Science of Remote Sensing 11 (2025) 100192

Contents lists available at ScienceDirect

Science of Remote Sensing

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

Synthesizing field plot and airborne remote sensing data to enhance national forest inventory mapping in the boreal forest of Interior Alaska

Pratima Khatri-Chhetri [a,*], Hans-Erik Andersen [b], Bruce Cook [c], Sean M. Hendryx [d], Liz van Wagtendonk [a], Van R. Kane [a]

[a] School of Environmental and Forest Sciences, University of Washington, Anderson Hall, 3715 W Stevens Way NE, Seattle, WA, 98195, USA
[b] USDA Forest Service, Pacific Northwest Research Station, Seattle, WA, USA
[c] NASA, Goddard Space Flight Center, Biospheric Sciences Laboratory, Greenbelt, MD, 20771, USA
[d] Scale AI, 155 5th St, San Francisco, CA, 94103, USA

ARTICLE INFO

Keywords:
Boreal forest
Deep learning
G-LiHT
Lidar
Forest type mapping

ABSTRACT

The boreal biome, the largest terrestrial biome on Earth, is increasingly vulnerable to climate change due to warming twice as rapidly as the global average. Climate change has increased the temperature, frequency, severity, and amount of area burned, which is leading to changes in the spatial extent of forest type and species range. These rapid ecological shifts necessitate fine-scale monitoring of forest type to detect potential type conversions and guide management interventions. In this study, we present a framework for forest type classification combining field plots and high-resolution remote sensing data using machine learning models in the boreal forest of Interior Alaska. For this purpose, we conducted forest type classification at three different levels, including 1. forest and nonforest, 2. hardwood, softwood, and nonforest, and 3. three dominant forest types, including paper birch, black spruce, white spruce, and nonforest. To achieve this goal, we compared the performance of two advanced modeling approaches, the convolutional neural network (CNN) and the XGBoost model. Our datasets included field and high-resolution topographic metrics including elevation, slope, aspect, and solar radiation and canopy height derived from lidar (1 m) and 44 vegetation indices derived from high-resolution (1 m) visible to near infrared (VNIR) hyperspectral data collected by NASA Goddard's Lidar, Hyperspectral and Thermal Imager (G-LiHT) sensor. The remote sensing data were collected under variable sky conditions (clear to overcast) throughout a 1-month growing-season period, and field data collected by United States Department of Agriculture (USDA) Forest Service, Forest Inventory and Analysis program (FIA). In this framework, we also studied the importance of topographic and remote sensing variables for the classification of forest types. We found the CNN model outperformed the XGBoost model in terms of overall accuracy and a macro average F1 score for all three different forest type classifications. The CNN model achieved an overall accuracy of 93.1% for forest or nonforest, 82.6% for hardwood, softwood, and nonforest, and 74.7% for three dominant forest types including paper birch, black spruce, and white spruce along with nonforest. Among the various topographic factors, we found that elevation was the most important factor for discriminating all forest types. In addition, we found that canopy height and vegetation indices including Photochemical Reflectance Index (PRI) (R531 & R570), Pigment Specific Normalized Difference (PSND) (R635 & R800), and Gitelson and Merzlyak (GM1) (R550 & R750) were important for differentiating between hardwood and softwood while Anthocyanin Reflectance Index (ARI1) (R550 & R700) was important for differentiating between forest and nonforest. The high-resolution forest type information can improve our ecological understanding of boreal forest dynamics, estimate above ground biomass, and carbon, and support the national forest inventory and forest managers.

* Corresponding author.
E-mail addresses: upanpra@uw.edu (P. Khatri-Chhetri), hans.andersen@usda.gov (H.-E. Andersen), bruce.cook@nasa.gov (B. Cook), sean.hendryx@scale.com (S.M. Hendryx), vanwag3@uw.edu (L. van Wagtendonk), vkane@uw.edu (V.R. Kane).

https://doi.org/10.1016/j.srs.2024.100192
Received 16 August 2024; Received in revised form 29 November 2024; Accepted 28 December 2024
Available online 30 December 2024

1. Introduction

The boreal forest biome is the largest terrestrial biome, covering nearly 14% (18.5 million km^2) of the entire vegetation cover of the Earth (Douglas et al., 2014; Mcguire & Chapin III, 2006). It plays a critical role in the global carbon cycle, storing nearly one-third of the terrestrial carbon (Pan et al., 2011; Walker et al., 2020). However, the future trajectory of boreal forest ecosystems remains uncertain due to increasing disturbances, thawing permafrost, changes in hydrology and warming. These changes are being observed in the boreal zone across the globe (Ekman et al., 2024).

This is particularly evident in boreal forest of Interior Alaska (Chapin et al., 2010; Johnstone et al., 2010), where temperatures in this region have surged nearly five times more than global mean over the past five decades, (Euskirchen et al., 2010). These dramatic changes led to an increase in fire extent, proliferation of deciduous trees following fires, and a decline in the growth and extent of dominant conifer tree species like spruce (Beck et al., 2011; Chapin et al., 2010; Johnstone et al., 2010; Walker et al., 2023). Shifts in forest type will change ecological processes and services, impact carbon uptake, and will have significant implications for the carbon and nutrient cycle, understory vegetation, and wildlife habitat (Barber et al., 2000; Johnstone et al., 2020; Kasischke et al., 2010). These rapid ecological changes require fine-scale monitoring of forests to detect potential type conversions and guide management interventions.

The boreal forest of Interior Alaska is characterized by a mosaic of coniferous and deciduous forest interspersed with nonforest patches (Fig. 1). This region accounts for approximately 17% of the total forested area in the United States which lacks spatially extensive fine scale forest type maps. However, previous studies have primarily focused on mapping vegetation composition at coarse resolution (250 m) (Ruefenacht et al., 2008) and forest types at moderate resolutions (30 m) (University of Alaska-Anchorage, 2017).

Alaska boreal vegetation cover varies at scales of small patches (Cahoon and Baer, 2022) with each forest type having key differences in response to disturbances such as wildfire, drought, and heat stress (Chapin et al., 2010; Johnstone et al., 2010, 2020; Walker et al., 2015). For example, black spruce forests are particularly vulnerable to dryness and drought which accelerate permafrost thawing and create soil conditions more conducive to deciduous forests (Douglas et al., 2014). Consequently, while previous studies provide valuable insights, their coarse resolutions limit their ability to study resilience and potential shifts in forest types driven by climate change. This limitation is especially critical in ecosystems where distinct forest types exhibit unique growth patterns responses to disturbance and climate change.

A combination of field data and high-resolution (~1m) remote sensing data affords the potential to produce precise and spatially extensive forest type maps for accurately monitoring shifts in this region. Previously, studies have shown the effectiveness of lidar and high-resolution imagery in classifying forest types at fine scales. Airborne lidar and hyperspectral data has been used to map individual tree species in boreal forests of Norway (Dalponte et al., 2014) and Finland (Mäyrä et al., 2021). Alonzo et al. (2018) used high-resolution UAV imagery with forest inventory analysis (FIA) plot data to model forest type at individual crown scale and achieved 85% accuracy for identifying tree species in the Bonanza Creek experimental forest, Alaska. Shoot et al. (2021) conducted a case study by combining FIA field plots with lidar and vegetation indices derived from hyperspectral data to classify forest types and found that the combination of lidar-derived metrics and vegetation indices produced more accurate classification. However, these studies do not address systematic fine scale boreal forest type mapping that can be extended over larger geographical extents and tended to be proofs of concepts studies covering small areas (<8900 ha).

Fig. 1. Landscape of Interior Alaska that is covered with mosaic of conifers and deciduous forest types along with nonforest patches (Cahoon and Baer, 2022).

Our research was motivated by the needs of the United States, Forest Service to monitor changes in the ecosystem across millions of hectares at fine resolution. In the continental US, the Forest Service uses a national Forest Inventory Analysis (FIA) program to monitor changes. Due to the vast and varied expanses of Interior Alaska, the FIA program estimated to take up to 15–20 years to cover the entirety of Interior Alaska. Hence, in Alaska, the FIA program was not implemented until 2014, significantly hindering the rate of progress of climate change impact monitoring. To overcome these challenges, the FIA program collaborated with National Aeronautics and Space Administration (NASA) to leverage high-resolution airborne remote sensing data to support the sparse FIA inventory. This is achieved by developing a novel approach of combining systematic field data sampling with high-resolution imagery acquired by NASA Goddard's Lidar, Hyperspectral and Thermal (G-LiHT) airborne imager in a strip sampling mode covering every FIA plot (Andersen et al., 2024; Cahoon and Baer, 2022). The airborne data acquired by G-LiHT affords an opportunity to upscale forest type information from field plot to larger geographical extents (Cook et al., 2013).

Our research builds upon previous boreal forest cover mapping studies by creating the first systematic framework for classifying forest type at the scale of small patches, across a large landscape (13.5 million hectares) leveraging FIA field, G-LiHT remote sensing, and topographic data. This framework integrates systematically sampled FIA field data and fine resolution remote sensing data (~1m), enabling deep learning models to map forest types in boreal forest zones at unprecedented scale. The high-resolution forest type information across larger areas can be used to improve our ecological understanding of boreal forest dynamics, improve estimates of aboveground biomass and carbon as well as support the FIA program and forest managers.

Other studies have explored utilizing deep learning models with high-resolution RGB (Schiefer et al., 2020) and hyperspectral (Fricker et al., 2019; Mäyrä et al., 2021) data to classify individual tree species in various forest ecosystems. However, the application of deep learning approaches for fine-scale mapping of dominant forest types in Interior Alaska remains relatively unexplored, particularly regarding their robustness across different environmental conditions. These conditions include differences in solar illumination, site quality, and sky condition ranging from clear to overcast, data acquired over a period of one-month during growing season. The availability of lidar, high-resolution hyperspectral data coupled with field data affords a unique opportunity to leverage deep learning models for classifying and mapping of dominant forest type at high-resolution across broader geographical extents. Moreover, the boreal forests in Interior Alaska are significantly influenced by topographical factors such as elevation, slope, aspect, and solar radiation (Calef et al., 2005; Douglas et al., 2014; Ohse et al., 2009). Therefore, we integrated remote sensing with topographic variables to enhance the classification of forest types. Further, we investigated the key topographic factors associated with the forest type to gain deeper insights into their ecological dynamics.

The primary aim of this research was to develop a framework that combines systematically sampled FIA field data and high-resolution remote sensing data that enables the assessment of the performance of state-of-the-art machine learning models for classification and mapping of forest types combining the high-resolution imagery and FIA field plot information in the boreal forest ecosystems of Interior Alaska. To achieve this objective, we compared the performance of a deep learning model, a convolutional neural network with a machine learning model, XGBoost for classifying forest type using G-LiHT remote sensing data. We used predictor variables including vegetation indices (VIs) derived from hyperspectral imagery and digital terrain model (DTM), slope, aspect, canopy height model (CHM), and solar radiation index (SRI) derived from lidar data. Our study included forest type classification at three levels, including 1. forest and nonforest, 2. hardwood and softwood and nonforest and 3. dominant forest types include paper birch, black spruce, white spruce, and nonforest.

Our overall objective was to determine how well forest types could be classified to support monitoring of changes in boreal forests. We did this by addressing the following three research questions.

1. Which machine learning model, convolutional neural network or XGBoost achieves higher accuracy for forest type classification using field and high-resolution remotely sensed data?
2. What are the most influential predictor variables associated with the forest type classification?
3. How do topographic factors slope, aspect, elevation, and solar radiation contribute to the spatial distribution of forest type, and which exhibits the highest significance?

2. Materials

2.1. Study area

We studied the Tanana unit of Interior Alaska, which covers 33.4 million acres of land, of which 20.8 million acers (62%) is covered with boreal forest (Cahoon and Baer, 2022) (Fig. 2). The study area has six dominant tree species: black spruce (*Picea mariana*), white spruce (*Picea glauca*), tamarack (*Larix laricina*), Alaska paper birch (*Betula neoalaskana*), balsam poplar (*Populus balsamifera*), and aspen (*Populus tremuloides*) (Chapin et al., 2006). Among them, black spruce is the dominant forest type covering 52% of forested area, followed by paper birch (19%) and white spruce (16%) (Cahoon and Baer, 2022).

The Tanana unit terrain consists of rolling hills, lowlands, flat bottomlands, uplands, and floodplains (Fig. 1). The elevation ranges from sea level to a nearly 6100 m high mountain. It has a continental climate, characterized by short, warm summers and long, very cold winters (Gallant, 1995). The mean annual air temperature is −3.3 °C with absolute extremes ranging from −51 °C to 38 °C (Jorgenson et al., 2001). The average annual precipitation is about 280 mm, half of which occurs during the summer (Douglas et al., 2014; Shulski and Wendler, 2007). However, most summer precipitation is rapidly lost due to evaporation (Hinzman et al., 2006). Snow accounts for nearly 40–45% of the mean annual precipitation (Douglas et al., 2014; Hinzman et al., 2006).

2.2. Field dataset

The United States Department of Agriculture (USDA) Forest Service, Forest Inventory and Analysis (FIA) program provides information about the forestlands of the United States with continuous monitoring. Individual FIA plots are on a hexagonal grid with one plot in the center and three subplots distributed in an inverted Y pattern (Bechtold and Patterson, 2005; Cahoon and Baer, 2022) (Fig. 3). In 2014, a joint NASA-FIA inventory pilot project was started to leverage high-resolution remote sensing technology to support the FIA inventory in Interior Alaska (Andersen et al., 2015). In Interior Alaska, there is one FIA plot per 97.12 km^2 (30,000 acres), which is sparser than the lower 48 states because of the remoteness of the area (Andersen et al., 2015).

The FIA field data for Tanana unit was collected during the growing season with leaf-on condition over four years (2014, 2016–2018), with most data gathered in 2014 and 2018 to coincide with G-LiHT data. There were total of 1128 FIA plots in the Tanana unit, which total around 4512 subplots (Table 1). The coordinates of each subplot were obtained using GLONASS-enabled Trimble GeoXH 6000 mapping-grade Global Navigation Satellite System (GNSS) receivers with <2 m geolocation error (Andersen et al., 2024; Finley et al., 2024; McGaughey et al., 2017). The field plot data were reprojected to the same coordinated systems as G-LiHT data for proper alignment and a total of 3557 FIA subplots were spatially aligned with G-LiHT data (Table 1).

Each FIA subplot was assigned with forest inventory information such as live or dead, tree height, DBH, species at tree level, and forest type information at the subplot level. The forest type is assigned to each subplot with the dominant stocking of live trees that are not overtopped

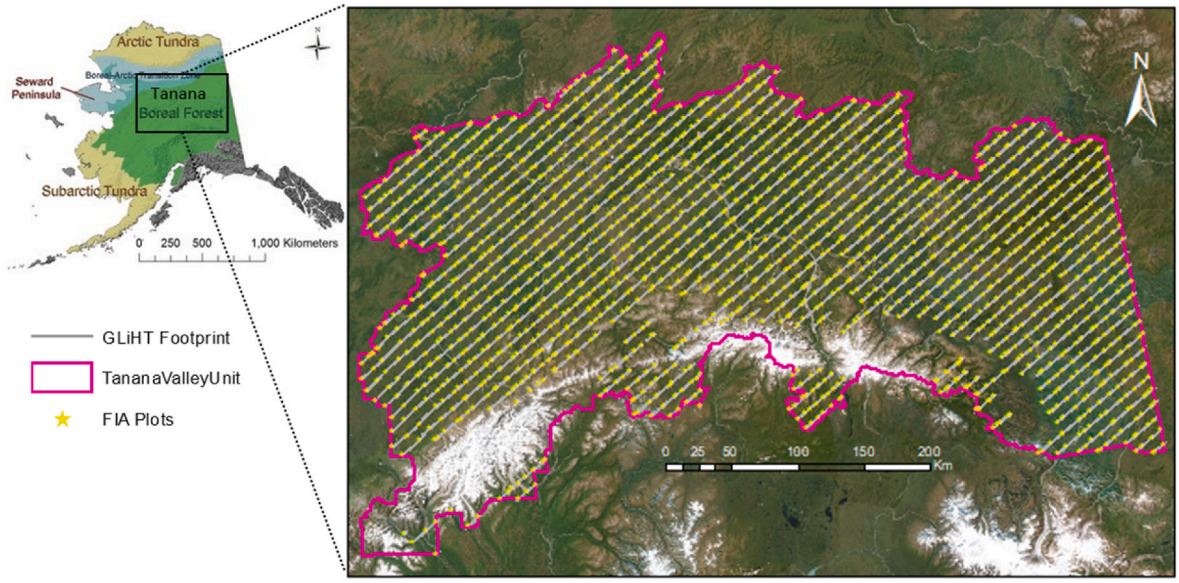

Fig. 2. Tanana unit of Interior Alaska with G-LiHT flight lines and representative FIA plot locations (actual FIA plot locations are obscured due to plot confidentiality requirements).

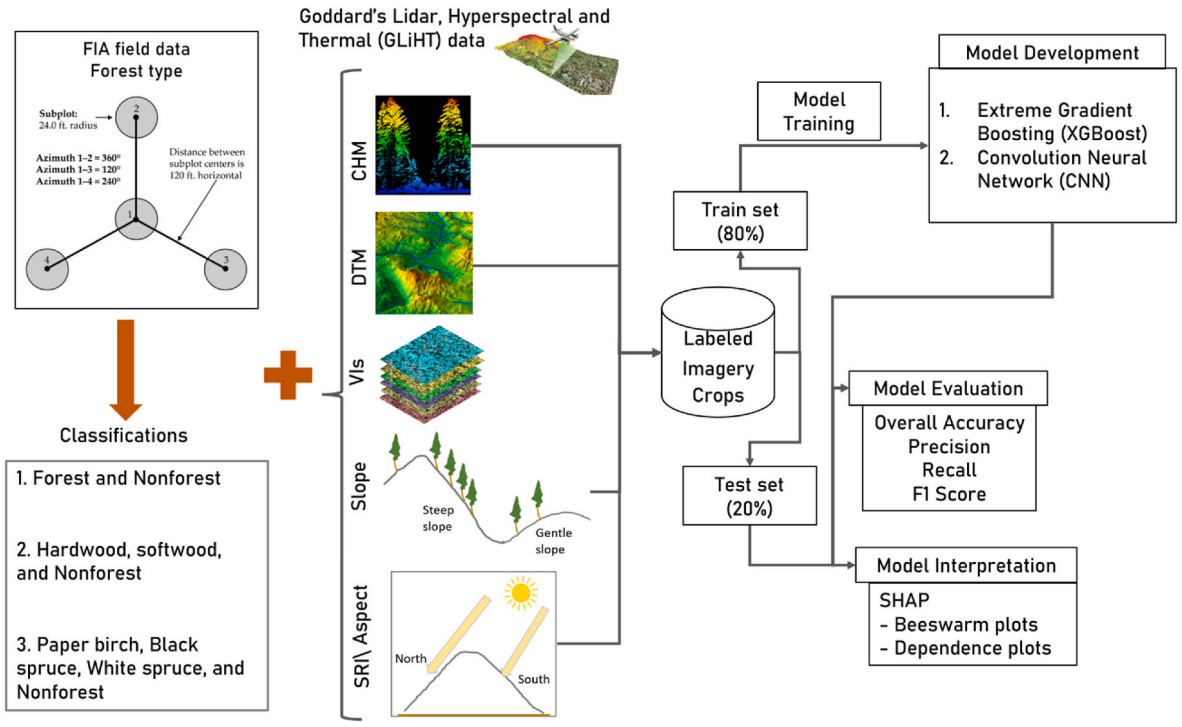

Fig. 3. Schematic diagram of the framework, which highlights the five main steps with data modality and models used for the classification of forest types.

Table 1
Number of total FIA subplots with forest type information used in the analysis. Field subplots represented the field measured plots and G-LiHT subplots represented the FIA subplot that had G-LiHT data available.

Land cover	Forest Type	Tree species	Field subplots	G-LiHT subplots
Nonforest	Nonforest	Nonforest	1752	961
Forest	Softwood	White spruce	523	487
Forest	Softwood	Black spruce	1470	1381
Forest	Softwood	Tamarack	35	35
Forest	Hardwood	Aspen	200	193
Forest	Hardwood	Paper birch	505	473
Forest	Hardwood	Balsam poplar	27	27
		Total	4512	3557

(Bechtold and Patterson, 2005). The accuracy of the forest type at each subplot level was assessed by quality assurance (QA) crew in the field (Bechtold and Patterson, 2005). If there are no trees or the plot does not meet the tree canopy cover threshold (10%) then it was classified as nonforest (USDA, 2018). Based on the FIA, the nonforest category includes shrubs, grasses, herbaceous ground cover, mountain, snow, and water. We used FIA forest type information to create our training dataset.

2.3. NASA Goddard's LiDAR, Hyperspectral and Thermal (GLiHT)

G-LiHT is an airborne imager system that has a suite of complementary instrument (lidar, VNIR imaging spectrometer, and thermal camera) that was designed to map the composition, structure and function of terrestrial ecosystems at fine spatial resolution (1 m) (Andersen et al., 2015; Cook et al., 2013). G-LiHT data was acquired at a nominal altitude of 335 m above ground level (AGL), covering a swath of ~350 m wide spaced 9200 m apart and oriented in a northeast-southwest direction. The G-LiHT swaths were spatially aligned with most of the FIA field plots data using GPS-INS (Inertial Navigation System) system while acquiring the data (Cook et al., 2013). A Riegl VQ-480 airborne laser scanner (Horn, Austria) was used to acquire ~3 to 6 pulses m^{-2}, and up to 8 returns per pulse were recorded for each ~10 cm diameter pulse. Lidar data was used to derive the canopy height model and bare earth model following methods described by Cook et al. (2013). Imaging spectrometer data in the visible through infrared range (420–920 nm) were acquired with a Hyperspec imaging spectrometer (Headwall Photonics, Fitchburg, MA, USA). Downwelling irradiance data was acquired with an Ocean Optics (Dunedin, FL, USA) USB4000-VIS-NIR spectrometer, and used to compute at-sensor reflectance as described by Cook et al. (2013). Reflectance data were used to derive 44 different vegetation indices that represent the unique characteristics of each forest type, and data were downloaded through the G-LiHT data center (https://glihtdata.gsfc.nasa.gov/). Multi-sensor data produced by G-LiHT provides a unique opportunity for scientists to develop synergistic algorithms to monitor forest conditions using data fusion (Cook et al., 2013).

2.3.1. Lidar derived metrics

For this study, we used various metrics derived from lidar (Table 2). We resampled the lidar point clouds and processed the data to produce 1 m resolution digital terrain model (DTM) using a classification of ground returns with a morphological filter (Cook et al., 2013). Similarly, derived a canopy height model (CHM) from the highest lidar return in every 1 m grid cell (Cook et al., 2013). The slope is a measure of the

Table 2
Lidar-derived canopy height and topographical metrics at 1 m spatial resolution were used for analysis. The remote sensing data were collected during summer of 2014 and 2018 with leaf-on conditions.

Lidar and Topographic Metrics	Detail
Canopy Height Model (CHM)	Highest return above ground (m AGL) in each grid cell, interpolated to pixel centroid (meter)
Digital Terrain Model (DTM)	Ground points (elevation) interpolated to pixel centroid (meter)
Slope	Steepness of surface (degrees)
Aspect	Direction in which the slope faces (degrees)
Solar Radiation Index (SRI)	Amount of solar radiation reaching the specific area (unitless)

steepness of a surface and the aspect represents the direction in which the slope faces. The solar radiation aspect index referred to as SRI is a metric used to quantify the orientation of land surfaces in relation to the solar radiation. Its values range from 0 to 1. We used *trasp* function available in SpatialEco package (Evans and Ram, 2021) to calculate SRI in R. The digital terrain model was used to derive the slope, aspect, and SRI. From here onwards, we will refer to DTM as elevation.

2.3.2. Vegetation indices

Vegetation indices that accentuate essential plant characteristics such as greenness, light use efficiency, leaf area, stress, biochemical pigmentation, chlorophyll, anthocyanin, and carotenoid content can be highly valuable for differentiating among forest types, particularly under varying illumination conditions. Vegetation indices has been used to differentiate between vegetation cover (Tian and Fu, 2020). We derived 44 different vegetation indices from hyperspectral data (1 m) to that highlight unique spectral and physiological characteristics as mentioned above (Cook et al., 2013). We have presented 24 commonly used and important indices selected by model (section 4.2) in Table 3 and other vegetation indices used are presented in Table A1.

3. Methods

The main aim of the study was to analyze the performance of 1) the convolutional neural network model and 2) XGBoost models for classifying forest type in Alaska's boreal forest using high-resolution G-LiHT derived metrics. We approached this study using methods that would be appropriate for mapping forest types over large ecosystems (~34 million acres). We therefore used the forest type identified at the subplot level by the FIA field data and combined it with the high-resolution G-LiHT derived data. Our method is divided into the following five sections: 1. Creating labels, 2. Model development, 3. Training models, 4. Model evaluation, and 5. Model Interpretation (Fig. 3).

3.1. Creating labels

We used the forest type information collected at the subplot level by the FIA program to create the labeled dataset. Each subplot was around 7 m (24 ft) in radius. From the in-situ forest type labels, we derived three different taxonomies: 1. forest and nonforest, 2. softwood, hardwood, and nonforest, and 3. dominant forests paper birch, black spruce, white spruce, and nonforest. Thus, in total we had three different set of classification labels for each subplot. For each of these classifications, we trained and evaluated models separately.

For the CNN model, we used the location of the FIA subplot as a center to crop each predictor data using a bounding box of size 16 x 16 m. We padded the 14 × 14 m subplots to by 2 such that the crops were a power of two. For the machine learning model, we derived seven summary statistics, including minimum, first quartile, median, third quartile, standard deviation, mean, and maximum of each predictor variable at the subplot level.

To create robust separation and avoid spatial autocorrelation between training and testing data, which could bias the classification results, we placed all subplots within a plot into either training or testing dataset (Marconi et al., 2022). This ensured that the subplots within the same plot would not end up in both the train and test set. We then randomly created a split of 80:20% data based on the plot location. We used a train set (80%) for training our models and a test set (20%) to evaluate the performance of different models.

3.2. Model development

We developed a custom deep learning (CNN) model using the PyTorch (Paszke et al., 2019) framework in Python for the classification of forest type. Our custom CNN model uses two convolutional layers with 3x3 kernels, followed by two linear layers. Each convolutional layer was followed by batch normalization, rectified linear units (ReLu) as the activation function, and max pooling. Batch normalization reduces training time and overfitting (Ioffe and Szegedy, 2015). These features were then converted to one-dimensional vectors before passing to the linear layer. Finally, we get the probability values of classes using the softmax in the last layer.

We compared the performance of the convolutional neural network

Table 3
The commonly used and important vegetation indices selected by the models that were used in the analysis, where R_xx means reflectance in that band xx, Ravg(xx_xx) and Rsum(xx_xx) mean average and sum of reflectance in those bands (xx_xx) respectively and D_xx means derivative of the spectral reflectance at the specific band xx.

Vegetation Index	Formula	Ref
Metrics describing pigments in vegetation		
Normalize Difference Vegetation Index (NDVI)	(R_800 – R_680)/(R_800 + R_680)	Rouse et al. (1974)
Simple Ratio Pigment Index (SRPI)	R_430/R_680	Peñuelas et al. (1993)
Metrics describing biochemical pigments		
Pigment Specific Normalized Difference (PSND)	(R_800 – R_635)/(R_800 + R_635)	Blackburn (1998)
Structure Insensitive Pigment Index (SIPI)	(R_800 – R_445)/(R_800 – R_680)	Peñuelas et al. (1995)
Plant Senescence Reflectance Index (PSRI)	(R_680 – R_500)/R_750	Merzlyak et al. (1999)
Normalized Phaeophytinization Index (NPQI)	(R_415 – R_435)/(R_415 + R_435)	Barnes et al. (1992)
Photochemical Reflectance Index (PRI)	(R_531 – R_570)/(R_531+ R_570)	Rondeaux et al. (1996)
Renormalized Difference Vegetation Index (RDVI)	(R_800−R_670)/sqrt(R_800 + R_670)	Roujean and Breon (1995)
Metrics describing chlorophyll pigments		
Derivative Chlorophyll Index (DCI)	D_705/D_722	Zarco-Tejada et al. (2002)
Datt Index (Datt1)	R_672 * (R_550 * R_708)	Datt (1998)
Datt Index (Datt2)	(R_850 – R_710)/(R_850 – R_680)	Datt (1999)
Gitelson and Merzlyak (GM1)	R_750/R_550	Gitelson and Merzlyak (1994a)
Gitelson and Merzlyak (GM2)	R_750/R_700	Gitelson and Merzlyak (1994b)
Greenness Index (GI)	R_554/R_677	Smith et al. (1995)
Maccioni (MAC)	(R_780 – R_710)/(R_780 – R_680)	Maccioni et al. (2001)
MERIS Terrestrial Chlorophyll Index (MTCI)	(R_754 – R_709)/(R_709 – R_681)	Dash and Curran (2007)
Normalized Pigment Chlorophyll Index (NPCI)	(R-_680 – R_430)/(R_680 + R-430)	Barnes et al. (1992)
Vogelmann (VOG)	D715/D705	Vogelmann et al. (1993)
Metrics describing anthocyanin and carotenoid content		
Anthocyanin Reflectance Index (ARI1)	(1/R_550) – (1/R_700)	Gitelson et al. (2007)
Anthocyanin Reflectance Index (ARI 2)	R_800 * ((1/R_550) – (1/R_700))	Gitelson et al. (2007)
Red Green Ratio Index (RGRI)	Rsum(600–699)/Rsum (500–599)	Gamon and Surfus (1999)
Carotenoid Reflectance Index (CRI1)	(1/R_510) – (1/R_550)	Gitelson et al. (2007)
Metrics describing leaf area, plant stress & background noise		
Carter Stress (CS2)	R_605/R_760	Carter (1994)
Elvidge and Zhikang (EZ)	R_(625–795)	Elvidge and Chen (1995)
Aoki (AOKI)	R_500/R_850	Akao et al. (1981)
Triangular Vegetation Index (TVI)	0.5*(120*(Ravg760_800 – Ravg530_570) – 200* (Ravg650_680 – Ravg530_570))	Broge and Leblanc (2001)

model with eXtreme gradient boosting (XGBoost), which is known to outperform other machine learning models for supervised classification. Image crop of canopy height, elevation, slope, aspect, and vegetation indices were the input for the CNN model. For the machine learning models, we used the summary statistics of each predictor variable. A total of 343 predictor variables were used for machine learning models to make fair comparison with the deep learning model. For XGBoost model, we used various parameters such as learning rate that controls the step size at which the optimizer updates to weights, gamma which minimize the loss reduction needed to make further partition on a leaf node, maximum depth of trees (max_depth) that controls the maximum depth of the trees in the model, colsample_bytree which controls the fraction of features used for each tree. We used randomized search for optimal hyperparameter selection.

3.3. Model training

For both models, we normalized the data with a z-score using the mean and the standard deviation of the training set. For training deep learning models, we used cross-entropy loss, and an Adam optimizer. We used augmentation techniques to avoid overfitting and improve the generalization of the model on unseen data. During training, we applied different augmentation techniques, such as random horizontal and vertical flipping, and Gaussian blur using the torchvision package (Marcel and Rodriguez, 2010). We trained deep learning models using NVIDIA 2080 Ti GPUs on the Hyak supercomputer of the University of Washington, Seattle, United States. Our models were trained for 100 epochs with a batch size of 16 and a learning rate of 1e-3, of which we saved the last epoch model for the evaluation model.

We implemented our machine learning models using scikit-learn (Pedregosa et al., 2011). Instead of using default hyperparameter values, we tuned hyperparameters during the training phase to improve the classification accuracy. We tuned the hyperparameter separately for each model, using a randomized search. We used a randomized search on five-parameter combinations with five-fold random cross-validation on training data and the best cross-validation score model was fit to the complete training set.

3.4. Model evaluation

To quantify model performance, we used overall accuracy (OA), precision (also known as user's accuracy (UA)), recall (also known as producer's accuracy (PA)), and F1 score. We also used a macro averages of precision, recall, and F1 score, which take the average across classes for multi-class classification to determine the best data modality and model (Mäyrä et al., 2021). These metrics were derived using the following formulas with true positives (TP), true negatives (TN), false positives (FP), false negatives (FN), and total numbers of items (N).

$$OA = \frac{TP + TN}{N} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1 = 2^{*}\frac{Precision^{*}Recall}{Precision + Recall} \tag{4}$$

3.5. Model interpretation and variable importance

Machine learning models are considered a black box because the process leading to model outputs is difficult to understand and interpret. However, there are methods like Shapley Additive exPlanations (SHAP) (Lundberg, 2017; Lundberg et al., 2018), that are model-agnostic and provides detailed and individualized explanations and attributes for individual variables, as opposed to providing only general interpretations across the entire dataset (Li, 2022). The SHAP method is a powerful tool for interpreting models confidently (Lundberg et al., 2018). For the local interpretation of predictors, we used the SHAP method with only the XGBoost model for simplicity and efficiency. We implemented SHAP by using the package called shapviz (Mayer and Stando, 2024) in R (version 4.3.1).

We used a beeswarm plot to interpret individual predictor's contribution to the model output. The SHAP value on the x-axis is the measure of the impact of each feature on the model's prediction. A SHAP value of zero means the feature has no impact. A positive SHAP value indicates that the feature is pushing the prediction towards one class and a negative SHAP value is pushing the prediction towards another class. Each row represents a feature used in the model ordered based on its importance. The x position of the dot shows the impact on the model's output and the color of the dot represents the value of the feature from low to high.

In addition, to understand which topographical variables were more associated with paper birch, black spruce, white spruce, and nonforest, we used a dependence plot with the XGBoost model using SHAP. The dependence plots help to understand the relationship between individual predictors and forest type. In the dependence plots, the y-axis represents SHAP values, indicating the impact of each feature on the model's prediction for a particular forest type. Positive SHAP values indicate that the feature contributes towards classifying the area as that forest type, while negative values indicate prediction away from that forest type.

4. Results

4.1. Classification results

We compared the performance of CNNs with XGBoost model for the classification of forest type using high resolution remote sensing data and FIA subplot for three different forest type classifications. All CNN models outperformed the XGBoost models with the highest macro average F1 score and overall accuracy for all classification levels (Fig. 4). A detailed breakdown analysis of model accuracy for all classification levels is provided in the subsections.

4.1.1. Classification of forest and nonforest

The forest and nonforest classification had the highest accuracy for both models compared to other classifications. The CNN model had an overall accuracy of 93.1% with a macro average F1 score of 0.91. Whereas the overall accuracy of XGBoost was 89.9% with a macro average F1 score of 0.86. The CNN model had the higher F1 score, precision, and recall for both forest and nonforest classes compared to XGBoost (Table 4). Generally, forest was classified with higher accuracies and nonforest was confused with forest over 20% of the time (Figure A1).

4.1.2. Classification of hardwood, softwood, and nonforest

The CNN model had over all accuracy of 82.6% with a macro average F1 score of 0.80. Whereas the overall accuracy of XGBoost was 79.0% with a macro average F1 score of 0.75. The F1 score for all classes was higher with the CNN model (Table 4). The hardwood and nonforest classes were mostly confused with softwood (Figure A2). The performance of the CNN model for all classes was more balanced and better than XGBoost.

4.1.3. Classification of paper birch, black spruce, white spruce and nonforest

Along with nonforest, we selected three dominant forest types including paper birch, black spruce, and white spruce as they comprise over 80% of the forested land in Interior Alaska for classification

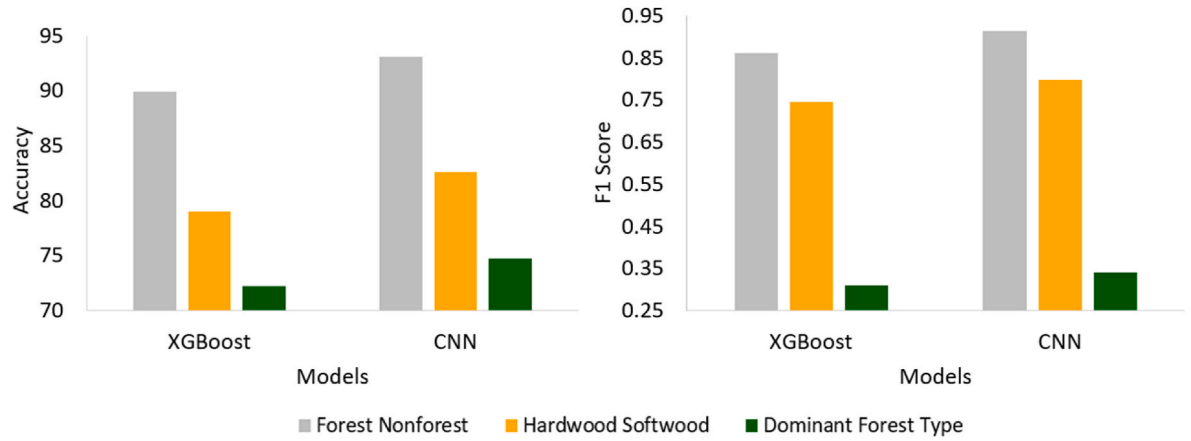

Fig. 4. Summary of the results showing CNN model performs better than XGBoost in terms of the overall accuracy and F1 score for each classification level. Dominant forest type includes paper birch, black spruce, white spruce and nonforest.

Table 4
Accuracy assessment of XGBoost and CNN model for different forest type classifications where OA means soverall accuracy and Macro Avg means macro average.

	Forest and nonforest					
		XGBoost			CNN	
	Precision	Recall	F1 score	Precision	Recall	F1 score
Forest	0.90	0.98	0.93	0.92	0.99	0.95
Nonforest	0.91	0.69	0.79	0.96	0.80	0.87
Macro Avg	0.91	0.83	0.86	0.94	0.89	0.91
OA			89.9 %			93.1 %
	Hardwood, softwood and nonforest					
		XGBoost			CNN	
	Precision	Recall	F1 score	Precision	Recall	F1 score
Hardwood	0.71	0.50	0.59	0.72	0.64	0.68
Nonforest	0.92	0.73	0.81	0.92	0.81	0.86
Softwood	0.77	0.92	0.84	0.81	0.90	0.85
Macro avg	0.80	0.72	0.75	0.82	0.78	0.80
OA			79.0 %			82.6 %
	Dominant forest type: Paper birch, black spruce, white spruce and nonforest					
		XGBoost			CNN	
	Precision	Recall	F1 score	Precision	Recall	F1 score
Paper Birch	0.62	0.44	0.52	0.62	0.67	0.64
Black spruce	0.10	0.88	0.18	0.11	0.83	0.20
Nonforest	0.04	0.78	0.07	0.04	0.83	0.07
White spruce	0.54	0.41	0.47	0.53	0.37	0.44
Macro avg	0.33	0.63	0.31	0.32	0.67	0.34
OA			72.2 %			74.7 %

(Cahoon and Baer, 2022). The CNN model had an overall accuracy of 74.7% with 0.34 macro average F1 score. Whereas XGBoost model had an overall accuracy of 72.2% with 0.31 macro average F1 score (Table 4). Typically, black spruce and nonforest were classified with higher accuracy, and white spruce was the most difficult species to classify correctly and was confused mostly with black spruce and some with birch (Figure A3).

4.2. *Global feature contribution*

To better interpret the optimal importance of different predictors with XGBoost, we used SHAP. In Fig. 5, the x-axis is the mean SHAP values, which represents the average impact on the model output. The canopy height model and elevation were consistently the top predictors across different classifications. Otherwise, feature contributions varied

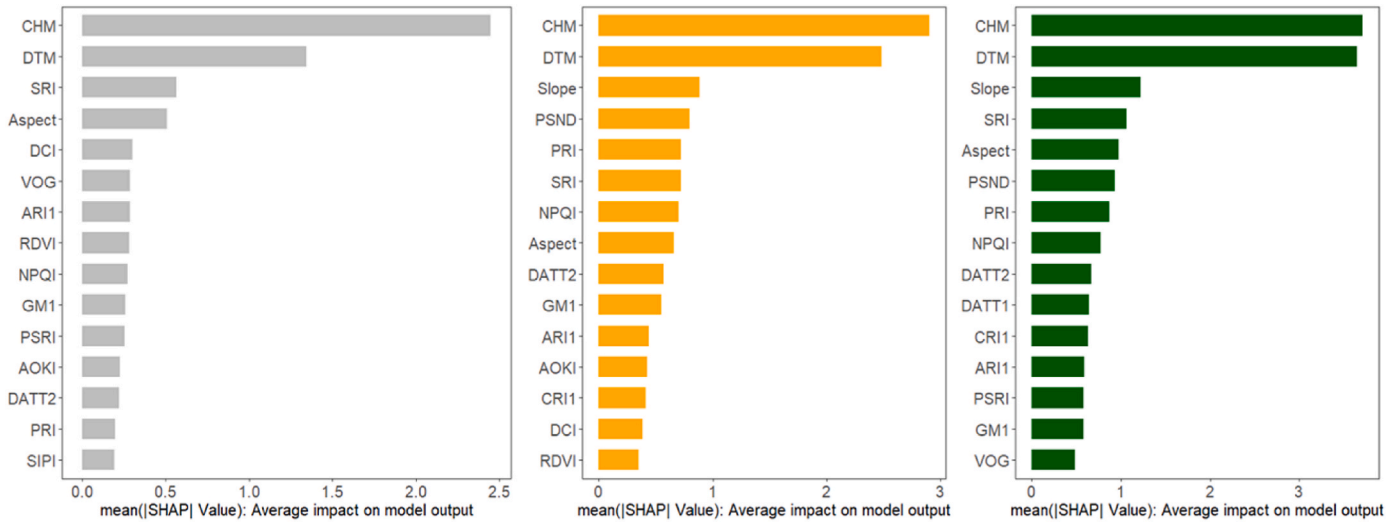

Fig. 5. The top 15 predictors for different level forest type classifications where (left) forest or nonforest, (middle) hardwood, softwood or nonforest, and (right) birch, black spruce, nonforest and white spruce. All acronyms in the above graphs are provided in Tables 2 and 3

across different classifications (Fig. 5). For forest and nonforest classification, solar radiation, aspect, Derivative Chlorophyll Index (DCI), Vogelmann (VOG), Anthocyanin Reflectance Index (ARI1), and Renormalized Difference Vegetation Index (RDVI) had the largest impact. For hardwood, softwood and nonforest classification, slope, Pigment Specific Normalized Difference (PSND), Photochemical Reflectance Index (PRI), Normalized Phaeophytinization Index (NPQI), solar radiation, and aspect had the largest impact. For paper birch, black spruce, white spruce and nonforest classification, slope, solar radiation, aspect, PSND, PRI, and NPQI had the largest impact on the model output. The variation in the canopy height and elevation for different forest types is shown in Fig. 6. The paper birch displayed the largest canopy height on average.

4.3. Local feature contribution

4.3.1. Forest and nonforest

For forest and nonforest classification, the widest spread of SHAP values and those further from zero are the most important features, including canopy height, elevation, and SRI (Fig. 7). The lower canopy height showed a positive SHAP value, indicating that lower canopy height values are important for predicting nonforest class. Conversely, the higher canopy height values spread towards negative SHAP values pushes the prediction towards the forest class. In contrast, higher elevation is spread towards positive SHAP values, indicating higher elevation lean towards nonforest and vice versa. The lower ARI1 values tend to lean towards nonforest. Features whose high and low values are spread in both positive and negative directions of graph suggest that their impact on the model's output depends on the combination with other features.

4.3.2. Hardwood and softwood

We found that elevation was the most important feature for hardwood classification, while canopy height was the most important feature for softwood classification (Fig. 8). For hardwood classification, lower elevation had positive SHAP values indicating that lower elevation pushes the model towards predicting the hardwood. Similarly, medium slope, higher canopy height and PRI values tend to lean towards hardwood. Interestingly, PSND with lower values tend to lean towards hardwood. Further, those features with high and low values spread in both directions suggested that their impact on model depends on the combination of other features.

For softwood, medium to low canopy height values tend to have a higher probability of being softwood. In contrast to hardwood, higher PSND and DATT2 values tend to predict softwood. Lower slope (less steep) and GM1 tends to lean towards softwood. Other presented features showed an interactive effect, which means their effect on the model depends on a combination of other features.

4.3.3. Paper birch, black spruce, and white spruce

We found that canopy height was the most important feature for classifying black spruce and white spruce while elevation was most important for paper birch (Fig. 9). Lower elevations had positive SHAP

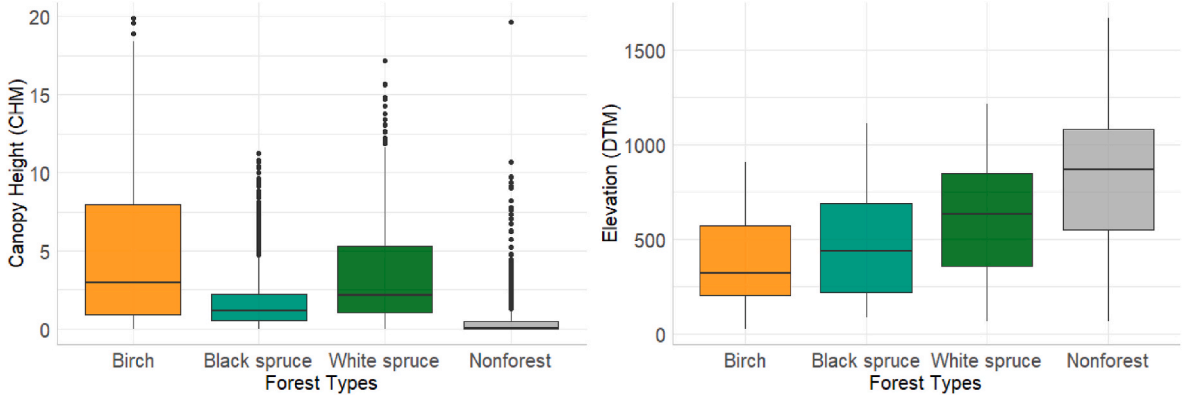

Fig. 6. Demonstrating box plot for the canopy height (CHM) and elevation (DTM) for three dominating forest types and nonforest classification.

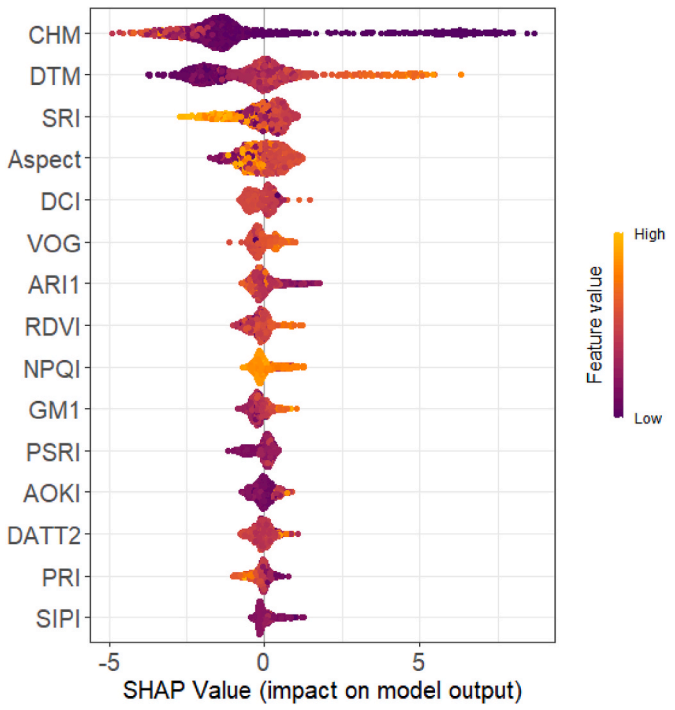

Fig. 7. Comparison of individual feature contributions (top 15) for forest and nonforest classification. Where each dot represents SHAP value of specific data point and its position on x-axis show impact on the model's output. Positive SHAP value indicates features pushing prediction towards nonforest and negative towards forest class. All acronyms in the above graphs are provided in Tables 2 and 3

values, indicating that lower elevation favors both paper birch and black spruce. While higher elevation values lean towards white spruce. Medium slope tends towards predicting paper birch. A higher canopy height favors paper birch, lower canopy height favors black spruce, and medium canopy height favors white spruce.

Among the vegetation indices, NPQI, PRI and DATT2 had interactive effects on paper birch classification. Whereas lower PSND and higher GM2 tend towards predicting paper birch. Similarly, higher PSND tend to predict black spruce and white spruce. In addition, other vegetation indices, such as PRI, DATT1 had an interactive effect for both black spruce and white spruce prediction.

4.4. *Topographical variables associated with forest types*

We used the dependence plots to understand the importance of each topographical variable since it captures non-linear interactions (Fig. 10). The elevation was consistently important for all forests and nonforest classes. For paper birch and black spruce, lower elevations were favorable, while for nonforest higher elevations were more favorable. For white spruce, there was no clear trend, but there were clusters of high and low values at various elevations. This indicated that elevation has a more complex relationship with the presence of white spruce.

In general slope was influential for all forest types but in a different way. For paper birch, the SHAP value tends to decrease as the slope increases, indicating that moderately steep areas were more likely to be associated with paper birch. For black spruce, with the increase in slope, the SHAP value decreases, indicating that less steep areas are more likely associated with black spruce. For white spruce, the SHAP values are concentrated around zero, with a slight trend where both low and high slopes have negative SHAP values. This suggested that moderate slopes may be favorable for white spruce. There is a wide distribution of SHAP values across the slope for nonforest with slight concentration of positive values at higher slope indicating that steeper slope may have higher likelihood of nonforest. Both solar radiation and aspect had less clear direct effects on forests and nonforest classification, suggesting their impact might depend on interaction with other features.

5. Discussion

5.1. *Overall findings*

In this study, we compared the performance of the convolutional neural network with XGBoost model for identifying forest type using the combination of FIA field plot and high-resolution G-LiHT remote sensing

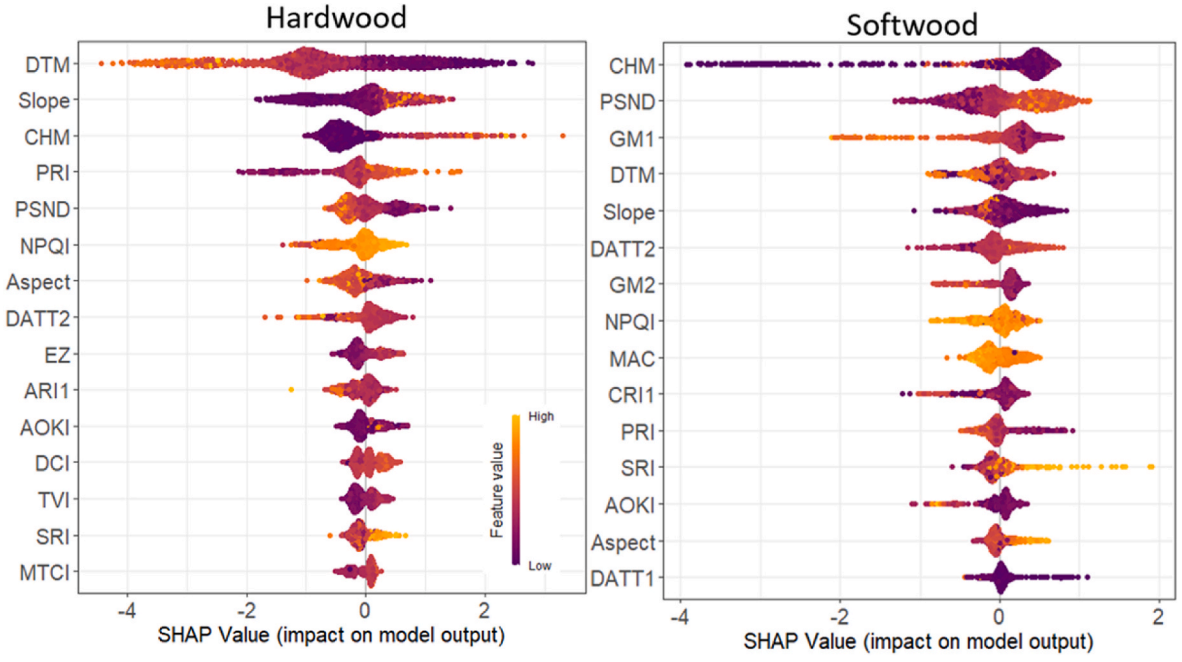

Fig. 8. Comparison of individual feature contributions (top 15) for hardwood, nonforest and softwood classification where class 1: hardwood and class 2: softwood. All acronyms in the above graphs are provided in Tables 2 and 3

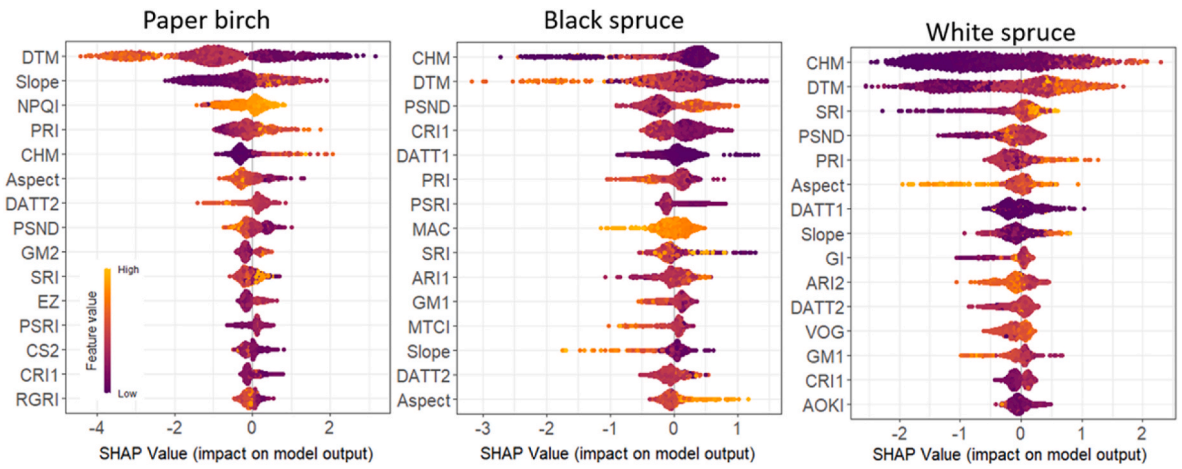

Fig. 9. Comparison of individual feature contributions (top 15) for forests and nonforest classification where class 1: paper birch, class 2: black spruce, and class 3: white spruce. All acronyms in the above graphs are provided in Tables 2 & 3.

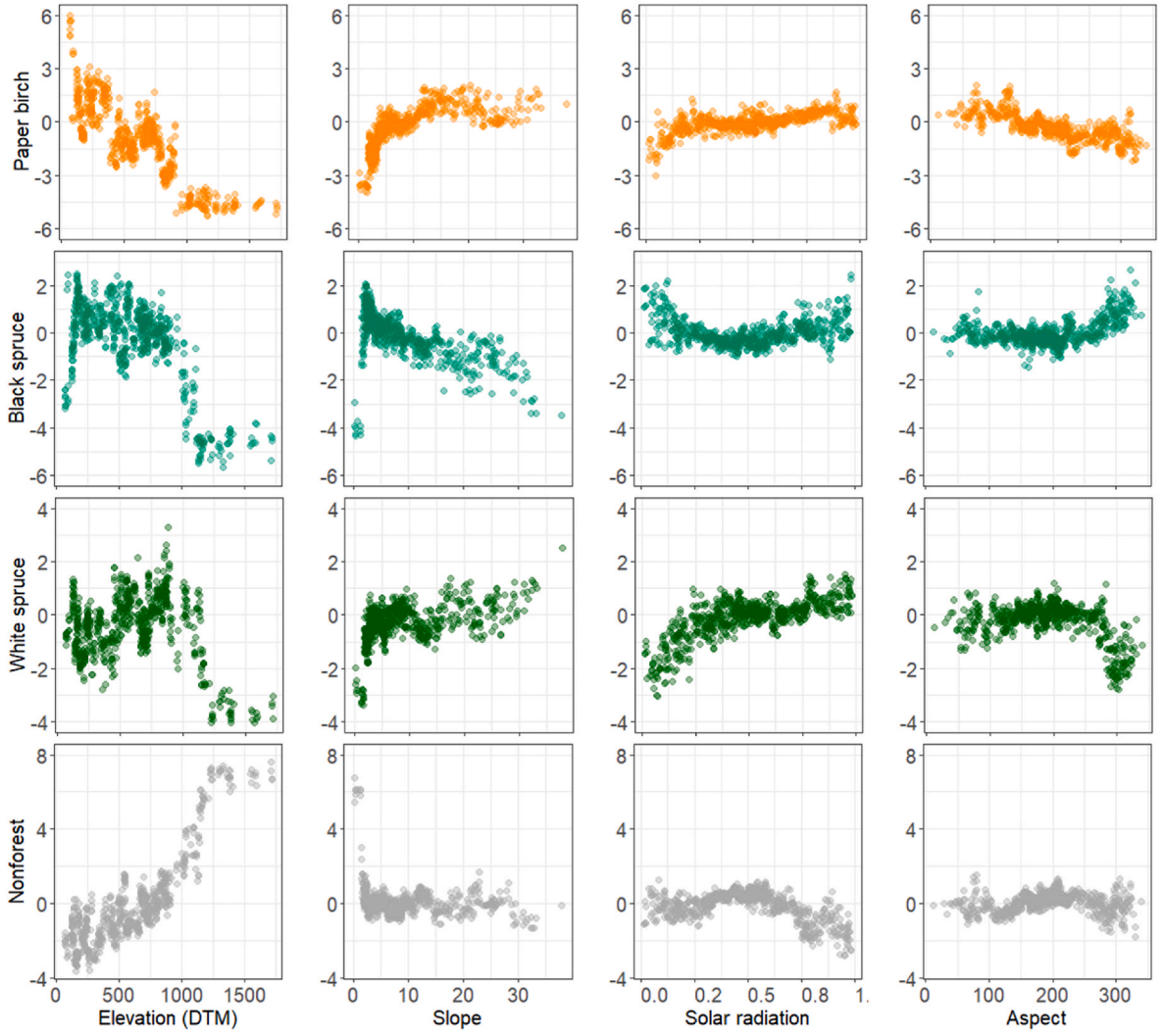

Fig. 10. Dependence plots demonstrate the association of topographic variables with different forest types.

data in the boreal forest of Interior Alaska. For the forest type classification, we conducted classification at three different levels including 1) forest and nonforest, 2) hardwood and softwood and nonforest, and 3) dominant forest type including paper birch, black spruce, white spruce and nonforest. In addition, we used the SHAP method to understand the global and local feature importance.

Our results showed that forest and nonforest classification was conducted with the higher accuracy (96.1%). Whereas we achieved an overall accuracy of 82.6% for hardwood, softwood, and nonforest and 74.7% for four class classification including paper birch, black spruce, white spruce, and nonforest. The CNN model outperformed the XGBoost model in terms of overall accuracy and macro average F1 score for all three different level classifications (Fig. 4). This results aligns with Mäyrä et al. (2021) who found CNN to outperform various commonly used machine learning models for the individual tree species classification. The key advantage of CNN model is that CNNs can automatically learn important features from the input data and are known as best performing model for image classification problems (Krizhevsky et al., 2017; LeCun et al., 2015; Mäyrä et al., 2021). The CNN models performed especially better in separating forest and nonforest, which can provide important information to forest managers. The CNN model also performed better in separating hardwood, softwood and nonforest, which can provide a beneficial method for industry, since commercial interest focuses on softwood. Further, among various features used, we found that canopy height and elevation were the topmost important predictors across all classifications (Fig. 5).

We also found that the commonly used Normalized Difference Vegetation Index (NDVI) was not useful for the boreal forest type classification. Instead, vegetation indices such as Anthocyanin Reflectance Index (ARI) to be useful for forest nonforest classification and Pigment Specific Normalized Difference (PSND), Photochemical Reflectance Index (PRI), and Gitelson and Merzlyak (GM1) to be most useful for differentiating between hardwood and softwood. These vegetation indices capture reflectance mostly in green, red, red-edge, and near-infrared bands and are particularly valuable for identifying and monitoring plant physiological traits, stress (Tian and Fu, 2020; Viinikka et al., 2020). This results aligns with Viinikka et al. (2020) who found vegetation indices other than NDVI such as PSND and PRI to be important for identifying boreal forest species. These vegetation indices are important for improving the classification accuracy and ecological assessments in the boreal forest ecosystem.

The framework developed in this study enables the generation of accurate and high-resolution forest type maps across unprecedented scales. These forest type maps are useful in a variety of contexts including estimation of aboveground biomass and carbon, supporting forest management plans, managing timber production, and harvesting (Triviño et al., 2015) and aid information to rural communities that depend on nearby forest resources (Shoot et al., 2021). The high-resolution forest type map can further our understanding of fine scale ecological processes and provide a strong basis for ecological studies.

5.2. Most important features

Our results showed that canopy height and elevation were the most important variable across all classifications. Lower canopy height, higher elevation, lower solar radiation was related to nonforest this could be due to nonforest includes grasses and shrubs whose canopy heights are generally lower than tree and are found above tree line at higher elevations with low temperatures.

Similarly, our study showed that lower elevation, medium slope, and higher canopies were associated with hardwood. The hardwood stands include aspen, paper birch and balsam poplar that are mostly found in less steep slope (Chapin et al., 2006). Further, we found medium to low canopy height was more associated with softwood stands. The softwood stands include black spruce, white spruce, and tamarack.

Further, vegetation indices that measure important plant features and characteristics like greenness, leaf pigmentation, and light use efficiency, proved highly valuable in differentiating among forest types in Interior Alaska. We will discuss specific vegetation indices later in our discussion.

5.3. Ecological interpretation

5.3.1. Potential sources of misclassification

In this section, we present potential ecological reasons to explain misclassification observed among forest types. Our classification was based identifying the dominant forest type within in subplots, each covering an area with 24 ft radius (around 7 m). Each subplot may have a mix of forest types, with one dominant species assigned to it. Some of the nonforest subplots were confused with the forest class which could be due to nonforest class also includes shrubs, grasslands and the areas that do not pass FIA's the minimum canopy threshold of 10% (USDA, 2018).

Among hardwood, softwood, and nonforest classification, softwood was classified most accurately (90%). While hardwood was confused with softwood class, this could be because hardwood forest stands such as aspen have a mixture of white spruce (Chapin et al., 2006). Similarly, birch dominated stands could have sporadically spread of white spruce and black spruce (Chapin et al., 2006; Douglas et al., 2014).

Similarly, among the three dominant forest type along with nonforest classification, black spruce and nonforest had the higher accuracy. The birch class was mostly misclassified as black spruce, this could be because birch stands could have scattered black spruce that replace the birch in later succession (Chapin et al., 2006). While white spruce was misclassified as both birch and black spruce. This could be because mixed white spruce stands could have birch (Douglas et al., 2014). In addition, both black spruce and late-successional white spruce are mostly available in cooler and moist areas (Chapin et al., 2006).

5.3.2. Importance of remote sensing variables

Among various remote sensing variables canopy height was the most important variable across all classifications. In addition, key plant traits allowed discrimination of forest types using vegetation indices. We found that Anthocyanin Reflectance Index (ARI1) with lower values, were more related to nonforest. This is because ARI1 is used for describing anthocyanin content in plants, which are higher in deciduous trees (Gould, 2004).

Similarly, our study showed that higher Photochemical Reflectance Index (PRI) and lower Pigment Specific Normalized Difference (PSND) were associated with hardwood. The PRI index is used as the indicator of photosynthetic radiation found in both hardwood and softwood forests (Gamon et al., 1997). However, softwood forest has low photosynthetic rate relative to the hardwood forest because of which higher PRI is more related to hardwood (Gamon et al., 1997). Further, the PSND index is highly correlated with chlorophylls (Blackburn, 1999) and hardwood forests during fall change color and have lower chlorophylls. That is why lower PSND values were more associated with hardwood forest.

Further, we found higher PSND and DATT2 were associated with softwood. The PSND and DATT2 is closed related to the higher

chlorophyll content in the leaf (Blackburn, 1999; Datt, 1999). Consequently, higher values of these indices are associated with softwood forests, where foliage maintains its green color without significant changes like hardwood forest. While lower Gitelson and Merzlyak (GM1) index is related to softwood because this index is related to chlorophyll content of yellow-green leaf (Gitelson et al., 1996a,b).

These findings align with previous study conducted in European boreal forest that has found that PSND was important for identifying the coniferous trees and PRI was important for identifying aspen and birch (Viinikka et al., 2020).

5.3.3. *Importance of topographic and solar radiation variables*

Among various topographical variables, including elevation, slope, aspect, and solar radiation, we found elevation was consistently important for all forest types and nonforest.

We found that lower elevation, moderately steep areas with east- and south-facing slopes had a higher probability of associating with paper birch. This aligns with previous studies that have found birch stands to be more common on upland and south-facing slopes (Douglas et al., 2014) while others have found on east- and west-facing slopes that are cool and moist (Chapin et al., 2006). Since birch can be found either in south-facing or west-, and east-facing, which has very different solar radiation therefore we were not able to find a clear relationship between solar radiation and paper birch.

We found that lower elevations, less steep slopes, north facing areas with medium solar radiation have a higher likelihood to be associated with black spruce. Among all topographic variables, elevation was strongly associated with black spruce that aligns with the results of previous studies (Calef et al., 2005). In general, black spruce stands are found to dominate the north-facing upland slopes with permafrost because they prefer cool and moist places (Chapin et al., 2006; Douglas et al., 2014).

Similarly, we found that moderate slope with either very high or low solar radiation is likely to be associated with white spruce. The lower elevations were associated with white spruce but not as strong as black spruce. Previous studies have shown that mixed white spruce stands can be found in uplands and lowlands and north-facing slopes (Douglas et al., 2014) and late successional white spruce are found in lower elevations on south-, east-, and west-facing slopes (Chapin et al., 2006). Since white spruce can be found in slope with different directions due to which we were not able to find a clear relationship between aspect and white spruce. White spruces are generally found in cooler and moist areas (Chapin et al., 2006). which aligns with our finding of white spruce being associated with moderate solar radiation. And higher elevation and lower solar radiation were associated with nonforest.

6. Conclusion

In this study, we presented a framework for forest type classification combining FIA field plots and high-resolution remote sensing data in the boreal forest of Interior Alaska. We conducted forest type classification at three different levels, including 1. Forest and nonforest, 2. Hardwood, softwood, and nonforest and 3. Three dominant forest types, including paper birch, black spruce, white spruce, and nonforest. In addition, we also identified the importance of topographic and remote sensing variables for each forest type classification. The study showed that forest type classification can be conducted with high accuracy using deep learning methods and high-resolution remote sensing data. The study also showed that elevation was the most important topographic factor associated with these forest types. In addition, we found that canopy height and vegetation indices including Photochemical Reflectance Index (PRI), Pigment Specific Normalized Difference (PSND), and Gitelson and Merzlyak (GM1) were important for differentiating between hardwood and softwood while Anthocyanin Reflectance Index (ARI1) was important for differentiating between forest and nonforest.

This is the first study to use a systematic sampling of field data with high-resolution airborne remote sensing data collected over large areas under variable sky condition to classify forest types in Interior Alaska, with particular emphasis in evaluating the performance of deep learning and machine learning models in the process. This research highlights the importance and uses of high-resolution remote sensing to complement field especially in cases where field locations are inaccessible by humans. In future, we aim to use these models to produce forest type classifications over the entire G-LiHT footprint. The development and evaluation of models using field and high-resolution remote sensing for mapping forest type will be crucial for long-term monitoring for forest ecosystem to detect changes and estimating above ground biomass, carbon, and supporting forest management plans.

Funding

Funding for this project was provided by the United States Department of Agriculture (USDA), Forest Service Pacific Northwest Research Station and the National Aeronautics and Space Administration (NASA). This research was part of the project “NASA-USFS Partnership to Advance Operational Forest Carbon Monitoring in Interior Alaska” funded by the NASA Carbon Monitoring System (CMS). Funding was provided through USDA Forest Service, Alaska Grant 18-JV-11261989-064, 19-JV-11261989-111, 20-JV-11261989-085, and 21-JV-11261989-068.

CRediT authorship contribution statement

Pratima Khatri-Chhetri: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hans-Erik Andersen:** Writing – review & editing, Methodology, Funding acquisition, Data curation, Conceptualization. **Bruce Cook:** Writing – review & editing, Data curation, Conceptualization. **Sean M. Hendryx:** Writing – review & editing, Investigation. **Liz van Wagtendonk:** Writing – review & editing, Investigation, Data curation. **Van R. Kane:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare no conflict of interest or financial support that may have influenced the results or interpretation of this study.

Acknowledgments

The authors would like to thank Interior Alaska FIA field crews for high-quality data collection, Maureen Kennedy for providing feedback on the manuscript and the University of Washington Research Computing Resource Team for providing access to the supercomputing. In addition, the authors would like to thank editors and reviewers for their thorough comments and suggestions, which greatly improved the final manuscript.

Appendix A

Table A1
Additional vegetation indices that were used in the analysis.

Vegetation Index	Formula	Ref
Carter Stress (CS1)	R_695/R_760	Carter (1994)
Carotenoid Reflectance Index (CRI2)	(1/R_510) – (1/R_700)	Gitelson et al. (2007)
Curvature Index (CI)	(R_683 * 2)/(R_675 * R_691)	Zarco-Tejada et al. (2002)
Difference Vegetation Index (DVI)	R_800 – R_680	Tucker (1979)
Filella and Penuelas (FP)	R_(680–780)	Filella and Penuelas (1994)
Modified Chlorophyll Absorption Ration Index Improved (MCARI2)	1.5*((2.5*(R_800 – R_670)) – (1.3*(R_800 –R_550)))/(sqrt(((((2*R_800 + 1)*(2*R_800 + 1)))) – (6*R_800 –(5*(sqrt(R_670))) – 0.5)))	Haboudane (2004)
Modified Red Edge Normalized Difference Vegetation Index (MRENDVI)	(R_750-R_705)/(R_750+R_705+(2*R_445))	Datt (1999)
Modified Red Edge Simple Ratio (MRESR)	(R_750 – R_445)/(R_705 – R_445)	Sims and Gamon (2002)
Modified Simple Ration (MSR)	((R_800/(R_678 – 1))/sqrt((R_800/R_760)+1))	Chen (1996)
Modified Triangular Vegetation Index (MTVI)	1.2*((1.2*(R_800 – R_550)) – (2.5*(R_670 –R_550)))	Haboudane (2004)
Modified Triangular Vegetation Index Improved (MTVI2)	(1.5*((1.2*(R_800 – R_550)) – (2.5*(R_670-R_550))))/(sqrt(((2*R_800 + 1)*(2*R_800 + 1))-(6*R_800 – (5*(sqrt(R_670)))) – 0.5))	Haboudane (2004)
Modified Chlorophyll Absorption Ratio Index (MCARI)	1.2*((2.5*(R_800 – R_670)) – (1.3*(R_800 – R_550)))	Haboudane (2004)
Normalize Difference Vegetation Index (NDVI)	(R_800 – R_680)/(R_800+R_680)	Rouse et al. (1974)
Optimized Soil Adjusted Vegetation Index (OSAVI)	(R_800 –R_680)/(R_800+R_680 + 0.16)	Peñuelas et al. (1994)
Pigment Specific Simple Ration (PSSR)	R_600/R_680	Blackburn (1998)
Ratio Vegetation Index (RVI)	R_800/R_680	Royle and Lathrop (2002)
Red Edge Inflection Point (REIP)	R_750/R_700	Vogelmann et al. (1993)
Red Edge Normalized Difference Vegetation Index (RENDVI)	(R_750 – R_705)/(R_750 + R_705)	Gitelson et al. (1996)
Vogelmann Red Edge Index (VOGREI)	R_740/R_720	Vogelmann et al. (1993)

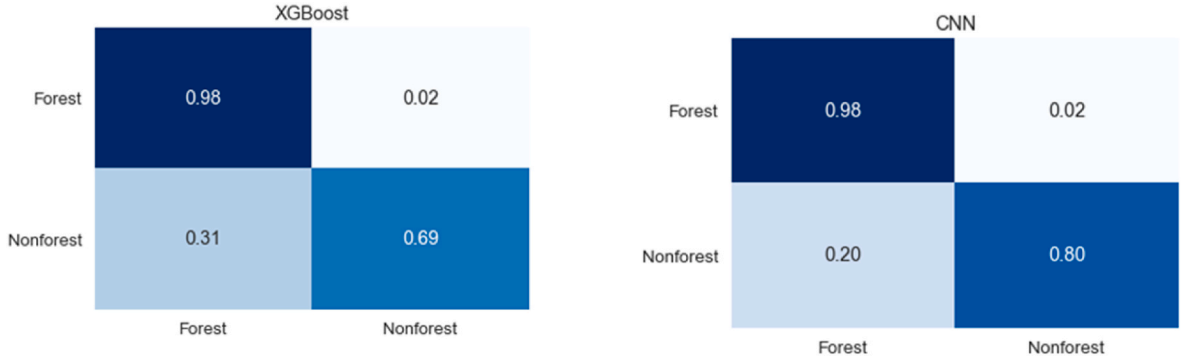

Fig. A1. Confusion matrix of forest and nonforest classification using XGBoost and CNN models. Rows indicate correct labels and columns indicate predicted labels.

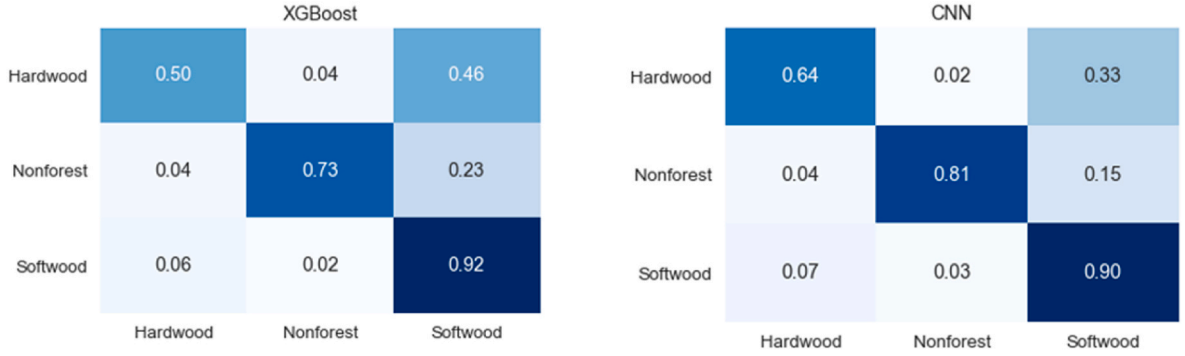

Fig. A2. Confusion matrix of hardwood, nonforest and softwood classification using XGBoost and CNN models. Rows indicate correct labels and columns indicate predicted labels.

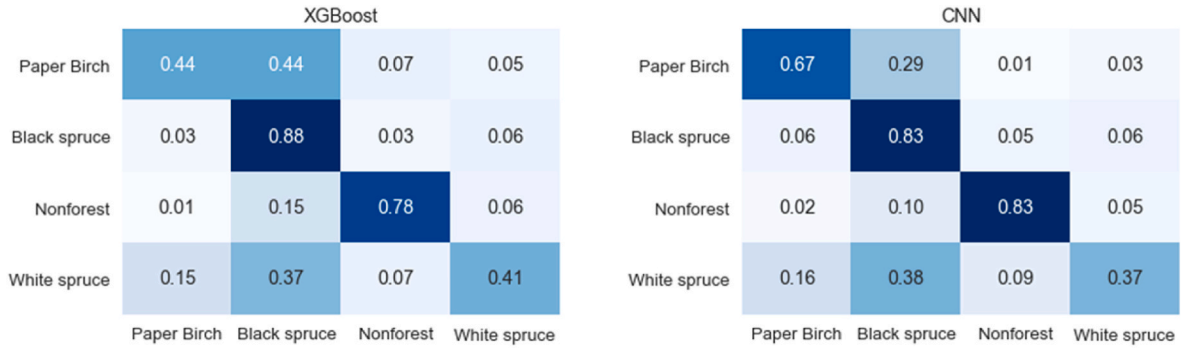

Fig. A3. Confusion matrix of paper birch, black spruce, nonforest and white spruce classification using XGBoost and CNN models. Rows indicate correct labels and columns indicate predicted labels.

Data availability

The source code used in our analysis is available at: https://github.com/Upanpra/forest-type-detection. FIA plots approximate locations are publicly available; however, the actual FIA plot coordinates used for analysis are confidential. Therefore, authors do not have permission to share data.

References

Akao, M., Aoki, H., Kato, K., 1981. Mechanical properties of sintered hydroxyapatite for prosthetic applications. J. Mater. Sci. 16 (3), 809–812. https://doi.org/10.1007/BF02402799.

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.

Andersen, H., Babcock, C., Pattison, R., Cook, B., Morton, D., Finley, A., 2015. The 2014 tanana inventory pilot: a USFS-NASA partnership to leverage advanced remote sensing technologies for forest inventory. In: Stanton, Sharon M., Christensen, Glenn A. (Eds.), Comps. 2015. Pushing Boundaries: New Directions in Inventory Techniques and Applications: Forest Inventory and Analysis (FIA) Symposium 2015, pp. 40–41. *2015 December 8–10; Portland, Oregon. Gen. Tech. Rep. PNW-GTR-931, 931*.

Andersen, H.-E., Ståhl, G., Cook, B.D., Morton, D.C., Finley, A.O., 2024. Model-assisted estimation of domain totals, areas, and densities in two-stage sample survey designs. Can. J. For. Res. https://doi.org/10.1139/cjfr-2024-0039 cjfr-2024-0039.

Barber, V.A., Juday, G.P., Finney, B.P., 2000. Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. Nature 405 (6787), 668–673. https://doi.org/10.1038/35015049.

Barnes, J.D., Balaguer, L., Manrique, E., Elvira, S., Davison, A.W., 1992. A reappraisal of the use of DMSO for the extraction and determination of chlorophylls a and b in lichens and higher plants. Environ. Exp. Bot. 32 (2), 85–100. https://doi.org/10.1016/0098-8472(92)90034-Y.

Bechtold, W.A., Patterson, P.L., 2005. The Enhanced Forest Inventory and Analysis Program—National Sampling Design and Estimation Procedures, vol. 80. USDA Forest Service, Southern Reserach Station.

Beck, P.S.A., Horning, N., Goetz, S.J., Loranty, M.M., Tape, K.D., 2011. Shrub cover on the north slope of Alaska: a circa 2000 baseline map. Arctic Antarct. Alpine Res. 43 (3), 355–363. https://doi.org/10.1657/1938-4246-43.3.355.

Blackburn, G.A., 1998. Spectral indices for estimating photosynthetic pigment concentrations: a test using senescent tree leaves. Int. J. Rem. Sens. 19 (4), 657–675. https://doi.org/10.1080/014311698215919.

Blackburn, G.A., 1999. Relationships between spectral reflectance and pigment concentrations in stacks of deciduous broadleaves. Rem. Sens. Environ. 70 (2), 224–237. https://doi.org/10.1016/S0034-4257(99)00048-6.

Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Rem. Sens. Environ. 76 (2), 156–172. https://doi.org/10.1016/S0034-4257(00)00197-8.

Cahoon, S.M.P., Baer, K.C., 2022. *Forest Resources of the Tanana Unit, Alaska: 2018* (PNW-GTR-1005; P. PNW-GTR-1005). U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. https://doi.org/10.2737/PNW-GTR-1005.

Calef, M.P., David McGuire, A., Epstein, H.E., Scott Rupp, T., Shugart, H.H., 2005. Analysis of vegetation distribution in Interior Alaska and sensitivity to climate change using a logistic regression approach. J. Biogeogr. 32 (5), 863–878. https://doi.org/10.1111/j.1365-2699.2004.01185.x.

Carter, G.A., 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. Int. J. Rem. Sens. 15 (3), 697–703. https://doi.org/10.1080/01431169408954109.

Chapin, F.S., McGuire, A.D., Ruess, R.W., Hollingsworth, T.N., Mack, M.C., Johnstone, J. F., Kasischke, E.S., Euskirchen, E.S., Jones, J.B., Jorgenson, M.T., Kielland, K., Kofinas, G.P., Turetsky, M.R., Yarie, J., Lloyd, A.H., Taylor, D.L., 2010. Resilience of Alaska's boreal forest to climatic changeThis article is one of a selection of papers from the Dynamics of Change in Alaska's Boreal Forests: resilience and Vulnerability in Response to Climate Warming. Can. J. For. Res. 40 (7), 1360–1370. https://doi.org/10.1139/X10-074.

Chapin, F.S., Oswood, M.W., Van Cleve, K., Viereck, L.A., Verbyla, D.L., 2006. Alaska's Changing Boreal Forest. Oxford University Press.

Chen, J.M., 1996. Evaluation of vegetation indices and a modified simple ratio for boreal applications. Can. J. Rem. Sens. 22 (3), 229–242. https://doi.org/10.1080/07038992.1996.10855178.

Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., Montesano, P., 2013. NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager. Rem. Sens. 5 (8), 4045–4066. https://doi.org/10.3390/rs5084045.

Dalponte, M., Ørka, H.O., Ene, L.T., Gobakken, T., Næsset, E., 2014. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. Rem. Sens. Environ. 140, 306–317. https://doi.org/10.1016/J.RSE.2013.09.006.

Dash, J., Curran, P.J., 2007. Evaluation of the MERIS terrestrial chlorophyll index (MTCI). Adv. Space Res. 39 (1), 100–104. https://doi.org/10.1016/j.asr.2006.02.034.

Datt, B., 1998. Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+b, and total carotenoid content in Eucalyptus leaves. Rem. Sens. Environ. 66 (2), 111–121. https://doi.org/10.1016/S0034-4257(98)00046-7.

Datt, B., 1999. Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. Int. J. Rem. Sens. 20 (14), 2741–2759. https://doi.org/10.1080/014311699211778.

Douglas, T.A., Jones, M.C., Hiemstra, C.A., Arnold, J.R., 2014. Sources and sinks of carbon in boreal ecosystems of interior Alaska: a review. Elementa: Science of the Anthropocene 2, 000032. https://doi.org/10.12952/journal.elementa.000032.

Ekman, E., Triviño, M., Blattert, C., Mazziotta, A., Potterf, M., Eyvindson, K., 2024. Disentangling the effects of management and climate change on habitat suitability for saproxylic species in boreal forests. J. For. Res. 35 (1), 34. https://doi.org/10.1007/s11676-023-01678-3.

Elvidge, C.D., Chen, Z., 1995. Comparison of broad-band and narrow-band red and near-infrared vegetation indices. Rem. Sens. Environ. 54 (1), 38–48. https://doi.org/10.1016/0034-4257(95)00132-K.

Euskirchen, E.S., McGuire, A.D., Chapin, F.S., Rupp, T.S., 2010. The changing effects of Alaska's boreal forests on the climate systemThis article is one of a selection of papers from the Dynamics of Change in Alaska's Boreal Forests: resilience and Vulnerability in Response to Climate Warming. Can. J. For. Res. 40 (7), 1336–1346. https://doi.org/10.1139/X09-209.

Evans, J.S., Ram, K., 2021. Package 'spatialEco'. *R CRAN Project*.

Filella, I., Penuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. Int. J. Rem. Sens. 15 (7), 1459–1470. https://doi.org/10.1080/01431169408954177.

Finley, A.O., Andersen, H.-E., Babcock, C., Cook, B.D., Morton, D.C., Banerjee, S., 2024. Models to support forest inventory and small area estimation using sparsely sampled LiDAR: a case study involving G-LiHT LiDAR in tanana, Alaska. J. Agric. Biol. Environ. Stat. 29 (4), 695–722. https://doi.org/10.1007/s13253-024-00611-3.

Fricker, G.A., Ventura, J.D., Wolf, J.A., North, M.P., Davis, F.W., Franklin, J., 2019. A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery. Rem. Sens. 11 (19), 2326. https://doi.org/10.3390/rs11192326.

Gallant, A.L., 1995. Ecoregions of Alaska. US Government Printing Office.

Gamon, J.A., Serrano, L., Surfus, J.S., 1997. The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. Oecologia 112 (4), 492–501. https://doi.org/10.1007/s004420050337.

Gamon, J.A., Surfus, J.S., 1999. Assessing leaf pigment content and activity with a reflectometer. New Phytol. 143 (1), 105–117. https://doi.org/10.1046/j.1469-8137.1999.00424.x.

Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996a. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Rem. Sens. Environ. 58 (3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7.

Gitelson, A.A., Merzlyak, M.N., Chivkunova, O.B., 2007. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. Photochem. Photobiol. 74 (1), 38–45. https://doi.org/10.1562/0031-8655(2001)0740038OPANEO2.0.CO2.

Gitelson, A.A., Merzlyak, M.N., Lichtenthaler, H.K., 1996b. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. J. Plant Physiol. 148 (3–4), 501–508. https://doi.org/10.1016/S0176-1617(96)80285-9.

Gitelson, A., Merzlyak, M.N., 1994a. Quantitative estimation of chlorophyll-a using reflectance spectra: experiments with autumn chestnut and maple leaves. J. Photochem. Photobiol. B Biol. 22 (3), 247–252. https://doi.org/10.1016/1011-1344(93)06963-4.

Gitelson, A., Merzlyak, M.N., 1994b. Spectral reflectance changes associated with autumn senescence of aesculus hippocastanum L. And acer platanoides L. Leaves. Spectral features and relation to chlorophyll estimation. J. Plant Physiol. 143 (3), 286–292. https://doi.org/10.1016/S0176-1617(11)81633-0.

Gould, K.S., 2004. Nature's Swiss army knife: the diverse protective roles of anthocyanins in leaves. J. Biomed. Biotechnol. 2004 (5), 314–320. https://doi.org/10.1155/S1110724304406147.

Haboudane, D., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. Rem. Sens. Environ. 90 (3), 337–352. https://doi.org/10.1016/j.rse.2003.12.013.

Hinzman, L.D., Viereck, L.A., Adams, P.C., Romanovsky, V.E., Yoshikawa, K., 2006. Climate and permafrost dynamics of the Alaskan boreal forest. In: Alaska's Changing Boreal Forest, pp. 39–61.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift arXiv preprint arXiv:1502.03167.

Johnstone, J.F., Celis, G., Chapin, F.S., Hollingsworth, T.N., Jean, M., Mack, M.C., 2020. Factors shaping alternate successional trajectories in burned black spruce forests of Alaska. Ecosphere 11 (5), e03129. https://doi.org/10.1002/ecs2.3129.

Johnstone, J.F., Hollingsworth, T.N., Chapin, F.S., Mack, M.C., 2010. Changes in fire regime break the legacy lock on successional trajectories in Alaskan boreal forest. Global Change Biol. 16 (4), 1281–1295. https://doi.org/10.1111/j.1365-2486.2009.02051.x.

Jorgenson, M., Racine, C.H., Walters, J.C., Osterkamp, T.E., 2001. Permafrost degradation and ecological changes associated with a warming climate in central Alaska. Clim. Change 48 (4), 551–579.

Kasischke, E.S., Verbyla, D.L., Rupp, T.S., McGuire, A.D., Murphy, K.A., Jandt, R., Barnes, J.L., Hoy, E.E., Duffy, P.A., Calef, M., Turetsky, M.R., 2010. Alaska's changing fire regime — implications for the vulnerability of its boreal forestsThis article is one of a selection of papers from the Dynamics of Change in Alaska's Boreal Forests: resilience and Vulnerability in Response to Climate Warming. Can. J. For. Res. 40 (7), 1313–1324. https://doi.org/10.1139/X10-098.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.

Li, Z., 2022. Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. Computers, Environment and Urban Systems 96, 101845.

Lundberg, S., 2017. A unified approach to interpreting model predictions arXiv preprint arXiv:1705.07874.

Lundberg, S.M., Erion, G.G., Lee, S.-I., 2018. Consistent individualized feature attribution for tree ensembles. https://doi.org/10.48550/ARXIV.1802.03888.

Maccioni, A., Agati, G., Mazzinghi, P., 2001. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. J. Photochem. Photobiol. B Biol. 61 (1–2), 52–61. https://doi.org/10.1016/S1011-1344(01)00145-2.

Marcel, S., Rodriguez, Y., 2010. Torchvision the machine-vision package of torch. Proceedings of the 18th ACM International Conference on Multimedia, pp. 1485–1488. https://doi.org/10.1145/1873951.1874254.

Marconi, S., Weinstein, B.G., Zou, S., Bohlman, S.A., Zare, A., Singh, A., Stewart, D., Harmon, I., Steinkraus, A., White, E.P., 2022. Continental-scale hyperspectral tree species classification in the United States national ecological observatory network. Rem. Sens. Environ. 282, 113264. https://doi.org/10.1016/j.rse.2022.113264.

Mayer, M., Stando, A., 2024. shapvis: SHAP Visualizations [Computer software]. https://cran.r-project.org/web/packages/shapviz/index.html.

Mäyrä, J., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Hurskainen, P., Kullberg, P., Poikolainen, L., Viinikka, A., Tuominen, S., Kumpula, T., Vihervaara, P., 2021. Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. Rem. Sens. Environ. 256, 112322. https://doi.org/10.1016/j.rse.2021.112322.

McGaughey, R.J., Ahmed, K., Andersen, H.-E., Reutebuch, S.E., 2017. Effect of occupation time on the horizontal accuracy of a mapping-grade GNSS receiver under dense forest canopy. Photogramm. Eng. Rem. Sens. 83 (12), 861–868. https://doi.org/10.14358/PERS.83.12.861.

Mcguire, A.D., Chapin III, F.S., 2006. Climate feedbacks in the Alaskan boreal forest. In: Alaska's Changing Boreal Forest. Oxford University Press, pp. 309–322.

Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Rakitin, V. Yu, 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. Physiol. Plantarum 106 (1), 135–141. https://doi.org/10.1034/j.1399-3054.1999.106119.x.

Ohse, B., Huettmann, F., Ickert-Bond, S.M., Juday, G.P., 2009. Modeling the distribution of white spruce (Picea glauca) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas. Polar Biol. 32 (12), 1717–1729. https://doi.org/10.1007/s00300-009-0671-9.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. science 333 (6045), 988–993.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., et al., 2019. PyTorch: an imperative style, high-performance deep learning library. arXiv:1912.01703 [Cs, Stat]. http://arxiv.org/abs/1912.01703.

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., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python, vol. 12, pp. 2825–2830.

Peñuelas, J., Frederic, B., Fiella, I., 1995. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. Photosynthetica 31 (2), 221–230.

Peñuelas, J., Gamon, J.A., Fredeen, A.L., Merino, J., Field, C.B., 1994. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. Rem. Sens. Environ. 48 (2), 135–146. https://doi.org/10.1016/0034-4257(94)90136-8.

Peñuelas, J., Gamon, J.A., Griffin, K.L., Field, C.B., 1993. Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. Rem. Sens. Environ. 46 (2), 110–118. https://doi.org/10.1016/0034-4257(93)90088-F.

Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Rem. Sens. Environ. 55 (2), 95–107. https://doi.org/10.1016/0034-4257(95)00186-7.

Roujean, J.-L., Breon, F.-M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Rem. Sens. Environ. 51 (3), 375–384. https://doi.org/10.1016/0034-4257(94)00114-3.

Rouse, J.W., Haas, H.R., Deering, D.W., Schell, J.A., Harlan, J.C., 1974. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. *No. E75-10354*.

Royle, D.D., Lathrop, R.G., 2002. Discriminating Tsuga canadensis hemlock forest defoliation using remotely sensed change detection. J. Nematol. 34 (3), 213.

Ruefenacht, B., Finco, M.V., Nelson, M.D., Czaplewski, R., Helmer, E.H., Blackard, J.A., Holden, G.R., Lister, A.J., Salajanu, D., Weyermann, D., Winterberger, K., 2008. Conterminous U.S. And Alaska forest type mapping using forest inventory and analysis data. Photogramm. Eng. Rem. Sens. 74 (11), 1379–1388. https://doi.org/10.14358/PERS.74.11.1379.

Schiefer, F., Kattenborn, T., Frick, A., Frey, J., Schall, P., Koch, B., Schmidtlein, S., 2020. Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. ISPRS J. Photogrammetry Remote Sens. 170, 205–215. https://doi.org/10.1016/j.isprsjprs.2020.10.015.

Shoot, C., Andersen, H.-E., Moskal, L.M., Babcock, C., Cook, B.D., Morton, D.C., 2021. Classifying forest type in the national forest inventory context with airborne hyperspectral and lidar data. Rem. Sens. 13 (10), 1863. https://doi.org/10.3390/rs13101863.

Shulski, M., Wendler, G., 2007. The Climate of Alaska. University of Alaska Press.

Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Rem. Sens. Environ. 81 (2–3), 337–354. https://doi.org/10.1016/S0034-4257(02)00010-X.

Smith, R., Adams, J., Stephens, D., Hick, P., 1995. Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. Aust. J. Agric. Res. 46 (1), 113. https://doi.org/10.1071/AR9950113.

Tian, L., Fu, W., 2020. Bi-temporal analysis of spatial changes of boreal forest cover and species in siberia for the years 1985 and 2015. Rem. Sens. 12 (24), 4116. https://doi.org/10.3390/rs12244116.

Triviño, M., Juutinen, A., Mazziotta, A., Miettinen, K., Podkopaev, D., Reunanen, P., Mönkkönen, M., 2015. Managing a boreal forest landscape for providing timber, storing and sequestering carbon. Ecosyst. Serv. 14, 179–189. https://doi.org/10.1016/j.ecoser.2015.02.003.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Rem. Sens. Environ. 8 (2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.

University of Alaska-Anchorage, A, 2017. Alaska center for conservation science. 2017. Alaska vegetation and wetland composite. https://accscatalog.uaa.alaska.edu/dataset/alaska-vegetation-and-wetland-composite.

USDA, 2018. Field Instructions for the Annual Inventory of Alaska 2018 Supplement for Interior Alaska. Department of Agriculture, Forest Service, Pacific Northwest Reserach Station, Portland, OR, USA.

Viinikka, A., Hurskainen, P., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Mäyrä, J., Poikolainen, L., Vihervaara, P., Kumpula, T., 2020. Detecting European aspen (populus tremula L.) in boreal forests using airborne hyperspectral and airborne laser scanning data. Rem. Sens. 12 (16), 2610. https://doi.org/10.3390/rs12162610.

Vogelmann, J.E., Rock, B.N., Moss, D.M., 1993. Red edge spectral measurements from sugar maple leaves. Int. J. Rem. Sens. 14 (8), 1563–1575. https://doi.org/10.1080/01431169308953986.

Walker, X.J., Baltzer, J.L., Bourgeau-Chavez, L., Day, N.J., Dieleman, C.M., Johnstone, J. F., Mack, M.C., 2020. Patterns of ecosystem structure and wildfire carbon combustion across six ecoregions of the North American boreal forest. Frontiers in Forests and Global Change 3, 87.

Walker, X.J., Mack, M.C., Johnstone, J.F., 2015. Stable carbon isotope analysis reveals widespread drought stress in boreal black spruce forests. Global Change Biol. 21 (8), 3102–3113. https://doi.org/10.1111/gcb.12893.

Walker, X.J., Okano, K., Berner, L.T., Massey, R., Goetz, S.J., Johnstone, J.F., Mack, M.C., 2023. Shifts in ecological legacies support hysteresis of stand type conversions in boreal forests. Ecosystems 26 (8), 1796–1805. https://doi.org/10.1007/s10021-023-00866-w.

Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L., Sampson, P.H., 2002. Vegetation stress detection through chlorophyll $a + b$ estimation and fluorescence effects on hyperspectral imagery. J. Environ. Qual. 31 (5), 1433–1441. https://doi.org/10.2134/jeq2002.1433.