendix C1); specifically, it separated the lower from the higher trees in the densest forest category in late summer.

Tasseled Cap analysis of Brightness (TCB), Greenness (TCG), and Wetness (TCW) revealed clear patterns among forest categories, particularly in peak summer, where the data were more tightly clustered compared to late summer (Fig. 5). A strong positive correlation between TCB and TCG was evident across all forest structure categories and both seasons. In late summer, the overall range of TCB and TCG values was reduced, suggesting a seasonal decline in vegetation vigor. Forest structure categories with higher canopy cover (especially 80–100%) consistently exhibited higher TCG and TCB values, while low-cover classes (5–20%) tended to occupy the lower end of these gradients.

Across both seasons, TCW values were negative and generally lower in late summer than in peak summer. A slight negative relationship between TCW and TCG was observed, more pronounced in peak summer. Again, high-canopy cover forests displayed higher TCW and TCG, while low-canopy cover classes consistently showed lower values,

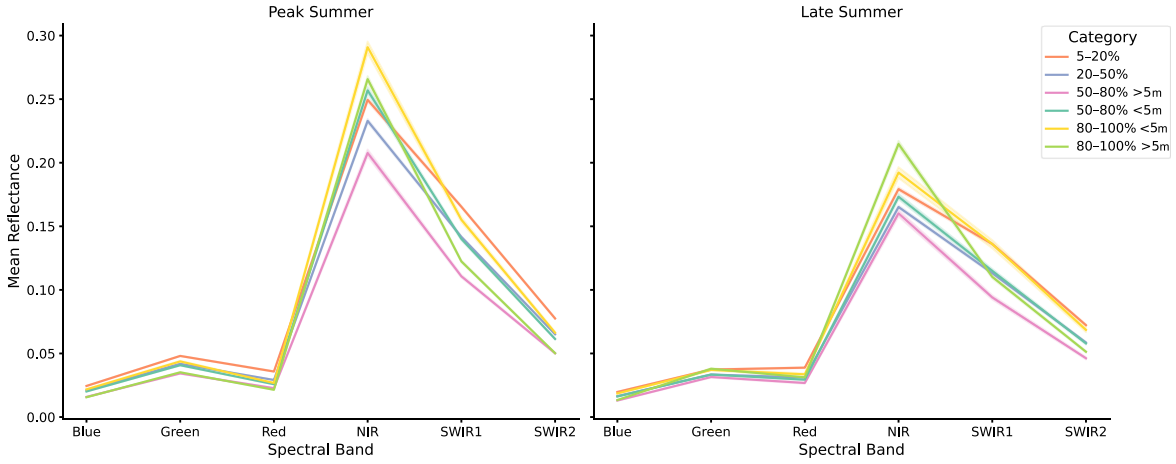

Fig. 2. HLS surface reflectance per forest category for peak summer (left) and late summer (right) interpolated across the visual (VIS), near infrared (NIR), and shortwave infrared (SWIR) spectral features. The surface reflectance is computed as the mean of the median surface reflectance per HLS Landsat and Sentinel-2 of all pixels in the category. In the categories, the percentage is the percentage of tree cover on that pixel, and '</>" are height sub-categories: above or below 5 m.

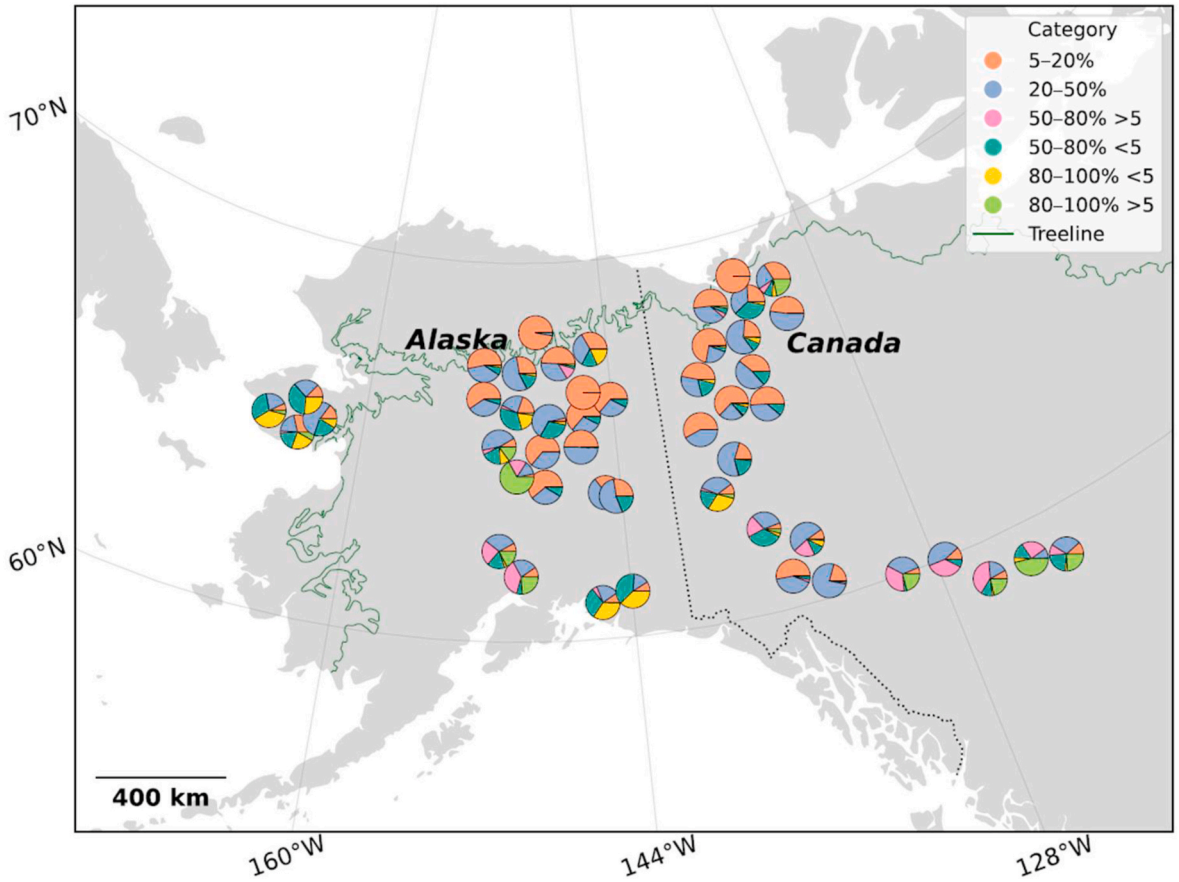

Fig. 3. Distribution of the forest structure categories across sites. Each pie chart corresponds to the proportions of forest structural categories within an unpiloted aerial vehicle (UAV) transect. The green line represents the Circum-Arctic Vegetation Map (CAVM) treeline (Raynolds et al., 2019), the border between the tundra and the boreal forest. The exact location of the sites may have been slightly moved to reduce overlap with other pie charts.

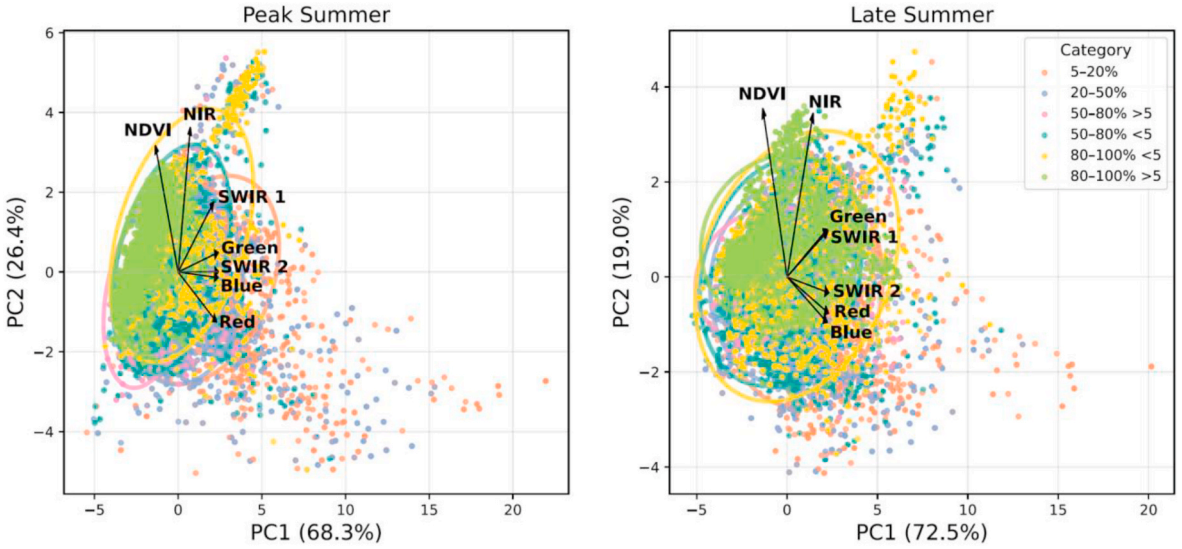

Fig. 4. Principal component 1 (PC1) versus 2 (PC2) for peak summer (left) and late summer (right). Each color code is a different forest structure category, with corresponding ellipses.

reflecting differences in canopy structure and moisture retention.

3.2. Quantitative upscaling: Crown Cover and Canopy Height

3.2.1. Regression models

Using the ten HLS spectral features, VIS, NIR, SWIR, spectral bands, NDVI, TC (Brightness, Greenness, Wetness), from our HLS forest

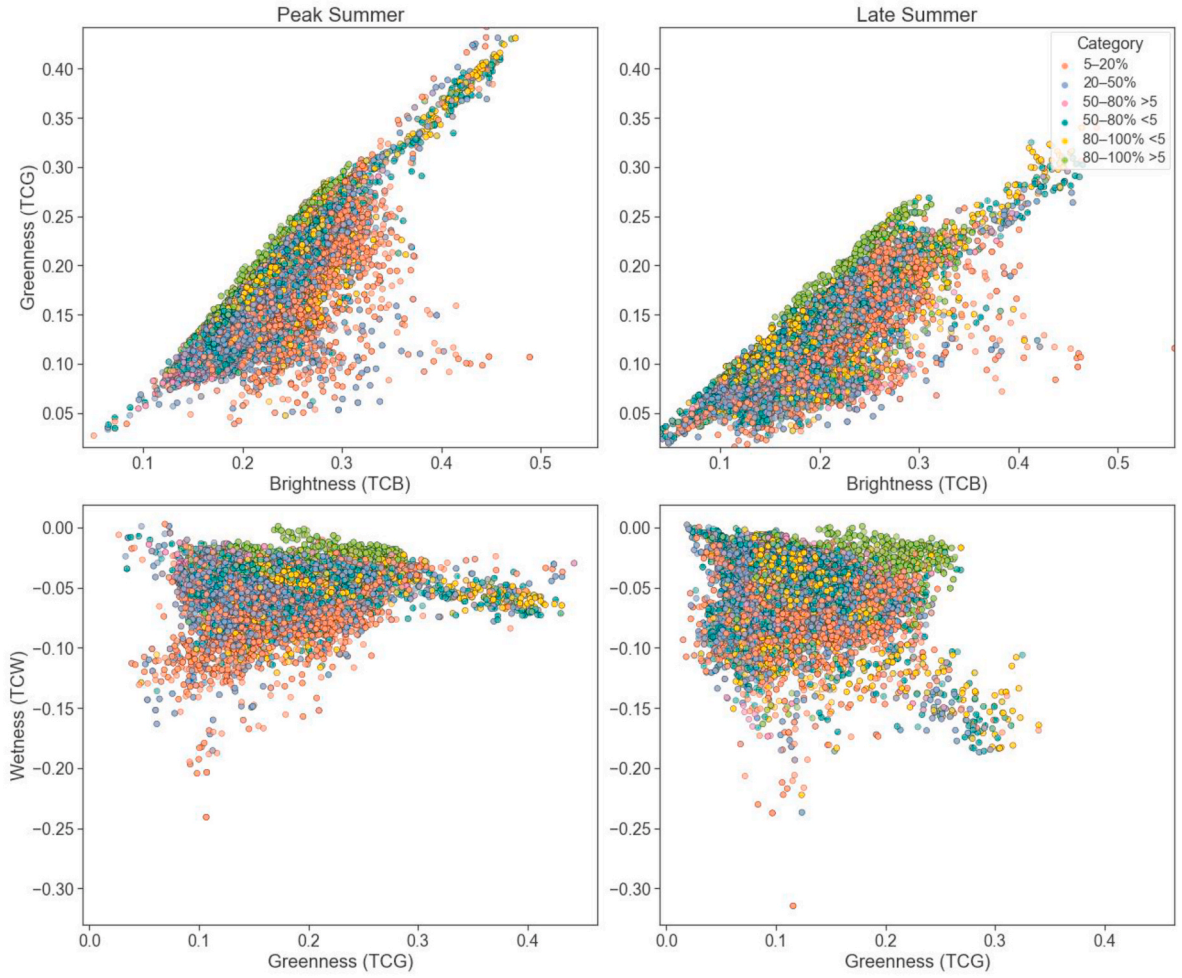

Fig. 5. Tasseled Cap (TC) coefficients derived from the HLS spectral features for peak summer (left column) and late summer (right column) pixels. The top is TC Brightness (TCB) vs. TC Greenness (TCG): (a) peak summer, (b) late summer, and the bottom row is TCG vs. TC Wetness (TCW): (c) peak summer, (d) late summer. Each color code is a different forest structure category. TCB and TCW both separate well the majority of forest structure categories, whereas TCG, which is most strongly linked to vital vegetation, does not.

structure dataset, we predicted Canopy Height and Crown Cover for peak and late summer, in total, four optimized random forest-based prediction models. Overall, the predictions worked well, with R^2 ranging from 0.60 to 0.66, and RMSE of 18.05% (peak summer) and 18.76% (late summer) for the Crown Cover predictions and 2.12 m to 2.07 m, respectively, for the Canopy Height (Fig. 6). Peak summer and late summer regressions showed similar results, with a slightly better-performing algorithm for Canopy Height in late summer and for Crown Cover in peak summer.

For the Canopy Height predictions, we observed a relatively good fit between predicted and 30 m UAV-LiDAR-derived values (Fig. 6a and b). However, large Canopy Height values (>10 m) tended to be underestimated while smaller Canopy Height values (1-5 m) tended to be overestimated (Fig. 6a and b, Fig. 7a and c).

For the Crown Cover predictions, we also observed a relatively good fit between predicted and measured values, but slightly worse than for the Canopy Height (Fig. 6b and d). The Crown Cover prediction slightly underestimated Crown Cover above 80%, and overestimated values between 20% and 80%, with the maximum overestimation around 40% crown cover for peak and late summer (Fig. 7b and d).

3.2.2. Key HLS spectral features

The important feature analysis suggested that across all four models and seasons, all HLS-derived spectral features played a significant role in predicting either Canopy Height or Crown Cover. Specifically, the TC Wetness and NIR features were important for predicting Canopy Height in both peak and late summer (Fig. 8 a&b). For Crown Cover, the spectral feature importance was even more distinct, with the photosynthetic activity of vegetation index NDVI, Red, and TCW showing relatively more importance than other features for peak summer (Fig. 8 c&d). TC Wetness shows a good separation between the forest structure categories, specifically for peak summer (Fig. 5 a&b). Regarding the Tasseled Cap variables, TC Wetness was the most important feature for all models, while TC Brightness was the least important predictor.

3.3. Qualitative upscaling: forest-type sub-categories classification

3.3.1. Classification results

We classified forest structure categories to assess the potential of upscaling for peak and late summer. Using the same ten spectral HLS features (VIS, NIR, SWIR, NDVI, three TC variables), we classified the six forest structure categories based on crown cover: 5-20%, 20-50%, 50-80% below 5 m, 50-80% above 5 m, 80-100% below 5 m, and 80-100% above 5 m.

We did not find significant differences between peak summer and late summer accuracies (Table 2). Specifically, overall accuracies of 0.60 and 0.61 were reached for the late summer and peak summer classifications, respectively. The dense and tall tree category, 80-100% > 5m, had the highest results in both peak and late summer, with precision, recall, and F-1 scores ranging from 0.69 to 0.78. The lowest density

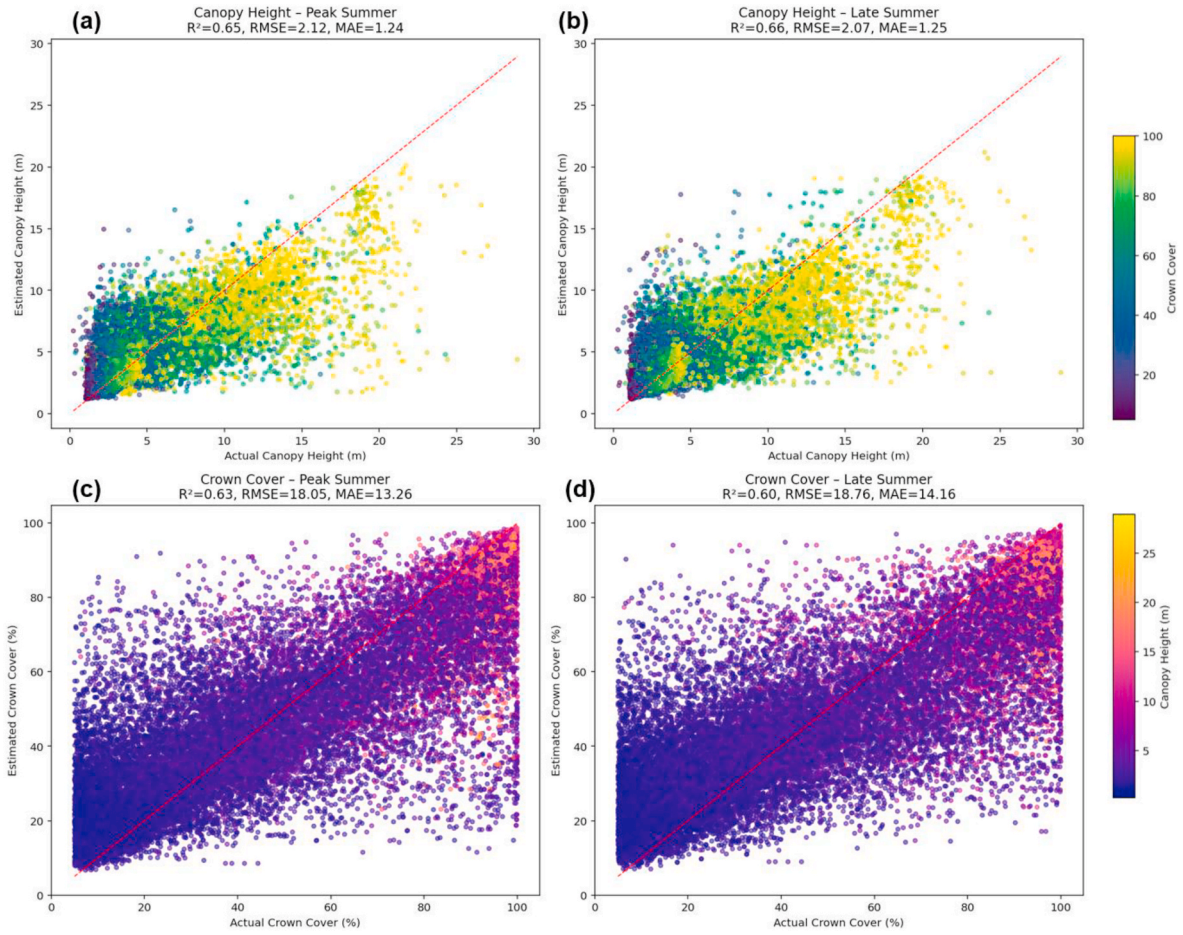

Fig. 6. Predicted forest structure, Canopy Height (a,b) and Crown Cover (c, d) using the random forest-based models: (top) Canopy Height: (a) peak summer and (b) late summer, and (bottom) Crown Cover: (c) peak summer and (d) late summer. The x-axes are the UAV-LiDAR-derived observed variable aggregated to 30 m (Enguehard et al., 2025), and the y-axes the predicted values for the HLS spectral features. For Canopy Height regressions, the data points are colored with the Crown Cover percentage, and conversely for the Crown Cover points. Associated R^2, root mean square error (RMSE), and mean absolute error (MAE) values are indicated on top of each sub-plot. The red dashed line represents the 1:1 perfect match.

crown cover category (5-20%) also had high accuracies, with accuracies varying between 0.67 and 0.69 for both seasons. Conversely, the forest category with the lowest accuracy was 50-80% $<$ 5 m, representing higher crown cover but lower trees.

3.3.2. *Confusion matrix and one-vs-rest ROC curve*

The One-vs-Rest ROC curves and confusion matrices support these findings (Fig. 9 b&d). The ROC curves showed that the 80–100% $>$ 5m class achieved the highest separability from other classes, with an area under the curve (AUC) of 0.95 in both seasons. The forest structure class with the lowest Crown Cover (5–20%) also had relatively high AUC values (0.87–0.88), consistent with their strong classification metrics. In contrast, the 50–80% $<$ 5m and 20-50% forest structure classes overlapped more with other categories in the ROC space, indicating lower discriminative power.

The confusion matrices further revealed that most misclassifications occur between structurally similar classes (Fig. 9a and c). For instance, samples from the 50–80% $<$ 5m forest structure class were frequently confused with the 50–80% $>$ 5m forest structure class, suggesting difficulty in distinguishing tree height when Crown Cover is similar. Similarly, moderate confusion was observed between the 5–20% and 20–50% forest structure classes. Overall, the matrices emphasized that classification accuracy was highest for well-separated structural classes (e.g., tall, dense canopies) and lower for intermediate or structurally ambiguous forest structure categories.

The feature importance analysis results showed that all spectral HLS-derived features play a significant role in predicting the classes, but that TC Wetness has the most importance among all predictors. The NDVI band was more important for late summer than peak summer, while the Red bandwas more important for peak summer (Appendix D1).

3.4. *Relationship between important features and predicted variables*

We explored the relationship between the predicted variables Canopy Height and Crown Cover, with the HLS predicting features of Blue, Green, Red, NIR, SWIR1, SWIR2, NDVI, TCB, TCG, and TCW. We observed a positive relationship between NDVI and Canopy Height and Crown Cover (Fig. 10), meaning that higher forest canopy and Crown Cover tended to have higher NDVI values. Specifically, we see that peak summer spectrally separated the forest categories better than late summer, where the data are sparser. Similarly, TCW showed a positive relationship with both predicted variables (Fig. 11), as well as the NIR feature in both peak and late summer (Appendix E3) and for the green feature in late summer (Appendix E5). Conversely, negative trends were observed for SWIR1, SWIR2, Red, TCB, and Blue for both peak and late summer, and Green in peak summer. (Appendix E1-8). SWIR, VIS, and NDVI HLS features showed a stronger relationship with Canopy Height than NIR.

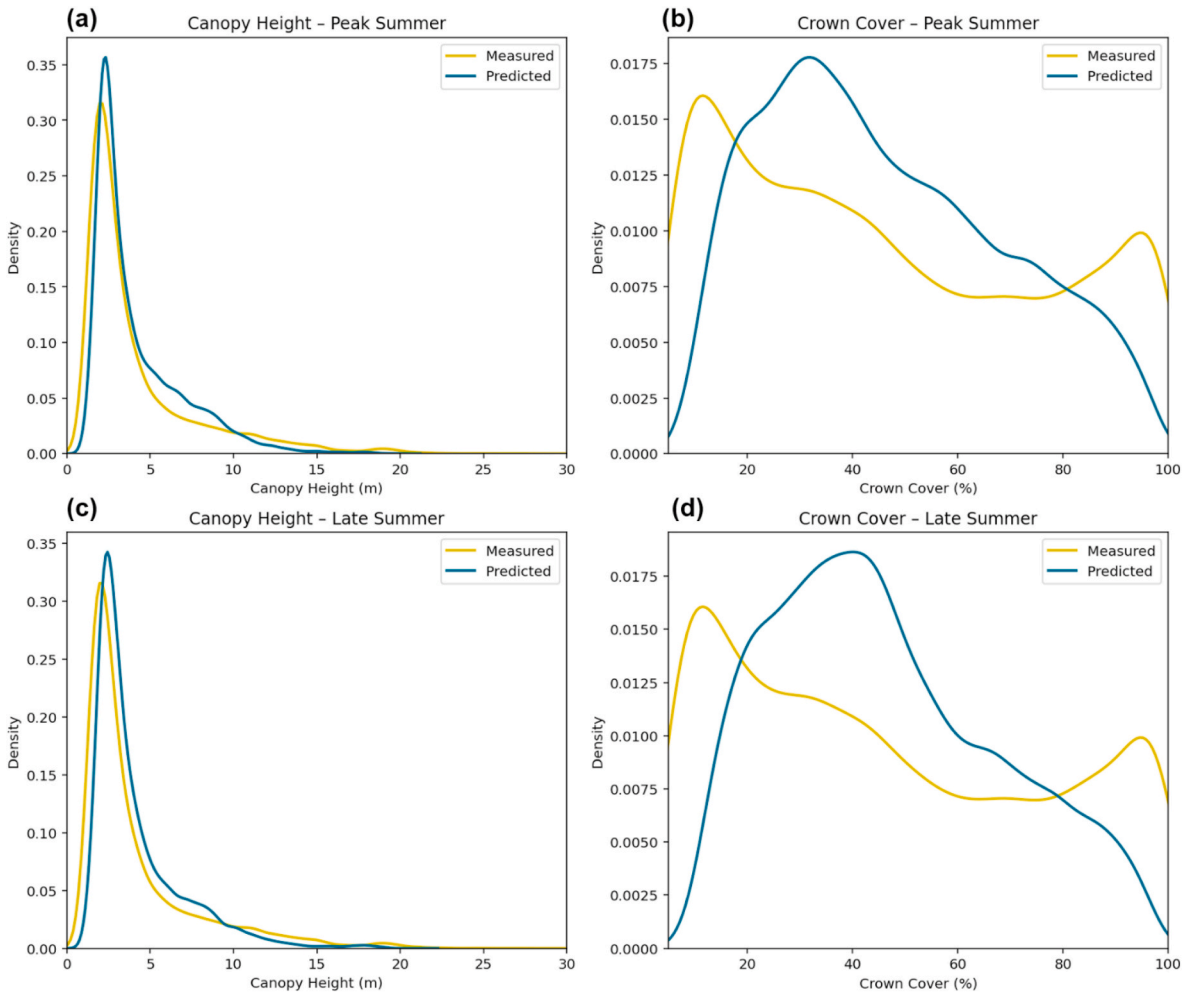

Fig. 7. Density distribution of UAV-LiDAR-derived 30 m aggregated variables (yellow) and predicted values of HLS spectral features (blue) for each random forest-based prediction model: (1) Canopy Height: (a) peak summer and (c) late summer, and (2) Crown Cover: (b) peak summer and (d) late summer.

3.5. Assessment of upscaled forest structure

3.5.1. Comparison with NASA ABoVE tree canopy cover

We compared the Crown Cover percentage from our HLS forest structure dataset (Enguehard et al., 2025) with the ABoVE TCC (Feng et al., 2022). We found a moderate positive relationship between the two datasets, with a Pearson coefficient of 0.49 and an RMSE of 26.51 (Appendix F1). Specifically, the ABoVE TCC Crown Cover tended to be higher than the HLS forest structure data between our 30 m 35-80% UAV-LiDAR observed Crown Cover, and lower outside of this range, specifically showing underestimation in the dense crown cover category higher than 80% (Fig. 12 & Appendix F1). The high RMSE indicated significant variation in the datasets.

3.5.2. Geographical comparisons of forest structure predictions

The median bias per site across our study region (Fig. 13) showed that the ABoVE TCC overestimates Crown Cover in the treeline ecotone in the more mountainous regions of Interior Alaska and in the mountainous northwestern forests extending from Alaska south through southern Yukon, interior British Columbia, and the Alberta foothills in NW Canada. ABoVE TCC Crown Cover seemed to be underestimated in the Taiga Cordillera in NW Canada, the Seward Peninsula, and Marine West Coast forest in Western Alaska. Interior Alaska showed both under- and over-estimated sites. Fig. 14 shows the median bias per site between our HLS-predicted and the 30 m UAV-LiDAR observed Crown Cover. Compared to the bias comparison of ABoVE TCC, we observed that there is much less overestimation at the northern treeline.

4. Discussion

Leveraging UAV-LiDAR data from 89 surveyed sites across the western North American Boreal Forest at its Northern Edge, we investigated the potential for upscaling key forest structural attributes, Canopy Height and Crown Cover, using Harmonized Landsat and Sentinel-2 multispectral imagery from peak and late summer seasonal temporal windows. We characterized the relationship of UAV-LiDAR transect-derived forest structure with HLS spectral features spanning all HLS VIS, NIR, and SWIR bands, the first three Principal Components, and Tasseled Cap indices, and assessed the relationship of the existing ABoVE Tree Canopy Cover (TCC) product to the UAV-LiDAR-derived crown cover at the sites of our transects. Our results demonstrate reasonable relationships between forest structural information derived from UAV-LiDAR and spectral information derived from 30-m resolution multispectral satellite data, highlighting the considerable potential for upscaling Canopy Height and Crown Cover to Landsat and Sentinel-2 satellite scales.

4.1. HLS forest structure key spectral characteristics

Boreal forests typically feature less dense and shorter overstory trees and therefore, a higher abundance of the lower layer of shrub-dominated understory. Specifically, in open-canopy areas, this

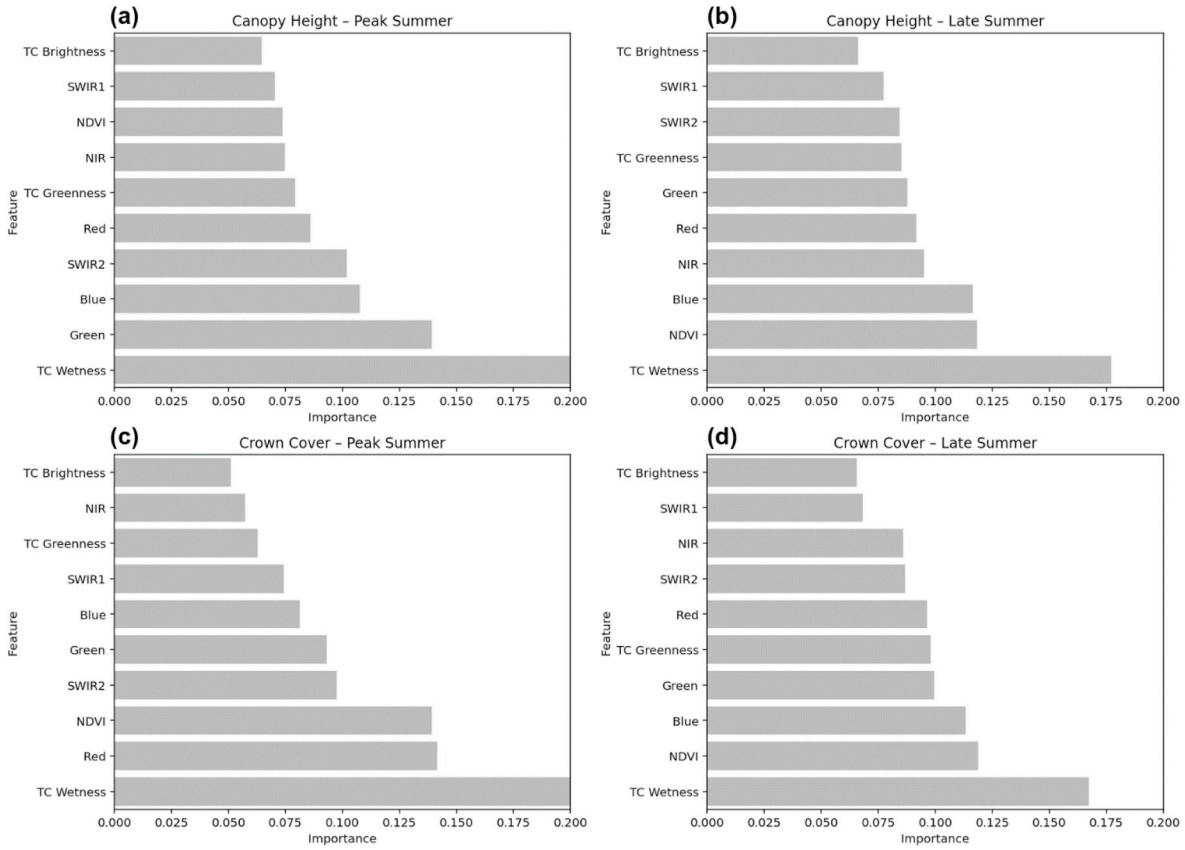

Fig. 8. Importance of each of the seven spectral features in the four models: (1) Canopy Height: (a) peak summer and (c) late summer, and (2) Crown Cover: (b) peak summer and (d) late summer.

Table 2
Classification performance of the random forest-based prediction models for the six forest structure categories for peak and late summer, showing precision, recall, F1-score, and overall accuracy.

Label	Late Summer classification			Peak Summer classification		
	Precision	Recall	F1-score	Precision	Recall	F1-score
20–50%	0.68	0.73	0.71	0.69	0.74	0.71
50–80% < 5	0.58	0.53	0.55	0.58	0.54	0.56
50–80% > 5	0.62	0.57	0.59	0.62	0.58	0.60
5–20%	0.78	0.75	0.77	0.79	0.75	0.77
80–100% < 5	0.65	0.60	0.63	0.65	0.60	0.62
80–100% > 5	0.81	0.85	0.83	0.81	0.84	0.83
accuracy			0.706			0.709
macro avg	0.69	0.67	0.68	0.69	0.68	0.68
weighted avg	0.70	0.71	0.70	0.71	0.71	0.71

understory can significantly contribute to multispectral reflectance (Pisek et al., 2021). Since we defined a horizontal threshold of 5 % minimum Crown Cover above 1 m height in our HLS dataset, we are reducing spectral contributions from surface elements such as leaf litter, moss, lichen, rock, and soil, as described as forest background contributors in (Pisek and Chen, 2009; Roberts et al., 2004). However, a high shrub understory is likely to contribute meaningfully to the spectral signal in our dataset.

The characteristics derived from HLS spectral features, including the spectral VIS, NIR, SWIR HLS bands and their PC and TC transformations, capture clear structural and seasonal patterns across the UAV-surveyed sites. Forests with high Crown Cover (especially 80–100%) consistently show higher TC Greenness, Brightness, and Wetness values in line with LaRue et al. (2018), reflecting greater biomass and photosynthetic activity from the dominating evergreen needle-leaf trees. NDVI also increases with both Canopy Height and Crown Cover, consistent with higher NIR surface reflectance and higher red absorption, and by this, low red surface reflectance due to enhanced pigment absorption in denser, greener canopies. However, the lowest density class (5-20%) shows high NIR surface reflectance in peak summer and even in late summer, also with a wide NDVI range that is more indicative of quite high vegetation biomass and vitality. In contrast, even the medium density class (50-80 %) with high trees (>5 m) is characterized by lower NIR surface reflectance.

Peak summer shows improved separability of forest structure classes and stronger vegetation signals than late summer. A slight negative relationship between TC Wetness and Greenness, particularly in peak summer, further reflects the interaction between vegetation cover and moisture retention. Additionally, the observed decrease in SWIR surface reflectance with increasing Canopy Height and Crown Cover aligns with their sensitivity to water absorption and lignin-cellulose content, traits more prominent in sparsely vegetated or drier areas (Stenberg et al., 2008). Together, these trends confirm that our merged UAV-LiDAR and HLS dataset captures ecologically meaningful gradients in forest structure and condition of the northern edge of the boreal forest, with HLS spectral metrics strongly linked to both canopy characteristics and seasonal dynamics.

4.2. *HLS-based canopy height and crown cover predictions*

Our results demonstrate that both forest structure variables–Canopy Height and Crown Cover– derived from high spatial-resolution UAV-LiDAR, can be predicted using 30-m pixel resolution HLS satellite-derived multispectral data. Here, we obtained model accuracies

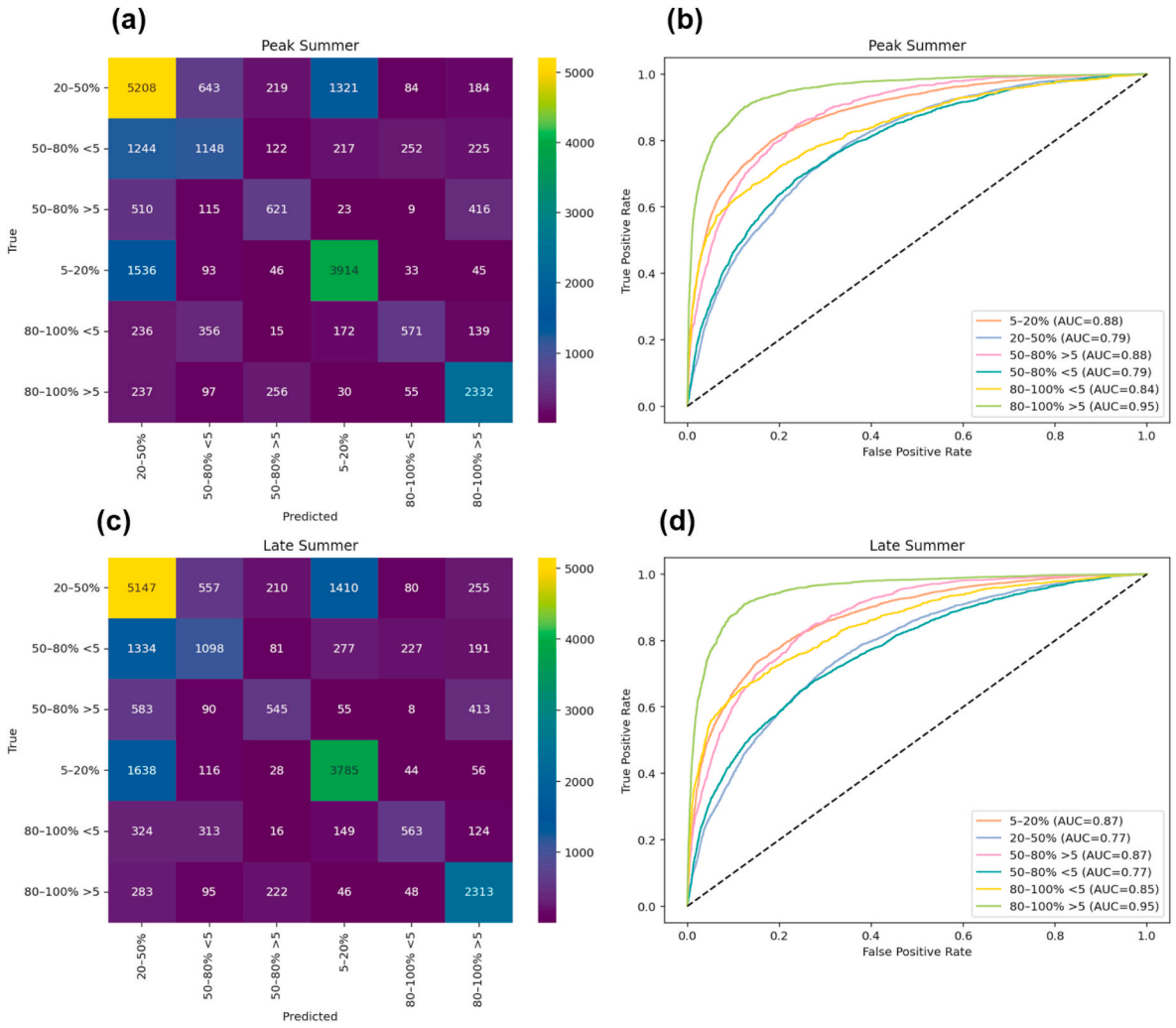

Fig. 9. Confusion matrix and One-vs-Rest rate of change (ROC) curves for peak and late summer: (top) peak summer (a) confusion matrix, (b) ROC diagram; and (bottom) late summer (c) confusion matrix and (d) ROC diagram.

ranging from 0.60 to 0.66 (R^2), showing similar to higher performance than previous studies focusing on forest-related variables. More general land cover classifications always yield higher accuracies; for example, the ESA Climate Change Initiative Land Cover (ESA, 2017) and the Copernicus Global Land Cover 100 m (Buchhorn et al., 2020a, 2020b) reach overall accuracies of up to 90%. However, they separate spectrally distinct classes (such as surface water, barren, grassland, forest, etc.). For forest-related variables, Potapov et al. (2021) predicted global forest height with GEDI and Landsat data and achieved an R^2 of 0.61 when validated against airborne LiDAR, but did not provide assessments in the boreal zone. Similarly, Matasci et al. (2018) upscaled lidar plots across Canada for key structural variables such as stand height, stem volume, and canopy cover. The authors validated their prediction on more than 20,000 lidar plots and reached R^2 values ranging from 0.49 to 0.61. Despite their relatively lower accuracies, those studies were recognized as robust large-scale applications.

While our work and dataset highlight the potential of UAV-LiDAR combined with HLS data for large-scale forest structure mapping at the northern boreal forest edge, we also demonstrate the forest type-driven inherent challenges to future users of our dataset. As we showed in the exploration of the HLS spectral features, the most abundant forest class, characterized by low tree heights (<5 m) and low crown cover, contains mixed spectral pixels with a high contribution of vegetation greenness from the green understory, which spans the full value range of HLS spectral features. On the other hand, the low-productivity northern boreal forests exhibit a very narrow height distribution, with a high frequency of trees below 5 m. This imposes a data-driven limitation on quantifying canopy height. In contrast, mean canopy heights in forests studied by Fakhri et al. (2025), Lang et al. (2019), and Li et al. (2020) ranged from 10 to 15 m up to 30-40 m and 50-70 m in Gabon mangrove and tropical forests. Even in the boreal forest study in Finland (Astola et al., 2019), the forest was still of the productive type with 13-15 m canopy height. Thus, our forest type–driven limitations stand in contrast to other forest structure retrievals in more productive forest types and will similarly challenge the quantification of structural attributes using advanced methods, such as linear optimization developed by Fakhri et al. (2025) or deep learning approaches in Lang et al. (2019) and Li et al. (2020).

In their study, Li et al. (2020) described that both random forests and deep learning models performed well, but the retrieval performance of canopy height was more strongly linked to the satellite sensors, with Sentinel-1 SAR showing the highest performance, followed by Sentinel-2, due to its higher spatial resolution and red-edge bands compared with Landsat-8. Similarly, in performance assessments by Astola et al. (2019) for Sentinel-2 and Landsat-8, Sentinel-2 outperformed Landsat-8 for forest structure retrievals due to the red-edge bands and finer pixel resolution. In our study, we did not undertake Sentinel-2 versus Landsat-8 comparisons, as HLS–Sentinel-2 has a slightly coarser spatial resolution of 30 m per pixel, linked to the native Landsat resolution, rather than the native Sentinel-2 10 m and 20 m

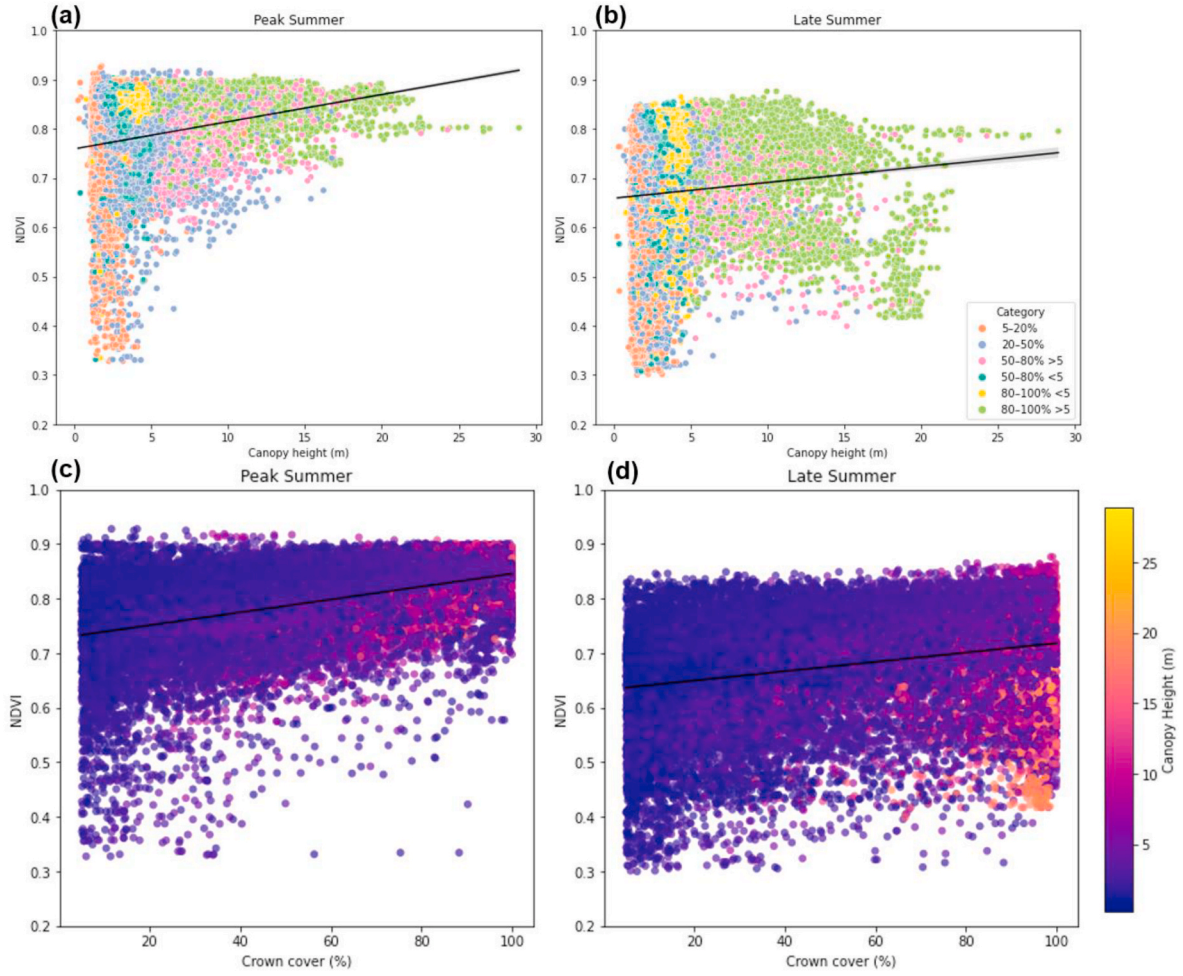

Fig. 10. Relationship between Canopy Height (top) and Crown Cover (bottom) with normalized difference vegetation index (NDVI) for peak summer (a-c) and late summer (b-c). For enhanced understandability, Canopy Height points are colored with the forest structure categories, and the Crown Cover points with the Canopy Height values. The Pearson correlation coefficients (r) were computed: (a) r = 0.251; (b) r = 0.113; (c) r = 0.444; (d) r = 0.251.

pixel lengths. HLS data, however, will allow users to process long time series of forest structure historically, once suitable algorithms are developed.

The next steps to derive forest structure for low-productivity forest types at the northern forest edge from our HLS dataset would be to develop models and implement larger-scale HLS data processing, similar to the large-scale Landsat processing described in Masek et al. (2021), Hermosilla et al. (2016, 2017, 2018, 2022), and Berner and Goetz (2022), at least covering our study regions. Fakhri et al. (2019, 2022, 2025) also demonstrated that model performance improves when locally or regionally optimized spectral indices are developed, specifically for semi-arid forest types. However, as discussed in Montesano et al. (2016), high uncertainties remain at the pixel level, and low tree cover is a defining attribute of the northern forest biome. Montesano et al. (2016) show in their assessment of global 30 m Landsat canopy cover (Sexton et al., 2013) (the basis for the northern boreal forest optimized ABoVE TCC (Feng et al., 2022)) that the performance compared with airborne LiDAR reference data had an uncertainty of around 25% crown cover, with a bias toward lower crown coverage. After calibrating ABoVE TCC with boreal reference data, uncertainty was reduced to 16.4% crown cover, a level comparable to the uncertainty obtained in our study.

The slightly higher predictive accuracy for Canopy Height in our study may stem from its strong relationship with continuous vegetation gradients – such as biomass, leaf area, and vertical structure – which are effectively characterized by NIR and visible spectral HLS bands. These spectral features are particularly sensitive to vegetation density and the organization of branches, needles, and leaves, resulting in different shadow contributions and brightness. Furthermore, internal leaf structure and water content strongly influence the TCW, NIR, and SWIR signals, all of which are correlated with canopy height. Specifically, the Tasseled Cap Wetness was a significant predictor of forest structure. At our sites, canopy height appears to be less correlated with brightness and more strongly associated with greenness and wetness. In general, other studies also highlight the complexities of correlating spectral response with canopy height (e.g., Shimizu et al., 2020), specifically at low canopy heights (Fakhri et al., 2025).

In contrast, Crown Cover, a more discrete and binary structural measure, primarily reflects the presence or absence of vegetation in a "horizontal" sense (see definitions in Li et al., 2023). This makes the main spectral signal more closely tied to NDVI and the red wavelength region, which are sensitive to surface greenness and chlorophyll content. NDVI effectively captures horizontal vegetation extent but seems relatively insensitive to vertical variation in canopy structure.

Although peak summer shows greater separability of forest classes in the PCA, we find no significant differences in model performance between the peak and late summer datasets. This likely reflects limited phenological variation in the western North American Boreal Forest during these periods, especially in evergreen-dominated systems where spectral signals remain relatively stable across the summer (Berra and Gaulton, 2021). These findings suggest that peak summer data alone are sufficient for accurate structural predictions, though incorporating both

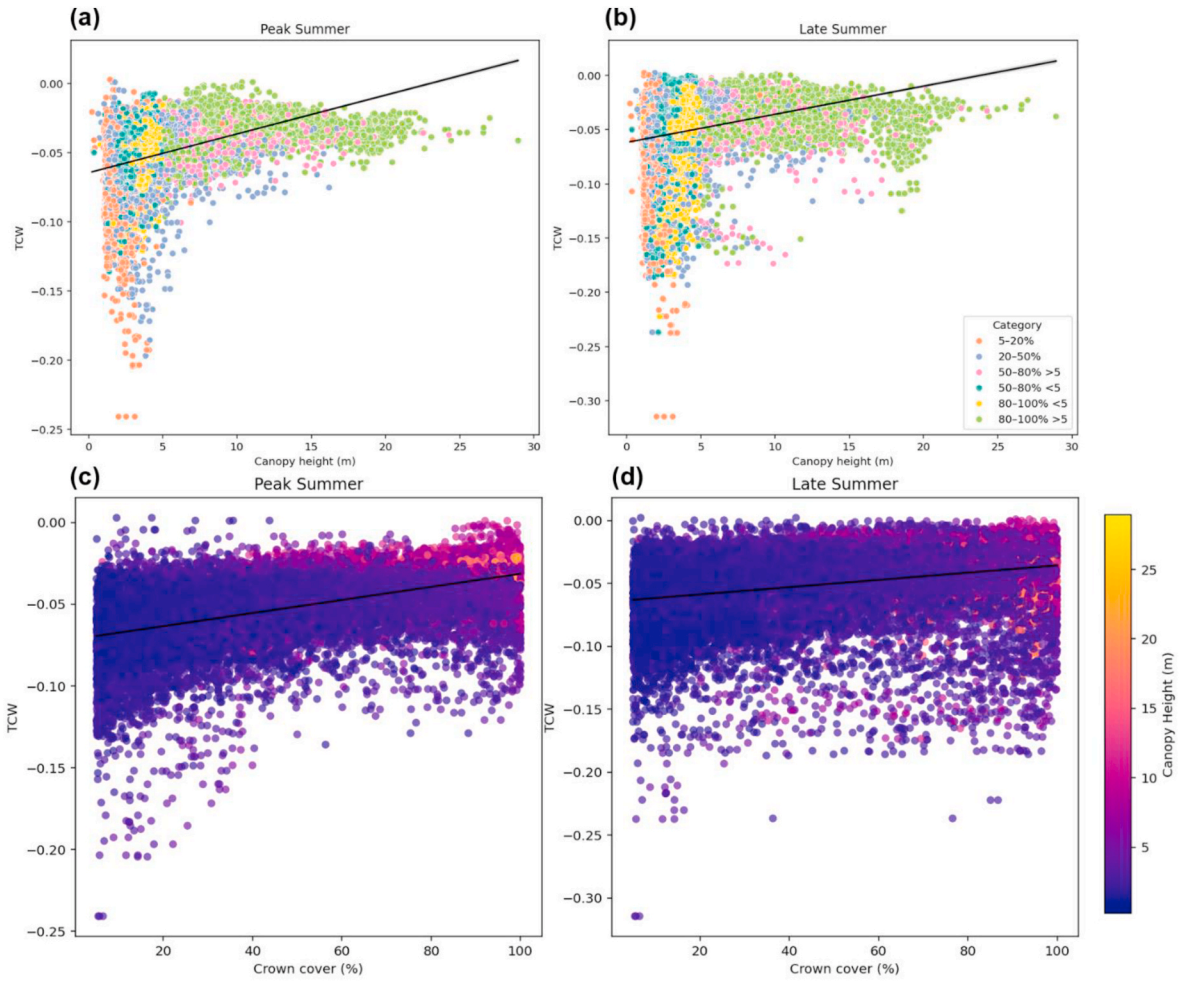

Fig. 11. Relationship between Canopy Height (top) and Crown Cover (bottom) with Tasseled Cap Wetness (TCW) for peak summer (a-c) and late summer (b-c). For enhanced understandability, Canopy Height points are colored with the forest structure categories, and the Crown Cover points with the Canopy Height values. The Pearson correlation coefficients (r) were computed: (a) r = 0.462; (b) r = 0.313; (c) r = 0.545; (d) r = 0.288.

phenological time windows could potentially improve generalizability and robustness.

4.3. HLS-based forest structural category classification

In their spectral and structural analysis of forests, Roberts et al. (2004) noted that while green foliage typically dominates the reflectance signal from forest pixels, variations in forest structure can significantly influence reflectance patterns, as also emphasized by Asner (1998). As such, the spectral components captured in our HLS forest structure dataset include contributions from non-photosynthetic vegetation (e.g., bark and branches), green vegetation, and variations arising from the bidirectional reflectance distribution function (Nicodemus, 1965), including shadow effects– all of which reflect the diverse structures represented in our dataset. To account for this complexity, it was necessary to define forest structure categories incorporating both Canopy Height and Crown Cover.

We classified forest structural categories from HLS spectral features using two Random Forest classifiers trained on our HLS Forest Structure dataset (Enguehard et al., 2025) for peak and late summer. Here again, we find no significant differences in the model performances between the peak and late summer datasets. Classification accuracy of the forest structure categories does not follow a clear gradient from sparse to dense cover. Instead, results are mixed: medium-to-high density classes (50–80% cover) are the most difficult to distinguish, often being confused with each other due to similar spectral characteristics. In contrast, the highest classification accuracy is observed in dense-canopy forest structure classes, which correspond to areas with the highest NIR reflectance values, most strongly expressed in the second principal component, TCG, and also TCW. At the northern forest edge, the lowest density class (20–50% cover) also has higher accuracy, however, showing high NIR and green reflectance even in late summer. This likely reflects the contribution of broadleaf shrubs, which are common in boreal ecosystems and can maintain high greenness despite low canopy closure. For example, some shrubby forest sites on the Seward Peninsula in Alaska retain notable greenness in late summer, in contrast to the overall decline in vegetation vitality observed across most other sites. Overall, our results demonstrate that the HLS forest structure dataset can reliably be used to investigate predictions of forest structural categories, particularly in areas with closed-canopy stands, but also in sparse forests at the northern forest edge. Medium-density stands remain more challenging to classify due to spectral overlap, but still achieve moderate performance with a precision of 0.58-0.62.

4.4. ABoVE tree canopy cover assessment

Montesano et al. (2016) optimized the estimation of Landsat TCC in sparse forests using a low Canopy Height threshold of 2 m derived from airborne LiDAR. In a similar approach, we used a 1 m Canopy Height threshold based on UAV-LiDAR to produce our Crown Cover layer. Montesano et al. (2016) emphasized the importance of including lower vegetation structure, particularly in the treeline ecotones, where the

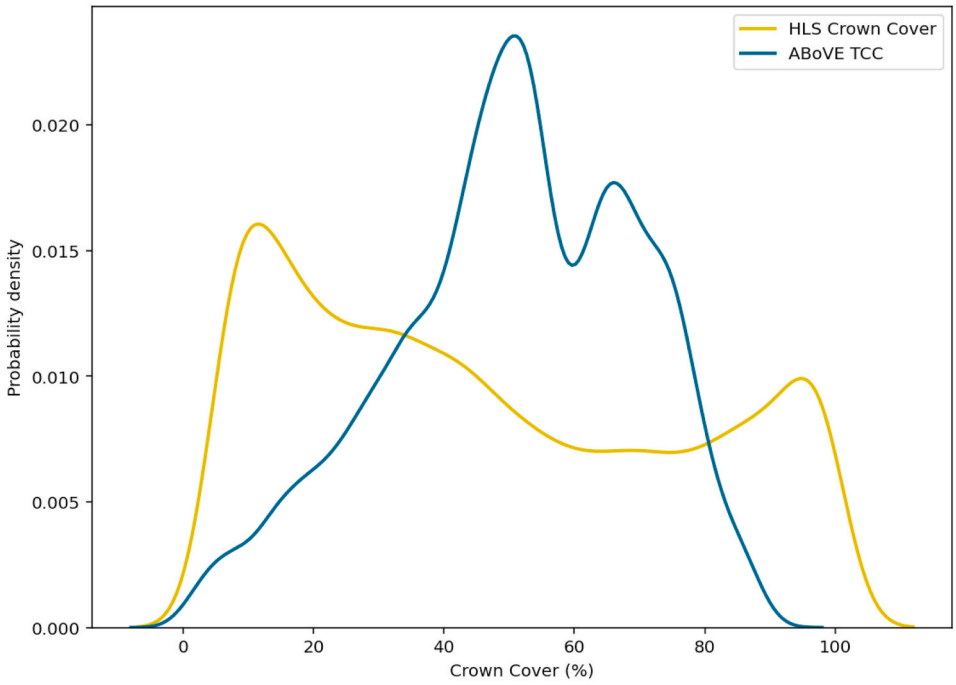

Fig. 12. Density distribution of the Crown Cover from the 30 m resolution ABoVE TCC (blue line) and our UAV-LiDAR observed 30 m aggregated HLS forest structure (yellow line). The y-axis is the probability density and the area under the curves integrates to 1.

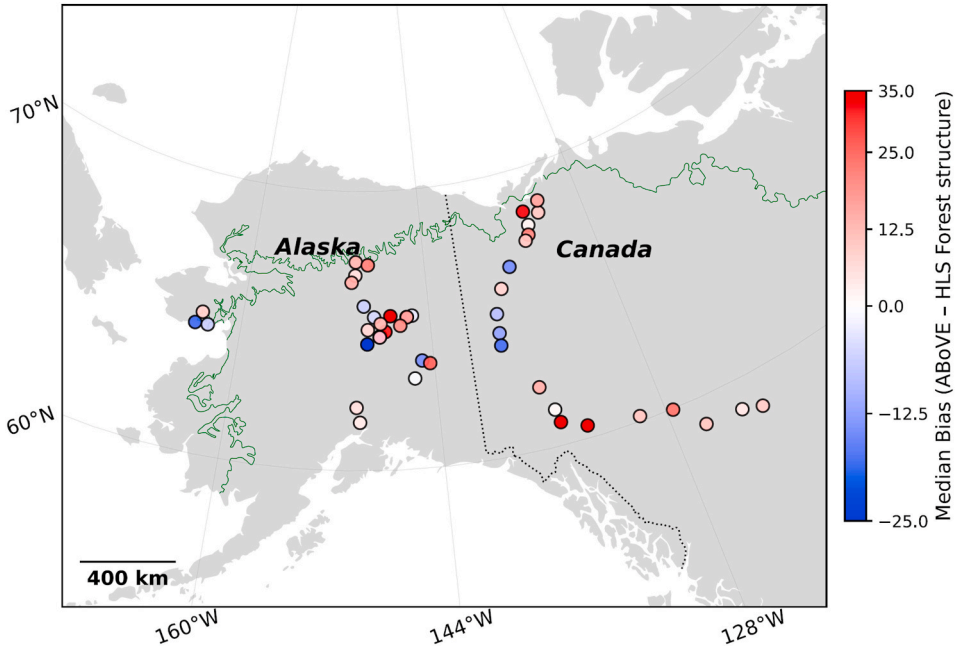

Fig. 13. Geographical distribution of the median bias (ABoVE TCC minus 30 m aggregated UAV LiDAR -derived Crown Cover) across the UAV transects. The exact location of the sites may have been slightly moved to prevent overlap with other pie charts. Red colors indicate an overestimation of Crown Cover by ABoVE TCC while blues indicate an underestimation. The green line represents the Circum-Arctic Vegetation Map (CAVM) treeline (Raynolds et al., 2019), the border between the tundra and the boreal forest.

definition of 'tree" becomes ambiguous and should not rely on the commonly applied global threshold of 5 m (Sexton et al., 2015).

The comparison between ABoVE TCC estimates (Feng et al., 2022) and our 30 m UAV-LiDAR-derived Crown Cover highlights key areas for the ABoVE TCC good performance at most of the denser forest sites, but also regions with larger positive and negative deviations. The observed bias, particularly the ABoVE TCC overestimation of tree canopy cover at the northern forest edge, suggests some limitations in the ABoVE TCC product's ability to capture the challenging spatial variability and sparse tree cover with abundant green understory in these transition zones. Montesano et al. (2016) already noted that global Landsat- and MODIS-derived Crown Cover product estimates suffer from confusion between dense understory vegetation and sparse tree cover. They already considerably reduced this error in their circum-boreal Landsat TCC, which serves as the foundation for the ABoVE TCC. The majority of the overestimation at the northern and elevational treeline seems to stay

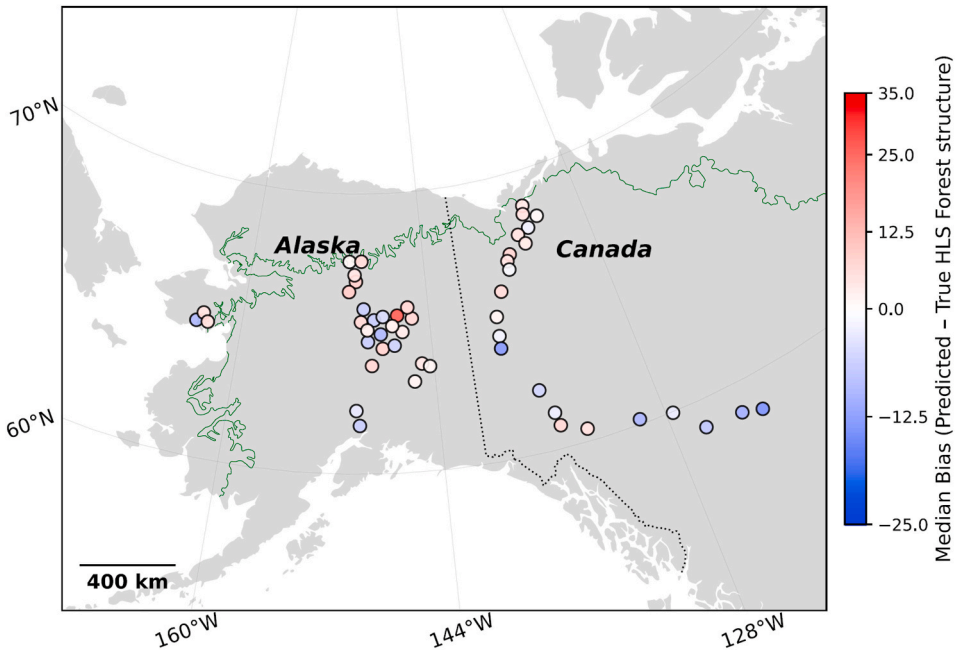

Fig. 14. Geographical distribution of the median bias (HLS-Predicted Crown Cover minus 30 m aggregated UAV LiDAR-derived Crown Cover) across the UAV transects. The exact location of the sites may have been slightly moved to prevent overlap with other pie charts. Red colors indicate an overestimation of Crown Cover by our HLS spectral features-based prediction, while blues indicate an underestimation. The green line represents the Circum-Arctic Vegetation Map (CAVM) treeline (Raynolds et al., 2019), the border between the tundra and the boreal forest.

in the range up to 25 % Crown Cover compared to our Crown Cover above a 1 m reference. On the contrary, ABoVE TCC shows underestimation at most dense forest sites. Montesano et al. (2016) already removed some of the systematic Landsat inherent TCC underestimation above 80% cover (Montesano et al., 2009; Sexton et al., 2013), considerably improving circumboreal Landsat TCC, and successively the ABoVE TCC (Feng et al., 2022). We see that ABoVE TCC underestimation is less pronounced for the densest forest category with trees above 5 m height compared to the dense forest category with a tree layer below 5 m.

Although the ABoVE dataset represents tree canopy cover from the year 2020 (up to four years before our data acquisition), we checked and ruled out disturbances such as wildfire and permafrost thaw at our sites and consider temporal changes in forest structure to be minimal over such a short period. Our findings demonstrate the value of UAV-LiDAR data for improving the accuracy of tree cover mapping and suggest strong potential for its use in training or validating remote sensing algorithms in boreal forests, especially in the northern edge of the boreal forest.

4.5. Perspectives

While our study is based on a limited set of 89 UAV-LiDAR sites, careful site selection ensured representation across key gradients of forest structure and canopy cover at the northern edge of the western North American boreal forest, representing Northern low productivity and unmanaged forests. This spatial diversity helps mitigate concerns about sample size by capturing structural variability relevant to regional-scale modeling.

The use of Harmonized Landsat and Sentinel-2 (HLS) data allows applications at high latitudes by significantly increasing the number of cloud-free observations during the summer season. This is particularly advantageous compared to using data from the Landsat or Sentinel-2 mission alone, which often suffers from limited useable imagery due to persistent cloud cover or wildfire smoke in the summer. However, the trade-off is a reduced spatial resolution (30 m in HLS compared to 10–20 m in Sentinel-2), which may limit the detection of fine-scale structural variation important for localized mapping or small-area analyses. In our study, we explore the relationships between 30 m × 30 m forest structure, i.e., canopy height and cover of the low productivity forests of the Northern Forest edge.

Our results align with the aspirational products identified by Radeloff et al. (2024), namely, forest types and habitat heterogeneity. While we did not generate wall-to-wall satellite products, we demonstrate the potential of upscaling UAV-LiDAR-derived forest types and structure to medium-resolution multispectral satellite data, such as Landsat and Sentinel-2, also for low-structure forest types of the Northern Forest edge. The observed distribution of forest categories across sites enables users to investigate habitat heterogeneity. These results can serve to determine key biodiversity patterns and conservation (Stein et al., 2014; Tuanmu and Jetz, 2015). Our study underscores the value of UAV-LiDAR as a critical link between fine-scale field inventories and coarser-resolution large-scale airborne campaigns such as NASA's ABoVE airborne LiDAR program (Miller et al., 2019; Montesano et al., 2023). By integrating these complementary datasets, UAV-LiDAR and HLS data, we can help bridge spatial gaps to forest stand level scales (e.g., Rautiainen and Heiskanen, 2013) and develop robust, transferable models to support large area boreal forest monitoring.

We provide an openly available training dataset (Enguehard et al., 2025) specifically designed for boreal forest structure at the northern treeline. This HLS forest structure dataset includes UAV-LiDAR-derived Canopy Height and Crown Cover measurements, spatially aggregated to match the 30 m resolution of HLS and Landsat data, ensuring alignment between structural reference data and satellite observations. This consistency enhances the utility of the dataset for training machine learning models and benchmarking satellite-based land cover and structure products, providing a much-needed reference in a region where ground-truth data remains scarce.

5. Conclusions

This study demonstrates the potential of integrating UAV-LiDAR-derived forest structure metrics with Harmonized Landsat and Sentinel-2 (HLS) multispectral imagery to upscale Canopy Height and

Crown Cover across the northern edge of the western North American boreal forest. By leveraging high-resolution UAV-LiDAR sites, we show that HLS spectral information is predictive of structural forest attributes. Specifically, we found that dense and sparse forests reach the highest accuracy, notably near the treeline.

In our study, we also highlight the forest type-driven limitations and challenges. Forest classes characterized by low tree heights (<5 m) and, consequently, low crown cover contain mixed spectral pixels, with a high contribution of vegetation greenness from the green understory that spans the full value range of HLS spectral features. In contrast, low-productivity northern boreal forests exhibit a very narrow height distribution, with the highest abundance of trees below 5 m. This poses a data-driven limitation for accurately quantifying canopy height. Forest structure variables and categories are most reliably predicted in denser forests within our study area. We note that the high proportion of green understory background reduces the sensitivity of HLS spectral features to canopy height, crown cover, and forest structure categories compared to productive forests (tropical, subtropical, temperate, and southern boreal forests).

Our work brings new insights into boreal forests structure at their northern edge, their relationships with Harmonized products like HLS, and potential to be upscaled. By providing an openly available HLS–LiDAR reference dataset and demonstrating scalable modeling approaches, our study contributes a valuable resource for forest structure mapping in data-sparse high-latitude ecosystems. Specifically, we fill a gap for the data-sparse non-managed low-productivity forests at the northern forest edge, covering large regions in North-West America. This work supports the broader integration of UAV-LiDAR in remote sensing workflows and highlights its relevance for refining forest monitoring and carbon accounting in boreal regions, especially at ecotones where accurate structural characterization remains a key challenge.

CRediT authorship contribution statement

Léa Enguehard: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stefan Kruse:** Writing – review & editing, Funding acquisition, Data curation. **Ronny Hänsch:** Writing – review & editing, Methodology. **Ulrike Herzschuh:** Writing – review & editing, Supervision, Funding acquisition. **Santosh Panda:** Writing – review & editing, Data curation. **Birgit Heim:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve the readability and language of the manuscript during the writing phase. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the ERC Consolidator Grant Glacial Legacy (772852), BMWi AI-vergreens, DFG-project No. 44865179, and 'Potsdamer InnoLab für Arktisforschung' No. F221-08-AWI/001/002 from the Brandenburg Ministry for Science, Research and Culture (MWFK) for funding this work. We appreciate the support of the DataHub Information Infrastructure funds for the support of the BorFIT and PC2RCHIVE projects.

Recognizing Indigenous sovereignty, we respectfully acknowledge the Indigenous communities on whose territories we conducted research across Alaska and Canada, namely the Inupiat, Gwich'in Nành, Dënéndeh, Koyukon, Tanana, Tanacross, Upper Tanana, Ahtna Nenn', Dena'ina Ełnena, Alutiiq (Sugpiaq), Hän, Lingít Aaní (Tlingit), Kaska Dena Kayeh, Tagish First Nation (Yukon), Tagish, Na-cho Nyak Dun, Selkirk, Tr'ondëk Hwëch'in, Ta'an Kwäch'än, Kwanlin Dün, Denendeh (Acho Dene Koe), Michif Piyii (Métis), Inuvialuit, Inuit Nunangat, Vuntut Gwich'in, and Tetlit Gwich'in. We are grateful for the opportunity to work in these territories and recognize the deep cultural, ecological, and historical connections Indigenous Peoples have with boreal forests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.srs.2026.100403.

Data availability

The code used to generate results is available on github: https://github.com/leaenguehard/HLS_LiDAR and Zenodo: https://doi.org/10.5281/zenodo.18620809.

The datasets that support this study are published in Pangaea (Felden et al., 2023):

Enguehard et al. (2025): Harmonized Landsat and Sentinel-2 tiles labeled with UAV-LiDAR-derived forest structural attributes in Alaska and Northwest Canada [dataset]. PANGAEA, https://doi.org/10.1594/PANGAEA.983738.

Kruse et al. (2025): Point clouds with ground point classification and individual tree segmentation of 25 northern boreal forest and tundra sites from UAV-based lidar surveys in northwestern Canada in 2022 [dataset]. PANGAEA, https://doi.org/10.1594/PANGAEA.977771.

Kruse et al. (2025): Point clouds with ground point classification and individual tree segmentation of 47 northern boreal forest and tundra sites from UAV-based lidar surveys in the eastern part of Alaska in 2023 [dataset]. PANGAEA, https://doi.org/10.1594/PANGAEA.980485.

Kruse et al. (2025): Point clouds with ground point classification and individual tree segmentation of 28 northern boreal forest and tundra sites from UAV-based lidar surveys in western and central parts of Alaska in 2024 [dataset]. PANGAEA, https://doi.org/10.1594/PANGAEA.980757.

References

Alonzo, M., Andersen, H.-E., Morton, D., Cook, B., 2018. Quantifying boreal forest structure and composition using UAV structure from motion. Forests 9 (3), 119. https://doi.org/10.3390/f9030119.

Astola, H., Häme, T., Sirro, L., Molinier, M., Kilpi, J., 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. Rem. Sens. Environ. 223, 257–273. https://doi.org/10.1016/j.rse.2019.01.019.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Rem. Sens. Environ. 64 (3), 234–253. https://doi.org/10.1016/s0034-4257(98)00014-5.

Baret, F., Weiss, M., Troufleau, D., Prevot, L., Combal, B., 2000. Maximum Information Exploitation for Canopy Characterization by Remote Sensing.

Belmonte, A., Sankey, T., Biederman, J.A., Bradford, J., Goetz, S.J., Kolb, T., Woolley, T., 2020. UAV-derived estimates of forest structure to inform ponderosa pine forest restoration. Remote Sensing in Ecology and Conservation 6 (2), 181–197. https://doi.org/10.1002/rse2.137.

Berner, L.T., Goetz, S.J., 2022. *Arctic-Boreal Vulnerability Experiment (ABoVE)ABoVE: Landsat Vegetation Greenness Trends, Boreal Forest Biome, 1985-2019* (Version 1, p. 0 MB) [CSV, GeoTIFF]. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/2023.

Berra, E.F., Gaulton, R., 2021. Remote sensing of temperate and boreal forest phenology: a review of progress, challenges and opportunities in the intercomparison of in-situ and satellite phenological metrics. For. Ecol. Manag. 480, 118663. https://doi.org/10.1016/j.foreco.2020.118663.

Bolton, D.K., Tompalski, P., Coops, N.C., White, J.C., Wulder, M.A., Hermosilla, T., Queinnec, M., Luther, J.E., van Lier, O.R., Fournier, R.A., Woods, M., Treitz, P.M., van Ewijk, K.Y., Graham, G., Quist, L., 2020. Optimizing landsat time series length for regional mapping of lidar-derived forest structure. Rem. Sens. Environ. 239, 111645. https://doi.org/10.1016/j.rse.2020.111645.

Bolton, D.K., Coops, N.C., Wulder, M.A., 2015. Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using landsat time-series and airborne lidar data. Rem. Sens. Environ. 163, 48–60. https://doi.org/10.1016/j.rse.2015.03.004.

Bonan, G.B., Pollard, D., Thompson, S.L., 1992. Effects of boreal forest vegetation on global climate. Nature 359 (6397), 716–718. https://doi.org/10.1038/359716a0.

Brandt, J.P., Flannigan, M.D., Maynard, D.G., Thompson, I.D., Volney, W.J.A., 2013. An introduction to Canada's boreal zone: ecosystem processes, health, sustainability, and environmental issues. Environ. Rev. 21 (4), 207–226. https://doi.org/10.1139/er-2013-0040.

Breiman, L., 2001. Random forest. In: Machine Learning. Kluwer Academic Publishers, pp. 5–32.

Brieger, F., Herzschuh, U., Pestryakova, L.A., Bookhagen, B., Zakharov, E.S., Kruse, S., 2019. Advances in the derivation of Northeast Siberian Forest metrics using high-resolution UAV-based photogrammetric point clouds. Remote Sens. 11 (12), 1447. https://doi.org/10.3390/rs11121447.

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., Smets, B., 2020a. Copernicus global land cover layers—collection 2. Remote Sens. 12 (6), 1044. https://doi.org/10.3390/rs12061044.

Buchhorn, Smets, Bertels, Roo, Lesiv, Tsendbazar, Herold, Fritz, 2020b. Copernicus global land service: Land Cover 100m: collection 3: epoch 2019: Globe (Version V3.0.1). Zenodo [Dat aset]. https://zenodo.org/records/3939050.

Chapin, F.S., Yarie, J., Van Cleve, K., Viereck, L.A., 2006. The conceptual basis of LTER studies in the Alaskan boreal forest. In: Chapin, F.S., Oswood, M.W., Van Cleve, K., Viereck, L.A., Verbyla, D.L. (Eds.), Alaska's Changing Boreal Forest. Oxford University Press, pp. 3–11. https://doi.org/10.1093/oso/9780195154313.003.0012.

Chen, N., Tsendbazar, N.-E., Hamunyela, E., Verbesselt, J., Herold, M., 2021. Sub-annual tropical forest disturbance monitoring using harmonized Landsat and Sentinel-2 data. Int. J. Appl. Earth Obs. Geoinf. 102, 102386. https://doi.org/10.1016/j.jag.2021.102386.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., Justice, C., 2018. The harmonized landsat and Sentinel-2 surface reflectance data set. Rem. Sens. Environ. 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002.

Cohen, W., Spies, T., 1992. Estimating structural attributes of douglas-fir/western hemlock forest stands from landsat and SPOT imagery. Rem. Sens. Environ. 41 (1), 1–17. https://doi.org/10.1016/0034-4257(92)90056-.

Coops, N.C., Tompalski, P., Goodbody, T.R.H., Queinnec, M., Luther, J.E., Bolton, D.K., White, J.C., Wulder, M.A., van Lier, O.R., Hermosilla, T., 2021. Modelling lidar-derived estimates of forest attributes over space and time: a review of approaches and future trends. Rem. Sens. Environ. 260, 112477. https://doi.org/10.1016/j.rse.2021.112477.

Delcourt, C.J.F., Combee, A., Izbicki, B., Mack, M.C., Maximov, T., Petrov, R., Rogers, B. M., Scholten, R.C., Shestakova, T.A., Van Wees, D., Veraverbeke, S., 2021. Evaluating the differenced normalized burn ratio for assessing fire severity using Sentinel-2 imagery in northeast Siberian larch forests. Remote Sens. 13 (12), 2311. https://doi.org/10.3390/rs13122311.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing 1, 100002. https://doi.org/10.1016/j.srs.2020.100002.

Dymond, C.C., Mladenoff, D.J., Radeloff, V.C., 2002. Phenological differences in Tasseled cap indices improve deciduous forest classification. Rem. Sens. Environ. 80 (3), 460–472. https://doi.org/10.1016/S0034-4257(01)00324-8.

Enguehard, L., Kruse, S., Farkas, L., Gloy, J., Heim, B., 2025. *Harmonized Landsat and Sentinel-2 tiles labeled with UAV-LiDAR-derived forest structural attributes in Alaska and Northwest Canada* (p. 438 data points). PANGAEA [Data set]. https://doi.org/10.1594/PANGAEA.983738.

Enguehard, L., Kruse, S., Heim, B., Nitze, I., Falco, N., Demir, B., Herzschuh, U., 2024. Unraveling boreal forest composition and drivers across scales in eastern Siberia. Environ. Res. Lett. 19 (7), 074050. https://doi.org/10.1088/1748-9326/ad5742.

ESA, 2017. Land cover CCI product user guide version 2 (Version 2) [Tech. Rep.]. maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.

Fakhri, A., Latifi, H., Samani, K.M., Fassnacht, F.E., 2025. Introducing a computationally light approach to estimate forest height and fractional canopy cover from Sentinel-2 data. J. Arid Environ. 228, 105343. https://doi.org/10.1016/j.jaridenv.2025.105343.

Fakhri, A., Sayadi, S., Naghavi, H., Latifi, H., 2022. A novel vegetation index-based workflow for semi-arid, sparse woody cover mapping. J. Arid Environ. 201, 104748. https://doi.org/10.1016/j.jaridenv.2022.104748.

Fakhri, S.A., Sayadi, S., Latifi, H., Khare, S., 2019. An optimized enhanced vegetation index for sparse tree cover mapping across a mountainous region. 2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), pp. 146–151. https://doi.org/10.1109/MetroAgriFor.2019.8909259.

Farkas, L., Kruse, S., 2024. *StefanKruse/PC2RCHIVE: PC2RCHIVE: a FAIR and easy point cloud archiving GUI* (Version v1.0). Zenodo. https://doi.org/10.5281/ZENODO.14534516 [Computer software].

Fassnacht, F.E., White, J.C., Wulder, M.A., Næsset, E., 2024. Remote sensing in forestry: current challenges, considerations and directions. Forestry: An Int.

Felden, J., Möller, L., Schindler, U., Huber, R., Schumacher, S., Koppe, R., Diepenbroek, M., Glöckner, F.O., 2023. PANGAEA - Data Publisher for Earth & environmental science. Sci. Data 10 (1), 347. https://doi.org/10.1038/s41597-023-02269-x.

Feng, Sexton, J.O., Wang, P., Channan, S., Montesano, P.M., Wagner, W., Wooten, M.R., Neigh, C.S., 2022. ABoVE: tree canopy cover and stand age from landsat, boreal forest biome, 1984-2020. https://doi.org/10.3334/ORNLDAAC/2012.

Feng, T., Duncanson, L., Montesano, P., Hancock, S., Minor, D., Guenther, E., Neuenschwander, A., 2023. A systematic evaluation of multi-resolution ICESat-2 ATL08 terrain and canopy heights in boreal forests. Rem. Sens. Environ. 291, 113570. https://doi.org/10.1016/j.rse.2023.113570.

Frazier, R.J., Coops, N.C., Wulder, M.A., Kennedy, R., 2014. Characterization of aboveground biomass in an unmanaged boreal forest using landsat temporal segmentation metrics. ISPRS J. Photogrammetry Remote Sens. 92, 137–146. https://doi.org/10.1016/j.isprsjprs.2014.03.003.

French, N.H.F., Kasischke, E.S., Hall, R.J., Murphy, K.A., Verbyla, D.L., Hoy, E.E., Allen, J.L., 2008. Using landsat data to assess fire and burn severity in the North American boreal forest region: an overview and summary of results. Int. J. Wildland Fire 17 (4), 443. https://doi.org/10.1071/WF08007.

Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A.Z., Schepaschenko, D.G., 2015. Boreal forest health and global change. Science 349 (6250), 819–822. https://doi.org/10.1126/science.aaa9092.

Goel, N., 1988. Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. Remote Sens. Rev. 4 (1), 1–212. https://doi.org/10.1080/02757258809532105.

Goel, N., Thompson, R., 2000. A snapshot of canopy reflectance models and a universal model for the radiation regime. Remote Sens. Rev. 18 (2–4), 197–225.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth engine: Planetary-scale geospatial analysis for everyone. Rem. Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest stand species mapping using the Sentinel-2 time series. Remote Sens. 11 (10), 1197. https://doi.org/10.3390/rs11101197.

Guimarães, N., Pádua, L., Marques, P., Silva, N., Peres, E., Sousa, J.J., 2020. Forestry remote sensing from unmanned aerial vehicles: a review focusing on the data, processing and potentialities. Remote Sens. 12 (6), 1046. https://doi.org/10.3390/rs12061046.

Hall, R.J., Skakun, R.S., Arsenault, E.J., Case, B.S., 2006. Modeling forest stand structure attributes using landsat ETM+ data: application to mapping of aboveground biomass and stand volume. For. Ecol. Manag. 225 (1–3), 378–390. https://doi.org/10.1016/j.foreco.2006.01.014.

Healey, S., Cohen, W., Zhiqiang, Y., Krankina, O., 2005. Comparison of tasseled Cap-based landsat data structures for use in forest disturbance detection. Rem. Sens. Environ. 97 (3), 301–310. https://doi.org/10.1016/j.rse.2005.05.009.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-Century forest cover change. Science 342 (6160), 850–853. https://doi.org/10.1126/science.1244693.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., Carroll, M., 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. The Moderate Resolution Imaging Spectroradiometer (MODIS): A New Generation of Land Surface Monitoring 83 (1), 303–331.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., 2022. Land cover classification in an era of big and open data: optimizing localized implementation and training data selection to improve mapping outcomes. Rem. Sens. Environ. 268, 112780. https://doi.org/10.1016/j.rse.2021.112780.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2017. Updating landsat time series of surface-reflectance composites and forest change products with new observations. Int. J. Appl. Earth Obs. Geoinf. 63, 104–111. https://doi.org/10.1016/j.jag.2017.07.013.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2018. Disturbance-informed annual land cover classification maps of Canada's forested ecosystems for a 29-Year landsat time series. Can. J. Rem. Sens. 44 (1), 67–87. https://doi.org/10.1080/07038992.2018.1437719.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Campbell, L.B., 2016. Mass data processing of time series landsat imagery: pixels to data products for forest monitoring. Int. J. Digit. Earth 9 (11), 1035–1054. https://doi.org/10.1080/17538947.2016.1187673.

Hudak, A.T., Lefsky, M.A., Cohen, W.B., Berterretche, M., 2002. Integration of lidar and landsat ETM+ data for estimating and mapping forest canopy height. Rem. Sens. Environ. 82 (2), 397–416. https://doi.org/10.1016/S0034-4257(02)00056-1.

Ju, J., Zhou, Q., Freitag, B., Roy, D.P., Zhang, H.K., Sridhar, M., Mandel, J., Arab, S., Schmidt, G., Crawford, C.J., Gascon, F., Strobl, P.A., Masek, J.G., Neigh, C.S.R., 2025. The harmonized landsat and Sentinel-2 version 2.0 surface reflectance dataset. Rem. Sens. Environ. 324, 114723. https://doi.org/10.1016/j.rse.2025.114723.

Kasischke, E.S., 2000. Boreal ecosystems in the global carbon cycle. In: Kasischke, E.S., Stocks, B.J. (Eds.), Fire, Climate Change, and Carbon Cycling in the Boreal Forest, 138. Springer, New York, pp. 19–30. https://doi.org/10.1007/978-0-387-21629-4_2.

Kauth, R.J., Thomas, G.S., 1976. The Tasselled cap—A Graphic Description of the spectral-temporal Development of Agricultural Crops as Seen by Landsat. *LARS Symposia*, 159.

Kruse, S., Enguehard, L., Juday, G., Santosh, P., Badola, A., Broers, J., Farkas, L., Schladebach, J., Hao, K., Jackisch, R., Döpper, V., Heim, B., 2025a. *Point Clouds with*

Ground Point Classification and Individual Tree Segmentation of 28 Northern Boreal Forest and Tundra Sites from UAV-based Lidar Surveys in Western and Central Parts of Alaska in 2024 (P. 1316 Data Points) [Text/tab-separated-values]. PANGAEA. https://doi.org/10.1594/PANGAEA.980757.

Kruse, S., Enguehard, L., Juday, G., Santosh, P., Badola, A., Broers, J., Farkas, L., Schladebach, J., Hao, K., Jackisch, R., Döpper, V., Heim, B., 2025b. *Point Clouds with Ground Point Classification and Individual Tree Segmentation of 47 Northern Boreal Forest and Tundra Sites from UAV-based Lidar Surveys in the Eastern Part of Alaska in 2023* (P. 2204 Data Points). PANGAEA. https://doi.org/10.1594/PANGAEA.980485 [Text/tab-separated-values].

Kruse, S., Gloy, J., Farkas, L., Schladebach, J., Hao, K., Döpper, V., Heim, B., 2025. *Point Clouds with Ground Point Classification and Individual Tree Segmentation of 25 Northern Boreal Forest and Tundra Sites from UAV-based Lidar Surveys in Northwestern Canada in 2022* (P. 1123 Data Points). PANGAEA. https://doi.org/10.1594/PANGAEA.977771 [Text/tab-separated-values].

Lang, N., Schindler, K., Wegner, J.D., 2019. Country-wide high-resolution vegetation height mapping with Sentinel-2. Rem. Sens. Environ. 233, 111347. https://doi.org/10.1016/j.rse.2019.111347.

LaRue, E.A., Atkins, J.W., Dahlin, K., Fahey, R., Fei, S., Gough, C., Hardiman, B.S., 2018. Linking landsat to terrestrial LiDAR: vegetation metrics of forest greenness are correlated with canopy structural complexity. Int. J. Appl. Earth Obs. Geoinf. 73, 420–427. https://doi.org/10.1016/j.jag.2018.07.001.

Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A., Gower, S.T., 2002. LiDAR remote sensing of above-ground biomass in three biomes. Global Ecol. Biogeogr. 11 (5), 393–399. https://doi.org/10.1046/j.1466-822x.2002.00303.x.

Li, L., Mu, X., Jiang, H., Chianucci, F., Hu, R., Song, W., Qi, J., Liu, S., Zhou, J., Chen, L., Huang, H., Yan, G., 2023. Review of ground and aerial methods for vegetation cover fraction (fCover) and related quantities estimation: definitions, advances, challenges, and future perspectives. ISPRS J. Photogrammetry Remote Sens. 199, 133–156. https://doi.org/10.1016/j.isprsjprs.2023.03.020.

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., Chen, H., 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. Int. J. Appl. Earth Obs. Geoinf. 92, 102163. https://doi.org/10.1016/j.jag.2020.102163.

Liu, K., Shen, X., Cao, L., Wang, G., Cao, F., 2018. Estimating forest structural attributes using UAV-LiDAR data in ginkgo plantations. ISPRS J. Photogrammetry Remote Sens. 146, 465–482. https://doi.org/10.1016/j.isprsjprs.2018.11.001.

Lu, J., Wang, H., Qin, S., Cao, L., Pu, R., Li, G., Sun, J., 2020. Estimation of aboveground biomass of Robinia pseudoacacia forest in the yellow river Delta based on UAV and backpack LiDAR point clouds. Int. J. Appl. Earth Obs. Geoinf. 86, 102014. https://doi.org/10.1016/j.jag.2019.102014.

Maesano, M., Santopuoli, G., Moresi, F., Matteucci, G., Lasserre, B., Scarascia Mugnozza, G., 2022. Above ground biomass estimation from UAV high resolution RGB images and LiDAR data in a pine forest in Southern Italy. iFor. Biogeosci. For. 15 (6), 451–457. https://doi.org/10.3832/ifor3781-015.

Majasalmi, T., Rautiainen, M., 2016. The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: a simulation study. Remote Sensing Letters 7 (5), 427–436. https://doi.org/10.1080/2150704X.2016.1149251.

Margolis, H.A., Nelson, R.F., Montesano, P.M., Beaudoin, A., Sun, G., Andersen, H.-E., Wulder, M.A., 2015. Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. Can. J. For. Res. 45 (7), 838–855. https://doi.org/10.1139/cjfr-2015-0006.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., et al., 2017. The ice, cloud, and land elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation. Rem. Sens. Environ. 190, 260–273. https://doi.org/10.1016/j.rse.2016.12.029.

Masek, J., 2001. Stability of boreal forest stands during recent climate change: evidence from landsat satellite imagery. J. Biogeogr. 28 (8), 967–976. https://doi.org/10.1046/j.1365-2699.2001.00612.x.

Masek, J., Ju, J., Roger, J.-C., Skakun, S., Vermote, E., Claverie, M., Dungan, J., Yin, Z., Freitag, B., Justice, C., 2021. HLS operational land imager surface reflectance and TOA brightness daily global 30m v2.0. NASA Land Processes Distributed Active Archive Center [Data set]. https://doi.org/10.5067/HLS/HLSL30.002.

Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Zald, H.S.J., 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using landsat composites and lidar plots. Rem. Sens. Environ. 209, 90–106. https://doi.org/10.1016/j.rse.2017.12.020.

Miller, C.E., Griffith, P.C., Goetz, S.J., Hoy, E.E., Pinto, N., McCubbin, I.B., Thorpe, A.K., Hofton, M., Hodkinson, D., Hansen, C., Woods, J., Larson, E., Kasischke, E.S., Margolis, H.A., 2019. An overview of ABoVE airborne campaign data acquisitions and science opportunities. Environ. Res. Lett. 14 (8), 080201. https://doi.org/10.1088/1748-9326/ab0d44.

Montesano, Neigh, C.S.R., Macander, M.J., Wagner, W., Duncanson, L.I., Wang, P., Sexton, J.O., Miller, C.E., Armstrong, A.H., 2023. Patterns of regional site index across a North American boreal forest gradient. Environ. Res. Lett. 18 (7), 075006. https://doi.org/10.1088/1748-9326/acdcab.

Montesano, P.M., Neigh, C.S.R., Macander, M., Feng, M., Noojipady, P., 2020. The bioclimatic extent and pattern of the cold edge of the boreal forest: the circumpolar taiga-tundra ecotone. Environ. Res. Lett. 15 (10), 105019. https://doi.org/10.1088/1748-9326/abb2c7.

Montesano, Neigh, C., Sexton, J., Feng, M., Channan, S., Ranson, K., Townshend, J., 2016. Calibration and validation of landsat tree cover in the taiga–tundra ecotone. Remote Sens. 8 (7), 551. https://doi.org/10.3390/rs8070551.

Montesano, P., Nelson, R., Sun, G., Margolis, H., Kerber, A., Ranson, K.J., 2009. MODIS tree cover validation for the circumpolar taiga–tundra transition zone. Rem. Sens. Environ. 113 (10), 2130–2141. https://doi.org/10.1016/j.rse.2009.05.021.

Mulverhill, C., Coops, N.C., Achim, A., 2023. Continuous monitoring and sub-annual change detection in high-latitude forests using harmonized landsat Sentinel-2 data. ISPRS J. Photogrammetry Remote Sens. 197, 309–319. https://doi.org/10.1016/j.isprsjprs.2023.02.002.

Neuenschwander, A., Pitts, K., Jelley, B., Robbins, J., Klotz, B., Popescu, S., Nelson, R., Harding, D., Pederson, D., & Sheridan, R. (2021). ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 5 [Data set]. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/ATLAS/ATL08.005.

Neuenschwander, A., Guenther, E., White, J.C., Duncanson, L., Montesano, P., 2020. Validation of ICESat-2 terrain and canopy heights in boreal forests. Rem. Sens. Environ. 251, 112110. https://doi.org/10.1016/j.rse.2020.112110.

Nicodemus, F.E., 1965. Directional reflectance and emissivity of an opaque surface. Appl. Opt. 4 (7), 767. https://doi.org/10.1364/ao.4.000767.

Omernik, James, Griffith, G., 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environ. Manag. 54 (6), 1249–1266. https://doi.org/10.1007/s00267-014-0364-1.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2025–2030.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V., Murayama, Y., Ranagalage, M., 2020. Sentinel-2 data for land cover/use mapping: a review. Remote Sens. 12 (14), 2291. https://doi.org/10.3390/rs12142291.

Pickell, P.D., Hermosilla, T., Frazier, R.J., Coops, N.C., Wulder, M.A., 2016. Forest recovery trends derived from landsat time series for North American boreal forests. Int. J. Rem. Sens. 37 (1), 138–149. https://doi.org/10.1080/2150704X.2015.1126375.

Pisek, J., Chen, J.M., 2009. Mapping forest background reflectivity over North America with Multi-angle imaging SpectroRadiometer (MISR) data. Rem. Sens. Environ. 113 (11), 2412–2423. https://doi.org/10.1016/j.rse.2009.07.003.

Pisek, J., Erb, A., Korhonen, L., Biermann, T., Carrara, A., Cremonese, E., Cuntz, M., Fares, S., Gerosa, G., Grünwald, T., Hase, N., Heliasz, M., Ibrom, A., Knohl, A., Kobler, J., Kruijt, B., Lange, H., Leppänen, L., Limousin, J.-M., et al., 2021. Retrieval and validation of forest background reflectivity from daily moderate resolution imaging Spectroradiometer (MODIS) bidirectional reflectance distribution function (BRDF) data across European forests. Biogeosciences 18 (2), 621–635. https://doi.org/10.5194/bg-18-621-2021.

Plowright, A., 2023. Forest@: Tools for Analyzing Remote Sensing Forest Data. R package version 1.0.2 [Computer software], Version 1.0.2. https://github.com/andrew-plowright/ForestTools.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and landsat data. Rem. Sens. Environ. 253, 112165. https://doi.org/10.1016/j.rse.2020.112165.

Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., Ohmann, J.L., 2010. Quantification of live aboveground forest biomass dynamics with landsat time-series and field inventory data: a comparison of empirical modeling approaches. Rem. Sens. Environ. 114 (5), 1053–1068. https://doi.org/10.1016/j.rse.2009.12.018.

Qiu, S., Zhu, Z., He, B., 2019. Fmask 4.0: improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. Rem. Sens. Environ. 231, 111205. https://doi.org/10.1016/j.rse.2019.05.024.

Radeloff, V.C., Roy, D.P., Wulder, M.A., Anderson, M., Cook, B., Crawford, C.J., Friedl, M., Gao, F., Gorelick, N., Hansen, M., Healey, S., Hostert, P., Hulley, G., Huntington, J.L., Johnson, D.M., Neigh, C., Lyapustin, A., Lymburner, L., Pahlevan, N., et al., 2024. Need and vision for global medium-resolution Landsat and Sentinel-2 data products. Rem. Sens. Environ. 300, 113918. https://doi.org/10.1016/j.rse.2023.113918.

Rautiainen, M., Heiskanen, J., 2013. Seasonal contribution of understory vegetation to the reflectance of a boreal landscape at different spatial scales. IEEE Geoscience and Remote Sensing Letters 10 (4), 923–927. https://doi.org/10.1109/lgrs.2013.2247560.

Raynolds, M.K., Walker, D.A., Balser, A., Bay, C., Campbell, M., Cherosov, M.M., Daniëls, F.J.A., Eidesen, P.B., Ermokhina, K.A., Frost, G.V., Jedrzejek, B., Jorgenson, M.T., Kennedy, B.E., Kholod, S.S., Lavrinenko, I.A., Lavrinenko, O.V., Magnússon, B., Matveyeva, N.V., Metúsalemsson, S., et al., 2019. A raster version of the circumpolar arctic vegetation map (CAVM). Rem. Sens. Environ. 232, 111297. https://doi.org/10.1016/j.rse.2019.111297.

Roberts, D.A., Ustin, S.L., Ogunjemiyo, S., Greenberg, J., Dobrowski, S.Z., Chen, J., Hinckley, T.M., 2004. Spectral and structural measures of Northwest forest vegetation at leaf to landscape scales. Ecosystems 7 (5). https://doi.org/10.1007/s10021-004-0144-5.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., Harlan, J.C., 1974. *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation* [Type II I report for the period September 1972-November 1974]. Remote Sensing Center Texas A&M University College Station, Texas 77843; Goddard Space Flight Center Greenbelt, Maryland20771.

Ruaro, R., Gubiani, É.A., Hughes, R., 2024. Omernik's ecoregion framework: a legacy for understanding regional patterns in attainable resource quality. Environ. Manag. 73 (2), 354–364. https://doi.org/10.1007/s00267-023-01871-4.

Senf, C., Mori, A.S., Müller, J., Seidl, R., 2020. The response of canopy height diversity to natural disturbances in two temperate forest landscapes. Landsc. Ecol. 35 (9), 2101–2112. https://doi.org/10.1007/s10980-020-01085-7.

Sexton, J.O., Noojipady, P., Anand, A., Song, X.-P., McMahon, S., Huang, C., Feng, M., Channan, S., Townshend, J.R., 2015. A model for the propagation of uncertainty from continuous estimates of tree cover to categorical forest cover and change. Rem. Sens. Environ. 156, 418–425. https://doi.org/10.1016/j.rse.2014.08.038.

Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., DiMiceli, C., Townshend, J.R., 2013. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. Int. J. Digit. Earth 6 (5), 427–448. https://doi.org/10.1080/17538947.2013.786146.

Shugart, H.H., Saatchi, S., Hall, F.G., 2010. Importance of structure and its measurement in quantifying function of forest ecosystems. J. Geophys. Res.: Biogeosciences 115 (G2). https://doi.org/10.1029/2009JG000993, 2009JG000993.

Shimizu, K., Ota, T., Mizoue, N., Saito, H., 2020. Comparison of multi-temporal PlanetScope data with landsat 8 and Sentinel-2 data for estimating airborne LiDAR derived canopy height in temperate forests. Remote Sens. 12 (11), 1876. https://doi.org/10.3390/rs12111876.

Stein, A., Gerstner, K., Kreft, H., 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecol. Lett. 17 (7), 866–880. https://doi.org/10.1111/ele.12277.

Stenberg, P., Rautiainen, M., Manninen, T., Voipoi, P., Mottus, M., 2008. Boreal forest leaf area index from optical satellite images: model simulations and empirical analyses using data from central Finland. Boreal Environ. Res. 13 (5).

Tuanmu, M., Jetz, W., 2015. A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. Global Ecol. Biogeogr. 24 (11), 1329–1339. https://doi.org/10.1111/geb.12365.

Van Geffen, F., Hänsch, R., Demir, B., Kruse, S., Herzschuh, U., Heim, B., 2025. A benchmark dataset for Sentinel-2 based forest type classification in the Siberian Summergreen-Evergreen forest transition zone. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 1–22. https://doi.org/10.1109/jstars.2025.3562912.

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. Rem. Sens. Environ. 16 (2), 125–141.

Vermote, E., Justice, C., Claverie, M., Franch, B., 2016. Preliminary analysis of the performance of the landsat 8/OLI land surface reflectance product. Landsat 8 Science Results 185, 46–56. https://doi.org/10.1016/j.rse.2016.04.008.

Villoslada, M., Berner, L.T., Juutinen, S., Ylänne, H., Kumpula, T., 2024. Upscaling vascular aboveground biomass and topsoil moisture of subarctic fens from unoccupied aerial vehicles (UAVs) to satellite level. Sci. Total Environ. 933, 173049. https://doi.org/10.1016/j.scitotenv.2024.173049.

Wang, J.A., Sulla-Menashe, D., Woodcock, C.E., Sonnentag, O., Keeling, R.F., Friedl, M. A., 2019. Above: Landsat-Derived Annual Dominant Land Cover Across Above Core Domain, 1984-2014. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1691, 15817.400714 MB.

Wang, J.A., Sulla-Menashe, D., Woodcock, C.E., Sonnentag, O., Keeling, R.F., Friedl, M. A., 2020. Extensive land cover change across arctic–boreal northwestern North America from disturbance and climate forcing. Glob. Change Biol. 26 (2), 807–822. https://doi.org/10.1111/gcb.14804.

White, J.C., Coops, N.C., Wulder, M.A., Vastaranta, M., Hilker, T., Tompalski, P., 2016. Remote sensing technologies for enhancing forest inventories: a review. Can. J. Rem. Sens. 42 (5), 619–641. https://doi.org/10.1080/07038992.2016.1207484.

Wu, Q., 2020. Geemap: a python package for interactive mapping with google Earth engine. J. Open Source Softw. 5 (51), 2305. https://doi.org/10.21105/joss.02305.

Wulder, M.A., Hermosilla, T., Stinson, G., Gougeon, F.A., White, J.C., Hill, D.A., Smiley, B.P., 2020. Satellite-based time series land cover and change information to map forest area consistent with national and international reporting requirements. Forestry: Int. J. Financ. Res. 93 (3), 331–343. https://doi.org/10.1093/forestry/cpaa006.

Wulder, M.A., Roy, D.P., Radeloff, V.C., Loveland, T.R., Anderson, M.C., Johnson, D.M., Healey, S., Zhu, Z., Scambos, T.A., Pahlevan, N., Hansen, M., Gorelick, N., Crawford, C.J., Masek, J.G., Hermosilla, T., White, J.C., Belward, A.S., Schaaf, C., Woodcock, C.E., et al., 2022. Fifty years of landsat science and impacts. Rem. Sens. Environ. 280, 113195. https://doi.org/10.1016/j.rse.2022.113195.

Zhai, Y., Roy, D.P., Martins, V.S., Zhang, H.K., Yan, L., Li, Z., 2022. Conterminous United States Landsat-8 top of atmosphere and surface reflectance tasseled cap transformation coefficients. Rem. Sens. Environ. 274, 112992. https://doi.org/10.1016/j.rse.2022.112992.