Lichen mapping facilitated by single Photon LiDAR (SPL)

(Final report)

Baoxin Hu, Qian Li, Junwei Li, Dylan Gaspar, and Rachelle Li
Dept of Earth and Space Science and Engineering, York University
Stephen Mayor and Marc Ouellette
Ontario Ministry of Natural Resources and Forestry

Forestry Futures Trust Ontario project # KTTD 13A-2024

Date: August 31, 2025

EXECUTIVE SUMMARY

This project evaluates the utility of Single Photon LiDAR (SPL) data for mapping lichen presence. SPL is a high-density active remote sensing technology capable of penetrating canopy gaps to characterize understory conditions. The study focused on the Hearst Forest Management Unit in northeastern Ontario, leveraging the province's wall-to-wall SPL dataset. This data was integrated with 201 ground-verified plots from the Vegetation Sampling Network (VSN), where lichen presence was rigorously confirmed through both field surveys and retrospective analysis of 360° hemispherical plot photography.

A deep learning (DL) based ITC delineation method was developed and combined with a classic non-machine learning method to derive individual crowns. A comprehensive suite of structural and intensity metrics from SPL point clouds and Canopy Height Models (CHM) using both area-based and Individual Tree Crown (ITC) approaches was derived. These features were designed to proxy the ecological requirements of lichens, specifically canopy openness, vertical complexity, and light availability. Three machine learning classifiers, such as Random Forest, Support Vector Machine, and Penalized Logistic Regression were trained and evaluated on an independent test set to predict lichen presence or absence. The Random Forest model demonstrated the strongest performance, achieving a ROC-AUC of 0.73 and confirming the operational viability of SPL for habitat mapping. Feature importance analysis revealed that the most influential predictors include mean treetop height, the 95th percentile of height, and the 90th percentile of intensity. These findings suggest that lichen distribution in mixed-wood forests is primarily constrained by canopy closure and can be distinguished by specific LiDAR intensity signals that differentiate lichen mats from other forest floor elements like moss or soil. A critical outcome of the study was the identification of a trade-off in decision thresholds. Using a standard probability threshold of 0.50, the model prioritized precision (0.79), a configuration suitable for operational forest inventories where overestimating resources must be minimized. However, adopting a conservation-oriented threshold of 0.30 significantly improved recall to 0.82. This adjustment ensures that critical caribou foraging habitats are not overlooked, addressing the primary risk in species-at-risk monitoring: the omission of viable habitat.

In conclusion, this research highlights the potential of the high-density SPL data for lichen mapping in mixed-wood forests. By successfully modelling lichen presence using LiDAR-derived structural features, this project establishes a scalable framework to support sustainable forest management and woodland caribou conservation planning across Ontario. Future work of vigorous test and integrating SPL with other data sources could further enhance this framework.

1. Introduction

This project evaluates the utility of Single Photon LiDAR (SPL) data for mapping terrestrial and arboreal macrolichens in the Ontario boreal forest. The research is motivated by the need to advance innovation and knowledge in sustainable forest management and biodiversity conservation in the face of global climate change. Lichens are a critical limiting factor for woodland caribou, a threatened species. Woodland caribou select habitat with lichen disproportionately to its availability on the landscape (Mayor et al. 2009). As lichens comprise the majority of the caribou diet in winter (Webber et al., 2022) and are necessary for maintaining winter energy balance (Parker et al., 2005), their accurate mapping is essential for habitat conservation. Beyond their value to wildlife, lichens significantly influence water and nutrient cycles, thereby affecting broader forest ecosystem dynamics (Palmroos et al., 2023). Furthermore, their sensitivity to climate change, air pollution, and stand-replacing disturbance, such as wildfire and clearcut forestry, establishes them as vital bioindicators of ecosystem health and biodiversity (Abas, 2021; Esseen et al., 2021).

Despite their ecological importance, mapping lichens accurately and effectively across broad spatial areas remains a significant challenge (McMullin et al., 2011). *In-situ* monitoring and inventory are laborious, resource-intensive, and inherently limited to a small number of plots. While studies have identified key environmental drivers of lichen coverage for various tree stands (e.g., Liu et al., 2000; Wigle et al., 2021; Esseen et al., 2021), few have successfully predicted lichen cover using remote sensing data. This gap is primarily due to limitations in the spatial resolution of traditional sensors and their inability to penetrate forest canopies to detect understory features.

Recent advances in geospatial technology, machine learning, and artificial intelligence (AI) have garnered increased attention for lichen mapping. Palmroos et al. (2023) recently evaluated the potential of airborne LiDAR and hyperspectral imagery for predicting aspen epiphytic lichen communities in Finnish boreal forests. Although promising, the accuracy of such methods has historically been constrained by low LiDAR point densities (~10 pts per squared meter) and the reliance on general stand-level structures rather than individual tree-level metrics. Additionally, the high cost of traditional LiDAR acquisition often limits its operational utility.

This study leverages the Ontario government's commitment to acquiring wall-to-wall SPL coverage across the province's forest management units (Bilyk et al., 2021). SPL is particularly well-suited for lichen mapping as lichen distribution is associated with fine-scale structural attributes—such as gap distances, understory density, and canopy light penetration—which depend heavily on tree density and time since disturbance. While SPL has been used to derive forest inventory attributes using area-based approaches (e.g., Queinnec et al., 2022), further analysis is required for applications demanding detailed lower-canopy information, such as wildfire fuels and wildlife habitat assessment (Irwin et al., 2021). The co-location of SPL data, Vegetation Sampling Network (VSN) field data, and field site photographs (Ontario Ministry of Natural Resources and Forestry, 2021) presents a unique opportunity to advance lichen mapping technology. To leverage this opportunity, this project pursues three primary investigations:

1. Characterize lichen cover from field data: Analyze arboreal and terrestrial macrolichen cover using photographs and data from VSN sites to serve as ground-truth for method development and validation.
2. Derive forest structural attributes from SPL: Develop methodologies to extract LiDAR-based metrics that reflect the characteristics of individual trees and forest plots, guided by knowledge of lichen habitat requirements.
3. Model lichen distribution: Develop and evaluate a Random Forest (RF) model to predict lichen cover using the derived SPL features.

2. Stuy area and VSN plots

The study area comprises the Hearst Forest in northeastern Ontario, a landscape representative of the boreal mixed-wood ecozone. The region is dominated by extensive stands of black spruce (*Picea mariana*) and jack pine (*Pinus banksiana*), interspersed with mixed-wood assemblages of trembling aspen (*Populus tremuloides*) and white birch (*Betula papyrifera*). The terrain is largely shaped by glacial deposits, characterized by poorly drained lowlands alternating with sandy uplands that support lichen-rich pine forests.

Climatically, the Hearst Forest lies within a continental boreal regime, experiencing long, cold winters (with mean January temperatures often below –20 °C) and short, warm summers. Annual precipitation ranges from approximately 850 to 950 mm, a significant portion of which falls as snow. These conditions promote the persistence of terricolous lichen communities, particularly

Cladonia spp., which are of ecological significance as primary forage for woodland caribou (*Rangifer tarandus caribou*). As part of Ontario's managed Crown forests, the Hearst region is subject to both industrial forestry operations and conservation initiatives, making it an ideal site for evaluating the integration of Single-Photon LiDAR (SPL) with ground-based vegetation sampling for biodiversity monitoring.

Lichen presence/absence data were obtained from the VSN plots. There are 201 VSN plots in the Hearst Forest, as shown in Figure 1. There are three plot configurations, Types A, B, and C. For the VSN data, understory vegetation is only characterized for Type C. There were only six C plots in the study area. In addition to field observations, 360-degree photographs were obtained and thus this imagery was interpreted for all plots to get Lichen information.

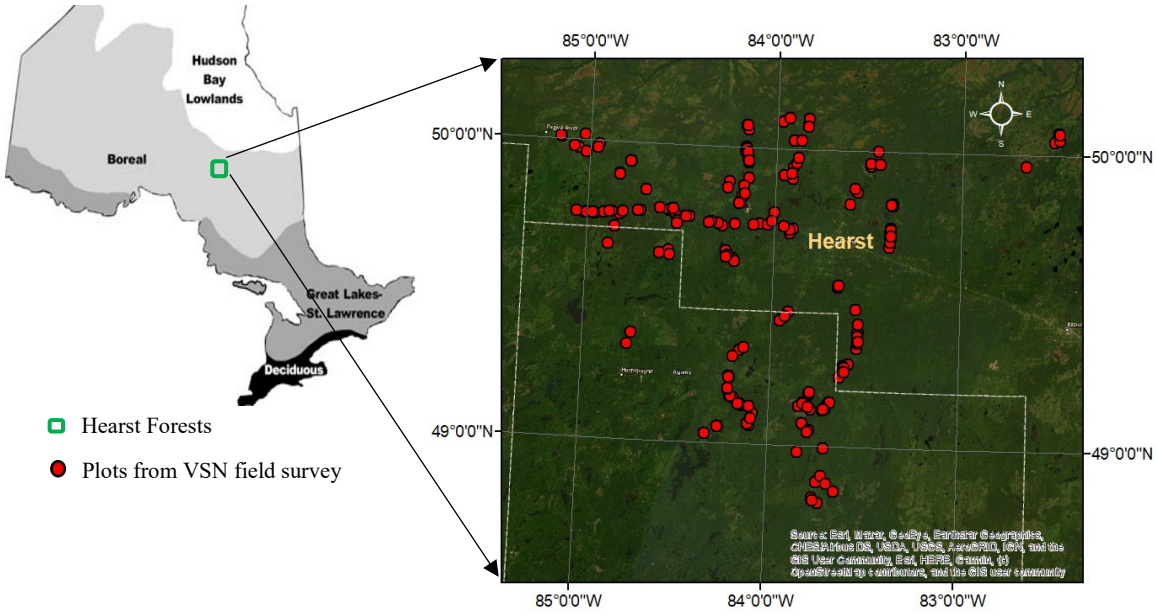

Figure 1. Study area located in Hearst Forest in northeastern Ontario.

3. Methodology

3.1 Photo-Based Ground Truth collection

To strengthen the reliability of the lichen presence/absence classification, 360° panoramic photographs captured at the center of each VSN plot were utilized. These hemispherical images, with an example shown in Figure 2, serve as a visual archive of site conditions, enabling retrospective inspection as an independent verification tool for lichen assessments.

Accurate interpretation required distinguishing between the main ecological and morphological types of lichens. Ecologically, lichens were categorized by substrate: epiphytic (tree trunks/branches), lithophytic (rocks), and terricolous (ground). The latter is of primary importance as winter forage for woodland caribou. Morphologically, three distinct body plans were identified: (1) Crustose: Forming thin, plaque-like crusts adhering closely to the substrate. (2) Foliose: Leaf-like or fern-like structures with distinct upper and lower surfaces. (3) Fruticose: Shrubby, three-dimensional forms, including the terrestrial mats critical for caribou and hanging epiphytes. Particular attention was paid to differentiating lichens from visually similar elements such as moss, fungi, and snow. While mosses often occupy similar substrates, they typically exhibit brighter green hues when hydrated and yellow-brown tones when senescent. In contrast, lichens generally appear in softer pastel tones of green, gray, or yellow green. These visual criteria formed the basis for reliable annotation.

The interpretation process followed a structured workflow. Raw panoramic images were organized into a metadata spreadsheet to record key variables, including:

- Binary Classification: Lichen presence vs. absence.
- Morphology & Context: Growth form (fruticose, foliose, crustose), coloration, and location within the image (e.g., ground mat, tree base).
- Site Metadata: Habitat context (e.g., bog, conifer stand), weather conditions, and potential disturbance signs.

Each image was manually reviewed by trained observers with academic backgrounds in ecology and forestry. Specific visual cue, such as the "varnish-like" texture of certain fungi versus the matte pastel of lichens, were used to resolve ambiguities. Table 1 illustrates the structure of the annotation spreadsheet used for Plot VSN601003.

A rigorous quality control procedure was implemented to maintain data integrity. Random subsets of images were re-checked to ensure logged attributes matched photographic evidence; inconsistencies triggered a re-evaluation of the entire batch.

This retrospective digital inspection proved vital for correcting field classification errors. For instance, in plots with high canopy closure (>80%), small fruticose mats that were overlooked during field surveys were often detected in the panoramic view. Conversely, several plots initially flagged as "present" in the field were reclassified as "absent" after high-resolution image analysis revealed that snow patches had been mistaken for crustose lichens. An example of the presence of fruticose

lichen mats on the forest floor is also shown in Figure 2. The guideline for photo interpretation is provided in Appendix A.

Figure 2. An example panoramic photograph with annotations marking the location of fruticose lichens on the forest floor in the yellow box.

Table 1. Example of the spreadsheet structure used for annotation of the Plot VSN601003.

VSN601003	Deciduous woodland filled with many saplings, dense canopy			
Photo ID	**Lichen Presence 1-Yes/0-No**	**Notes**	**Image Quality**	**Quadrant**
VSN601003202101	1	Lichen on live tree trunk	Fair	3, 4
VSN601003202102	1	Lichen on downed log, mult. different species	Fair	4
VSN601003202103	1	Downed log in midground, yellow lichen	Fair	3, 4
VSN601003202104	1	Epiphytic on birch tree log in midground	Good	3, 4
VSN601003202105	1	White lesions on downed broadleaf in back	Fair	3
VSN601003202106	0		Fair	
VSN601003202107	1	Lichen on fallen branches to lower left of photo	Fair	3
VSN601003202108	1	Lichen on branches towards middle right of photo	Fair	4
VSN601003202109	1	Lichen on branches in lower left and tree trunk to the right	Fair	3, 4
VSN601003202110	1	Lichen on trunks of left, middle, and right deciduous saplings	Good	3, 4
VSN601003202111	1	Lichen on sapling trunks throughout photo, grey plaques	Fair	3, 4
VSN601003202112	0		Fair	
VSN601003202113	1	Grey lichen plaques on downed logs in midground	Good	3
VSN601003202114	1	Foliose lichen on downed tree in lower left midground	Good	3
VSN601003202115	1	Light patches on trunk of sapling in middle of photo	Fair	3, 4
VSN601003202116	0		Fair	
VSN601003202117	1	Mint and grey patches on decaying log, centre left back	Fair	3, 4
VSN601003202118	1	Grey plaques on young trees to the right	Good	3, 4

3.2 SPL data processing

A SPL system is an efficient and potentially cost-effective means forest monitoring (Swatantran et al., 2016) due to its many advanced features compared with the traditional LiDAR instrument. SPL instruments have high sensitivity that permits the emission of weak laser pulses since only a few backscattered photons are needed (Wästlund et al., 2018). Its short pulse width and high-resolution timing detectors enable a good resolution in range. Furthermore, SPL instruments split a laser pulse to a 10 × 10 array of laser beamlets that illuminate the surface, and a corresponding 10 × 10 array of detectors is used to record the backscattered photons (Boretti, 2023). These properties enable SPL instruments to obtain high point density even at a high altitude. For forest applications, the only disadvantage of SPL is its working wavelength, typically at 532 nm (green). The traditional LiDAR instruments work in a near-infrared wavelength. Leaves of green vegetation reflect more near-infrared radiation than green radiation. Compared with the traditional LiDAR data, the point density of SPL data tends to be high. In addition, the proportion of the first returns of SPL data is high.

SPL data covering the Hearst Forest were processed to ensure precise spatial correspondence with VSN field plots. To derive predictors for lichen presence modelling, SPL point clouds were clipped to 20 × 20 m plot windows centered on the GPS coordinates recorded during field survey. The choice of a fixed-area square window ensured spatial consistency across plots, facilitated raster-based analyses, and aligned with area-based LiDAR feature extraction approaches commonly applied in forest inventory studies (Queinnec et al., 2022; Penner et al., 2013).

Accurate co-registration between LiDAR subsets and ground plots was critical. We standardized all plot centers to a single project Coordinate Reference System (CRS) matching the CHM/LAZ headers (UTM). Field-collected GPS positions, originally recorded in WGS84 latitude/longitude or legacy UTM variants, were reprojected to the project CRS using the PROJ library (via pyproj) with datum grid shifts enabled where applicable. Historical UTM columns were retained solely for diagnostic purposes to quantify discrepancies and verify zone/datum consistency. Using the validated plot centers, 20 × 20 m plot polygons were generated to clip both the Canopy Height Models (CHM) and raw point-cloud (LAZ) data. This ensured that each LiDAR subset accurately represented the forest structure of its corresponding VSN plot. Figures 3 and 4 illustrate examples of SPL-derived CHM and point cloud subsets for plots VSN601001 and VSN601003, respectively.

These corrected, plot-level LiDAR datasets served as the foundation for extracting the structural metrics used in subsequent lichen presence modelling.

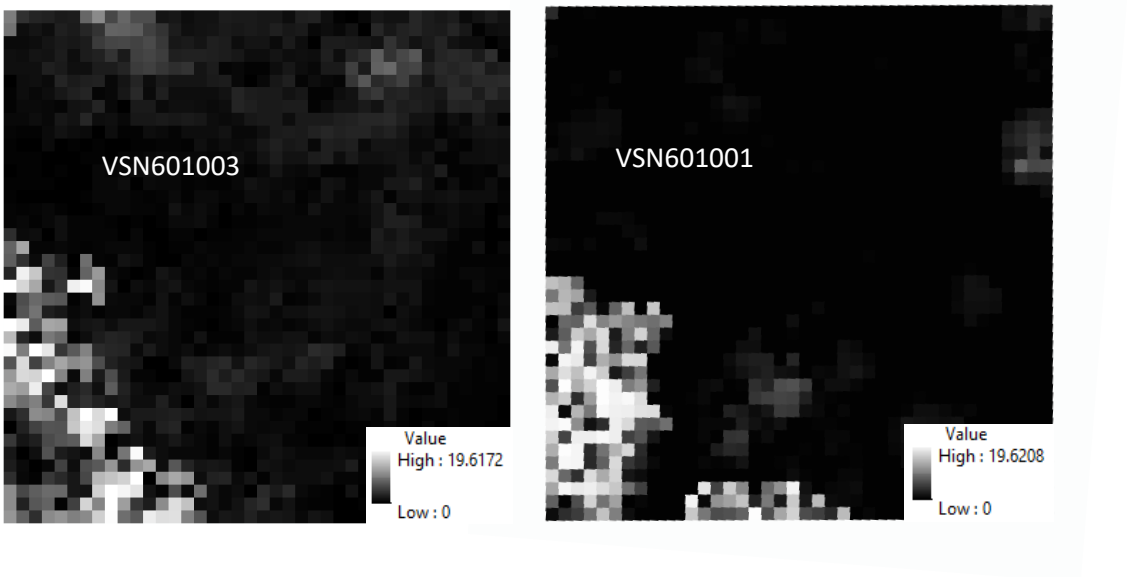

Figure 3. Example of SPL-derived CHM subsets clipped to 20 × 20 m windows for two VSN plots

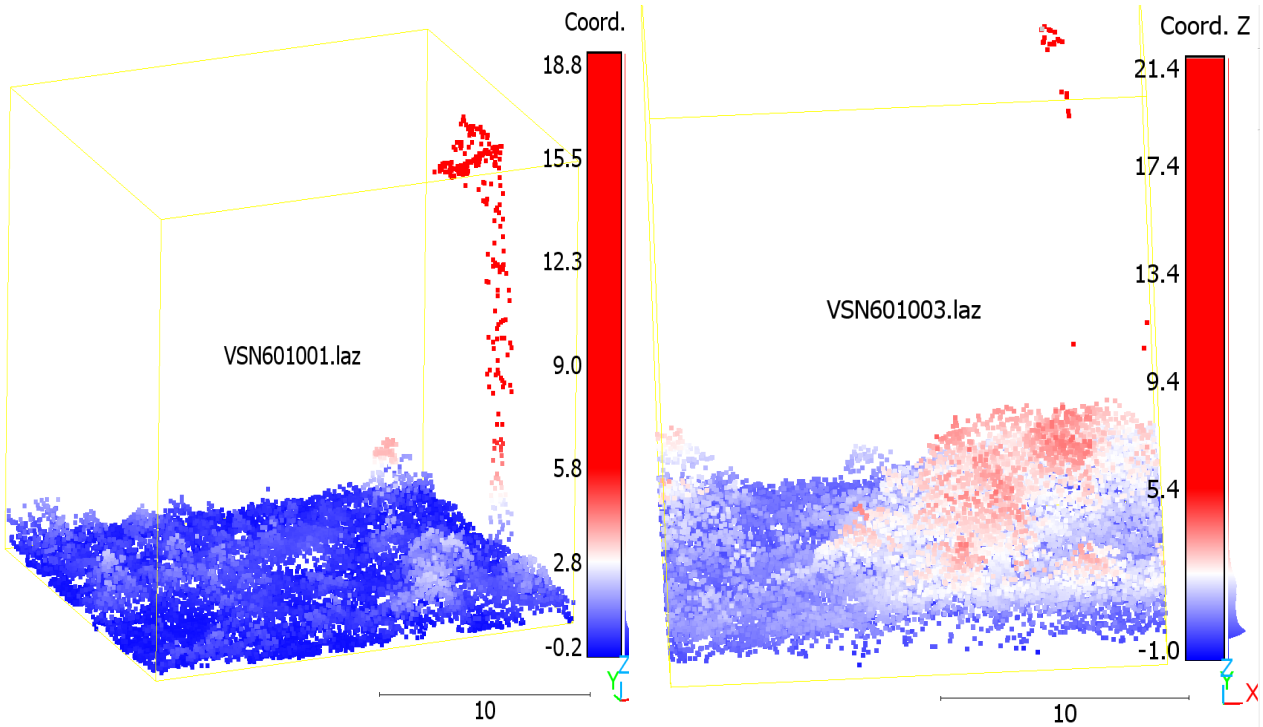

Figure 4. Example of SPL point cloud subsets clipped to 20 × 20 m windows for two VSN plots

3.3LiDAR feature extraction

3.3.1 Individual tree crown detection and extraction

In this project, features were extracted using area-based and individual tree-based approaches. As a result, individual tree crown (ITC) delineation was first carried out. Two approaches were investigated: non-machine learning method (referred as classic method hereafter) and deep learning (DL) method. For both methods, the CHM at the spatial resolution of 0.5m by 0.5 was used. For the CHM imagery, pre-processing included pit filling and contrast enhancement, followed by construction of a conservative canopy candidate mask that combined a minimum height threshold, a minimum connected-area constraint, and a local roughness filter to suppress spurious low-relief surfaces (Chen et al., 2006). For DL method, LiDAR data cloud was also used. In the following, the classic method was described. For the DL method, please see the publication by Li et al., (2025).

Treetop candidates were then identified using a multi-scale Laplacian-of-Gaussian (LoG) operator with local-height–difference (LHD) background removal (Vega et al., 2014). Scale-adaptive percentile thresholds were applied per LoG scale, and candidates were screened via Hessian-

eigenvalue tests to retain blob-like peaks (Dalponte et al., 2016). A two-stage, scale-aware non-maximum suppression was used to merge nearby responses while preserving distinct peaks over flat-topped or coalescent crowns; for plateau-like maxima, the representative peak within the connected support was selected using the LHD score (Yu et al., 2010). The surviving peaks served as markers for a marker-controlled watershed segmentation run on a smoothed gradient (Sobel) image of the CHM, with a compactness prior to limit radial leakage, an approach widely used in individual-tree crown delineation (Li et al., 2012; Silva et al., 2016). Post-segmentation quality control removed sub-threshold fragments, flagged crowns touching plot borders, and applied a soft size prior (expressed as a permissible radius range that increases with treetop height) to identify over-expanded segments.

The crowns derived by these two methods were combined. If there were inconsistent, manual observation were engaged to choose the most appropriate ones. In future work, this process will be automated.

Per-tree (ITC-level) attributes exported for modelling comprised planar coordinates and pixel indices of the treetop, peak height and the LoG detection scale (with response score), crown area and equivalent radius, the height-based radius prior and a boolean “oversize” indicator, an edge-contact flag, and within-crown height summaries (mean, maximum). Plot-level aggregations derived from these objects included tree count, total crown area and cover ratio, mean crown radius and its coefficient of variation, proportions of edge and oversize trees, and mean treetop height. Together, these ITC descriptors quantify crown geometry and vertical structure at tree resolution in a manner directly compatible with downstream analyses of lichen presence.

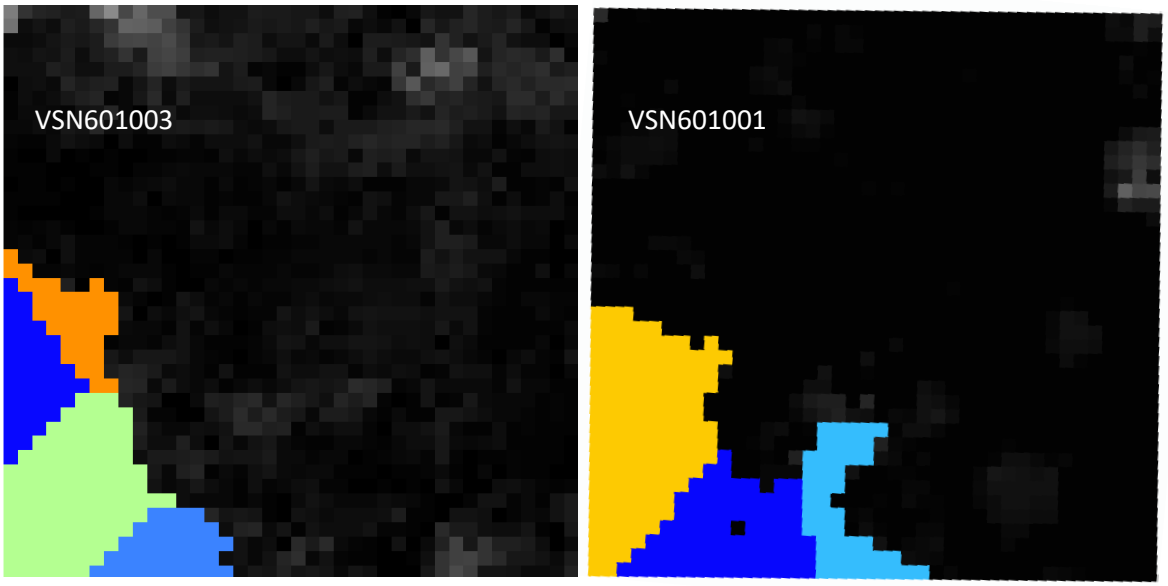

Figure 5. Example of ITC delineation results (masks in yellow, blue, green colors) using the proposed method and CHM for two VSN plots named VSN 601001 and VSN 601003.

3.4Feature Extraction

To characterize forest structure relevant to lichen presence, we derived a comprehensive set of structural and intensity features from both the Canopy Height Model (CHM) and the raw SPL point clouds (LAZ). These features were designed to capture vertical heterogeneity, canopy openness, and ground-level conditions known to influence lichen occurrence. A summary of all extracted features is provided in Table 2.

3.4.1 CHM-derived features: From the rasterized CHM, we calculated metrics describing the canopy surface geometry and complexity:

- Descriptive Statistics: Mean, standard deviation, minimum, maximum, interquartile range (IQR), skewness, and kurtosis. These capture general vertical variability.
- Canopy Openness & Gaps: Canopy closure and gap metrics (count, fraction, and maximum gap size) were extracted at two height thresholds (3 m and 5 m) to quantify light penetration and understory habitat suitability.
- Texture & Roughness: Multi-scale roughness indices (standard deviations within 1 m and 3 m windows) and texture variance were computed to characterize fine-scale canopy surface complexity.

3.4.2 Point Cloud-Derived Features: At the point-cloud level (LAZ), we computed metrics to describe the vertical distribution of returns and signal intensity:

- Height Distribution: Mean, percentiles (–), IQR, skewness, and kurtosis of return heights.
- Vegetation Stratification: Ratios of returns were calculated for specific layers to represent shrub (< 0.5 m, < 1 m), mid-story (0.15–5 m), and high-canopy (> 5 m) dominance. Foliage Height Diversity (FHD) was calculated using 0.5 m vertical bins as an index of vertical complexity.
- Penetration: The proportion of first returns below 2 m was used to indicate canopy penetration into the understory.

- Intensity: Intensity-related features (mean, standard deviation, coefficient of variation, percentiles) were calculated for the whole column and specifically for returns below 1 m. These metrics are critical as the reflectance properties of SPL signals at 532 nm can enhance the separability of lichen mats from other understory vegetation.

3.4.3 Individual tree-based Features: Aggregated metrics derived from the Individual Tree Crown (ITC) delineation (Section 3.4) provided object-based structural context:

- Stand Density: Tree density per plot and total crown area.
- Canopy Geometry: Cover ratio (proportion of ground covered by crowns), mean equivalent crown radius, and the coefficient of variation for crown size.
- Dominance & Competition: Mean treetop height (representing dominant canopy height), the percentage of edge trees, and the proportion of "oversized" crowns.

3.4.5 Feature selection: Given the large number of extracted features, a selection process was implemented to reduce redundancy and focus on ecologically interpretable predictors. Correlated variables were screened using pairwise correlation thresholds, and those with high collinearity were excluded. Remaining features were prioritized based on their ecological relevance to canopy openness, vertical stratification, and light availability—factors repeatedly identified as critical for terricolous and epiphytic lichen communities. This process ensured a parsimonious feature set for subsequent modelling while retaining the key structural gradients most relevant to lichen presence.

Table 2. Summary of LiDAR-derived features used in lichen presence modeling from 20 × 20 m plots

Category	Feature	Description	References
CHM derived features (Plot-level)	chm_mean, chm_std, chm_max, chm_min	Basic canopy height statistics (mean, variability, extremes) from CHM.	Hillman & Nielsen 2020; Wężyk et al. 2019
	chm_iqr, chm_skew, chm_kurtosis	Distribution shape of canopy height values (heterogeneity, asymmetry, tail weight).	He et al. 2021
	chm_num_pixels	Number of valid CHM pixels within plot (proxy for ground area).	Wężyk et al. 2019
	canopy_cover_gt3m, canopy_cover_gt5m	Proportion of canopy area above 3 m / 5 m thresholds (stand openness).	Hillman & Nielsen 2020; Węgrzyn et al. 2020
	gap_fraction_lt3m, gap_fraction_lt5m	Fraction of open gaps below 3 m / 5 m threshold.	Robichaud 2019

Category	Feature	Description	References
	gap_count_lt3m, gap_count_lt5m	Number of gaps (<3 m / <5 m) as indicator of forest discontinuity.	Robichaud 2019
	gap_max_area_lt3m, gap_max_area_lt5m	Largest gap area within plot.	Hillman & Nielsen 2020
	chm_local_std_1m, chm_local_std_3m	Local height variability in 1 m / 3 m windows (structural roughness).	Wężyk et al. 2019
	chm_texture_variance	Mean of local variance filter over CHM, proxy for texture.	He et al. 2021
Point cloud derived features (Plot-level)	z_mean, z_std, z_max, z_min	Vertical height distribution statistics of all LiDAR returns.	Hillman & Nielsen 2020
	z_iqr, z_skew, z_kurtosis	Vertical structural heterogeneity and distribution shape.	Wężyk et al. 2019
	z_p10, z_p25, z_p50, z_p75, z_p95	Percentile heights describing canopy profile.	He et al. 2021
	fhd_0p5m	Foliage height diversity index at 0.5 m binning (vertical complexity).	Rautiainen et al. 2024
	ratio_below_0p5m, ratio_below_1m	Fraction of returns below ground/shrub layers (0.5 m, 1 m).	Hillman & Nielsen 2020
	mid_story_ratio	Fraction of returns from mid-story (0.15–5 m).	Węgrzyn et al. 2021
	high_canopy_ratio	Fraction of returns >5 m (overstory dominance).	Hillman & Nielsen 2020
	num_points	Total LiDAR returns per plot.	He et al. 2021
	ratio_first_returns, ratio_last_returns, ratio_multi_returns	Proportion of first, last, and multiple returns, reflecting canopy penetration and density.	Wężyk et al. 2019
	first_returns_below_2m	Fraction of first returns <2 m (understory penetration).	Hillman & Nielsen 2020
	intensity_mean, intensity_std, intensity_cv	Mean, variability, and coefficient of variation of return intensity.	Rautiainen et al. 2024
	intensity_p10, intensity_p50, intensity_p90	Percentiles of intensity distribution.	He et al. 2021
	intensity_low1m_mean, intensity_low1m_p90	Intensity statistics for returns below 1 m, reflecting understory reflectance.	Robichaud 2019
Tree crown metrics (ITC-level)	n_trees	Number of detected tree crowns per plot.	Palmroos et al. 2023
	sum_crown_area_m2, cover_ratio	Total crown area and proportion of plot covered by crowns.	Wężyk et al. 2019
	mean_radius_m, cv_radius	Average and variability of crown equivalent radius.	Węgrzyn et al. 2020
	pct_edge_trees	Proportion of trees at plot boundary (edge effect).	Hillman & Nielsen 2020
	pct_oversize	Proportion of oversized crowns relative to expected CHM structure.	Wężyk et al. 2019
	mean_peak_h	Mean treetop height across ITCs.	Węgrzyn et al. 2021

3.5 Modeling

To predict lichen presence from SPL-derived features, we implemented three complementary classifiers: Random Forest (RF) (Breiman, 2001), a penalized logistic regression with elastic-net regularization (Logit-EN) (Zou et al., 2005), and a radial-basis-function support vector machine (SVM-RBF) (Razaque et al., 2021) using the scikit-learn framework (Prodregosa, 2011). They were selected due to complementary inductive biases, interpretable linear baselines, nonparametric interaction capture, and flexible kernel decision boundaries. They are well suited to tabular ecological data with correlated predictors and nonlinear responses, while yielding calibrated probabilities for decision-relevant thresholds, and provide robust, interpretable evidence for ecological drivers of lichen presence in forest applications (He et al., 2021; Palmroos et al., 2023).

The feature matrix consisted of structural and intensity-based predictors derived from 20 × 20 m clipped SPL point clouds and CHM. The binary response variable was lichen presence/absence, as verified by VSN Type C plots and 360° plot photographs. Prior to modeling, non-numeric fields were excluded, and missing values were imputed using column means. The dataset included 201 plots, with 99 classified as lichen-present and 102 as lichen-absent. Following stratification by lichen presence, the data were randomly split into training (80%) and testing (20%) subsets, resulting in balanced representation across folds.

We trained a random forest, an elastic-net penalized logistic regression, and an RBF-kernel SVM on the imputed feature set. For RF, tree number, depth, node size and feature subsampling were tuned by stratified 5-fold cross-validation on the training split; the final configuration favored shallow-to-moderate depth with $\sqrt{p}$ feature subsampling to balance bias and variance. For Logit-EN and SVM, scaling was applied within pipelines; penalty strength, mixing ratio (for EN), and kernel width (γ) were selected by the same cross-validated procedure. To obtain well-calibrated probabilities for thresholding and PR metrics, models were post-calibrated (isotonic for RF; Platt/sigmoid for linear and SVM). Full hyperparameter grids and selected values are provided in Table 3.

Table 3. The hyperparameters table used in RF, Logit-EN, and SVM-RBF models.

Model	Tuning protocol	Selected hyperparameters
RF	5-fold CV on train	n_estimators=600; max_depth=12; min_samples_leaf=2; max_features=$\sqrt{p}$; bootstrap
Logit-EN	5-fold CV on train	C=1.0; l1_ratio=0.5; solver=SAGA; class_weight=balanced
SVM-RBF	5-fold CV on train	C=1.0; γ=scale; probability=True; class_weight=balanced

Given the asymmetric costs in conservation detection tasks (missed presences typically worse than false alarms), the decision threshold for each model was selected on the training set using 5-fold stratified cross-validation to maximize β=1.5. Concretely, we scanned thresholds from 0.01 to 0.99, computed fold-wise, and adopted the median of fold-optimal thresholds as the operating point; for comparison, we also report metrics at the conventional 0.50 threshold. On the held-out test set, we evaluated: ROC AUC, average precision (area under the precision–recall curve), Brier score (for probabilistic calibration), accuracy and F1 (for the two operating points). To interpret predictors consistently across models, we computed permutation importance on the test set using average precision as the scoring function with 200 repeats per feature. For each model we the share-of-total importance (each positive importance divided by the sum over features, expressed as %). For communication, we provide Top 20 horizontal bar plots of share-of-total importance.

4. Results

4.1 Feature importance

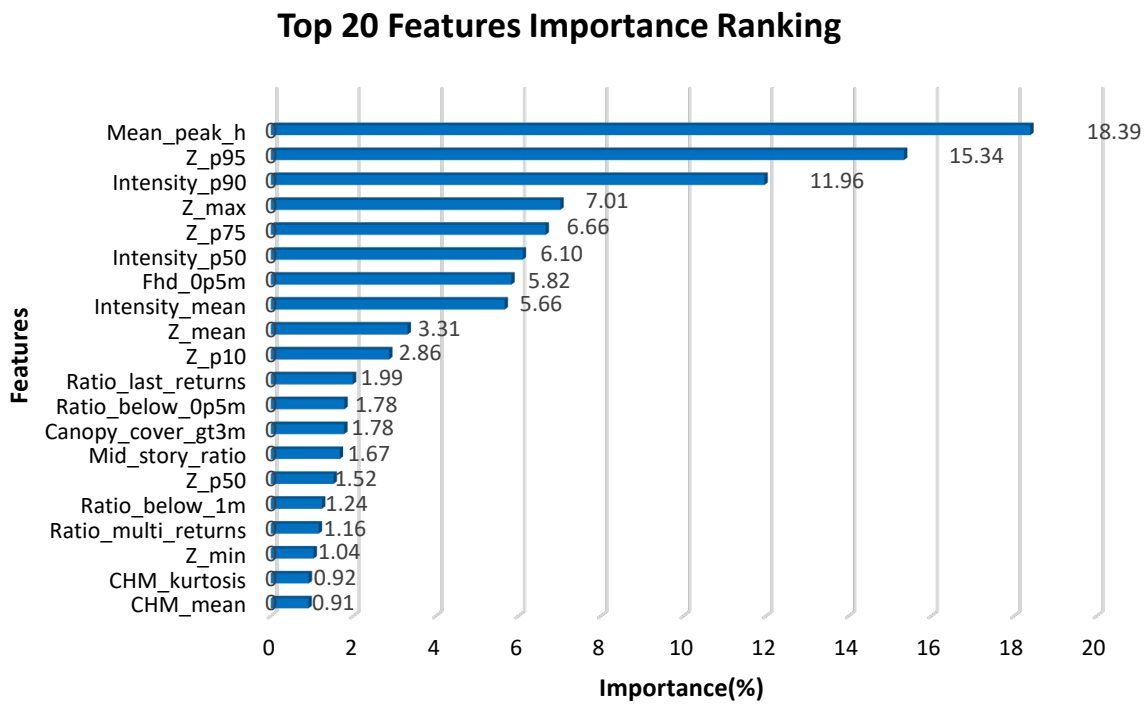

Figure 6. The share-of-total importance ranking of top 20 features used in RF

On the held-out test set, permutation importance (average-precision scoring; 200 repeats) indicated a concentrated contribution of a small subset of predictors to Random Forest performance (see Figure. 6 for the ranking). The three highest-ranked variables were Mean_peak_h (18.39% share of total importance), Z_p95 (15.34%), and Intensity_p90 (11.96%); together they accounted for 45.69% of the total. Adding Z_max (7.01%) and Z_p75 (6.66%) increased the cumulative share to 59.36%. The next group comprised Intensity_p50 (6.10%), FHD_0p5m (5.82%), and Intensity_mean (5.66%), raising the cumulative total to 77.60%. Mid-range height and layer-fraction variables contributed smaller but non-zero shares (e.g., Z_mean 3.31%, Z_p10 2.86%, Ratio_last_returns 2.13%, Canopy_cover_gt3m and Ratio_below_0p5m each 1.78%). Remaining predictors each contributed ≤1.67%, with the lowest shares observed for CHM_kurtosis (0.92%) and CHM_mean (0.91%). All percentages refer to the normalized (share-of-total) permutation importance computed on the test partition.

4.2 Lichen prediction

The Random Forest classifier trained on SPL-derived structural features and CHM metrics showed strong discriminatory power for predicting lichen presence/absence. Threshold-independent evaluation indicated a ROC-AUC of 0.73 and an Average Precision (AP/PR-AUC) of 0.67 on the test set (n = 41), demonstrating robust separation between presence and absence plots across the full range of probability thresholds. The Brier score of 0.17 further suggested that predicted probabilities were reasonably well calibrated.

When predictions were evaluated using a fixed threshold of 0.50, the model achieved an overall accuracy of 0.69 with an F1-score of 0.66 for the lichen-present class (Table 4). At this threshold, presence plots were predicted to be high precision (0.79) but moderate recall (0.57), while absence plots were classified with high recall (0.82) but lower precision (0.63). This pattern indicates that the model was conservative in assigning presence, thereby reducing false positives but missing some true presence cases.

By contrast, applying the cross-validated optimal threshold of 0.3, selected to maximize the F1.5 score and prioritize recall, substantially increased sensitivity to lichen presence. At this operating point, the model achieved a recall of 0.82 for presence plots, ensuring that most lichen occurrences were detected. However, this improvement in recall was accompanied by reduced precision (0.58) and a lower overall accuracy of 0.61. The F1-score for the presence class at this

threshold was 0.68, comparable to the 0.50 threshold, highlighting the trade-off between minimizing omission errors (false negatives) and controlling commission errors (false positives).

Across both thresholds, the ROC-AUC and AP remained unchanged (0.73 and 0.67, respectively), reflecting the inherent threshold independence of these global performance measures. These values demonstrate that the model's ability to separate lichen-present from lichen-absent plots was stable regardless of the decision threshold, while accuracy, precision, recall, and F1 varied with threshold selection.

Table 4. Performance of the Random Forest classifier under two decision thresholds.

Threshold	Accuracy	Precision (Presence)	Recall (Presence)	F1 (Presence)	ROC-AUC	PR-AUC/AP	Brier Score
0.50 (fixed)	0.69	0.79	0.57	0.66	0.73	0.67	0.17
0.30 (CV-optimal, F1.5)	0.61	0.58	0.82	0.68	0.73	0.67	0.17

Table 5. Performance comparison of the three models at the conventional 0.50 threshold.

Model	Threshold	ROC AUC	PR AUC	Brier	Accuracy	F1
RF	0.50	0.63	0.63	0.32	0.69	0.66
Logit-EN	0.50	0.50	0.54	0.26	0.59	0.52
SVM-RBF	0.50	0.55	0.58	0.28	0.61	0.65

Table 5 presents the performance of the three models at the conventional 0.50 threshold. The Random Forest (RF) yielded the highest overall performance, with ROC AUC and PR AUC of 0.63, accuracy of 0.69, and an F1 score of 0.66. The SVM with RBF kernel achieved moderate results, with ROC AUC of 0.55, PR AUC of 0.58, accuracy of 0.61, and an F1 score of 0.65. Logistic Regression with elastic-net penalty (Logit-EN) showed the lowest performance, with ROC AUC of 0.50, PR AUC of 0.54, accuracy of 0.59, and an F1 score of 0.52. Overall, the results demonstrate that SPL-derived structural metrics integrated with ground-based VSN observations can successfully predict lichen presence at the plot scale. While the model performance is constrained by the limited

number of test samples, the predictive ability (ROC-AUC > 0.73) indicates potential for scaling up to broader landscapes.

5 Discussion and conclusion

5.1 Contribution of SPL-derived features to lichen presence

Ecologically, the dominance of tree-height maxima (Mean_peak_h, Z_p95, Z_max) is consistent with evidence that overstory stature and cover regulate understory composition, lichen cover declines as stands become taller/denser, while open, lower canopies favor terricolous lichens over bryophytes. Independent field-based studies similarly identified canopy cover as a primary predictor of lichen abundance and reported strong negative effects of increasing cover, aligning with our model's non-zero importance for canopy metrics and high weight on upper-canopy statistics (which covary with closure).

The intensity suite (Intensity_p90, Intensity_p50, Intensity_mean) showed substantial importance, indicating additional separability beyond height alone. Prior ALS work documented class-level differences in normalized intensity between lichen and other surfaces and emphasized the value of careful intensity handling for vegetation discrimination, which helps explain the high ranks of our upper-tail intensity percentiles.

Mid-story and near-ground fractions (mid_story_ratio, ratio_below_0p5m/1m) contributed at intermediate levels, consistent with the idea that light transmission through the vertical profile and the balance of returns across layers capture openness relevant to lichen growth. The modest, yet non-zero, influence of return-order metrics (ratio_last_returns, ratio_multi_returns) is plausible given that first returns tend to represent the top of canopy whereas subsequent returns sample sub-canopy layers, linking these attributes to understory exposure.

Finally, the relatively low weights for aggregate CHM moments (e.g., CHM_mean, CHM_kurtosis) suggest that tree-resolved and percentile-based descriptors carry more discriminative signal for lichen presence than bulk canopy summaries in our mixed-wood plots, a pattern compatible with studies that favored canopy cover and height extremes over coarse height averages when relating overstory structure to lichen abundance.

The RF importance profile supports a mechanistic picture where dominant height/upper-canopy structure and high-percentile intensity are primary drivers, vertical complexity and near-ground exposure provide secondary information, and coarse canopy aggregates are least informative.

This ordering is consistent with the literature linking open, shorter canopies and distinct reflectance properties to higher lichen occurrence.

5.2 Lichen presence prediction using proposed model

This study investigates that SPL-derived plot and ITC level features, when integrated with Vegetation Sampling Network (VSN) ground data, provide a robust framework for predicting lichen presence in boreal mixed-wood forests. The model achieved stable threshold-independent performance (ROC-AUC = 0.73, PR-AUC = 0.67), indicating good overall discriminative ability. However, the selection of decision thresholds substantially influenced ecological interpretations, particularly with respect to omission versus commission errors. At the fixed threshold of 0.5, the model favored precision in predicting lichen presence (0.79) but produced lower recall (0.57). This conservative configuration minimizes false positives and may be useful in operational inventory contexts where overestimating resources is undesirable. Yet, it also risks underestimating true lichen presence, potentially biasing habitat assessments. By contrast, the cross-validated optimal threshold of 0.3 emphasized recall, correctly identifying 82% of lichen-present plots. This threshold reduced omission errors but increased commission errors, lowering precision to 0.58. From an ecological monitoring and conservation standpoint, prioritizing recall is often preferable, as omission errors can result in critical habitats being overlooked. This distinction is particularly relevant for woodland caribou (*Rangifer tarandus caribou*), which rely heavily on terricolous lichens for winter forage (Węgrzyn et al., 2020). Also, this aligns with biodiversity monitoring and habitat conservation objectives, where omission errors have greater ecological costs than commission errors (Cosgrove et al., 2024). In contrast, the 0.50 threshold scenario is better suited to forest inventory and management reporting, where accuracy and resource estimation are prioritized.

The importance of gap fraction, canopy openness, and vertical entropy in our model is consistent with prior findings. Studies in pine-dominated forests demonstrated that increasing tree height and canopy closure suppress terricolous lichens (Węgrzyn et al., 2021). Similarly, a Canadian study found that the detectability of lichens declines steeply once canopy closure exceeds 77–88% (Robichaud, 2019). Our results align with these patterns: omission errors at the 0.50 threshold were concentrated in dense-canopy plots where SPL-derived gap fractions were low, suggesting limited understory light penetration. Conversely, the 0.297 threshold reduced such omissions by lowering

the decision boundary, effectively capturing presence in more marginal canopy conditions. While hyperspectral remote sensing approaches have successfully mapped epiphytic lichens through reflectance properties (Rautiainen et al., 2024), they remain limited by canopy occlusion and difficulties in discriminating lichens from mosses and bryophytes. SPL, by contrast, does not directly detect lichens spectrally but models habitat suitability through structural proxies such as canopy cover fraction and understory openness. This distinction highlights the complementary potential of structural LiDAR metrics and spectral imagery. Multi-scale studies on reindeer lichens (He et al., 2021) suggest that integrating SPL with hyperspectral or UAV-based datasets could improve species-level discrimination and support abundance estimation, rather than binary presence/absence classification.

Our Random Forest models achieved robust performance (ROC-AUC = 0.73, PR-AUC = 0.67), confirming that SPL-derived features capture key ecological gradients influencing lichen distributions. Importantly, the study highlights that threshold selection is not merely a statistical choice but an ecological one: while a conventional threshold (0.5) favors conservative precision-oriented predictions, a cross-validated optimal threshold (0.3) prioritizes recall and better safeguards against omission errors that could lead to underestimation of critical lichen habitats. These findings align with previous work demonstrating the negative influence of tree height and canopy closure on terricolous lichens and reinforce the value of LiDAR-derived canopy openness for habitat assessment (Robichaud, 2019; Węgrzyn et al., 2021). They also complement multi-scale remote sensing approaches that combine structural and spectral data for lichen fractional cover mapping, suggesting that SPL can provide a structural baseline upon which spectral methods may be layered to achieve finer species-level or abundance estimates (He et al., 2021).

Beyond methodological contributions, this work has direct implications for biodiversity monitoring and conservation. Terricolous lichens are a keystone forage resource for woodland caribou and a sensitive indicator of forest integrity (Węgrzyn et al., 2020). By operationalizing SPL-based lichen mapping across Ontario's provincial dataset, managers can identify and monitor potential lichen habitats even under partially closed canopies where optical methods struggle (Korpela, 2008). Such structural mapping complements emerging hyperspectral applications and can inform both provincial forest inventories and species-at-risk recovery planning.

5.3 Limitation

Although the model produced reliable and consistent results, several limitations must be acknowledged. First, the analysis was based on a relatively small and spatially constrained set of labeled field plots (n=201). This restricted the representation of the full range of lichen habitat conditions and potentially reduced the model's ability to capture rare or transitional sites. Second, the use of binary presence/absence labeling simplified the complex ecological gradient of lichen cover, potentially overlooking within-plot variability. Third, the modelling relied solely on structural metrics from SPL. While SPL provides detailed canopy and understory information, it lacks spectral sensitivity, which may lead to confounding lichens with other bright understory elements such as bryophytes or exposed mineral soil. Fourth, in this study, the terrestrial and arboreal lichens were combined due to the lack of ground truth data from individual categories.

Finally, minor spatial or temporal mismatches between field plots and LiDAR data acquisition may have contributed to residual uncertainty.

6. Conclusion

This study establishes a scalable, operational pathway for lichen mapping in Ontario's boreal mixed-wood forests. By combining high-density Single Photon LiDAR (SPL) data with standardized VSN field plots, we successfully modelled lichen presence/absence, overcoming the limitations of traditional optical remote sensing in closed-canopy environments. The Random Forest model identified canopy height, intensity distribution, and vertical complexity as key predictors, reflecting the known ecological dependence of lichens on light availability and open forest structures.

Future work should focus on expanding the number and diversity of labeled plots by working with other forests and further test the algorithms. Combining SPL with complementary remote sensing datasets—such as airborne hyperspectral imagery, UAV-based multispectral data, or Sentinel-2 reflectance products—could enhance the capacity to distinguish lichen species and estimate their abundance rather than just their presence. Broader integration of structural, spectral, and ecological data will improve large-scale biodiversity mapping and support more effective conservation planning for woodland caribou (*Rangifer tarandus caribou*) and other lichen-dependent species. This work shows the potential of the provincial SPL datasets as a vital tool for sustainable forest management and biodiversity monitoring across Ontario.

7.Reference

Abas, A., 2021. A systematic review on biomonitoring using lichen as the biological indicator: A decade of practices, progress and challenges, *Ecological Indicators* 121, 107197.

Bilyk, A., R. Pulkki, C. Shahi, et al., 2021. Development of the Ontario forest resources inventory: A historical review, *Can. J. For.Res*.51,198–209.

Breiman, L., 2001. Random forests. *Machine Learning*, 45(1), 5–32.

Cosgrove, C. F., Coops, N. C., & Martin, T. G., 2024.. Using the full potential of Airborne Laser Scanning (aerial LiDAR) in wildlife research. *Wildlife Society Bulletin*, 48(3), e1532.

Dalponte, M., and Coomes, D. A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods in Ecology and Evolution*, 7(10), 1236–1245.

Esseen, P.-A., M. Ekström, A. Grafström et al., 2021. Multiple drivers of large-scale lichen decline in boreal forest canopies, *Global Change biology*, DOI: 10.1111/gcb.16128.

He, L., Chen, W., Leblanc, S. G., Lovitt, J., Arsenault, A., Schmelzer, I., Fraser, R. H., Latifovic, R., Sun, L., & Prévost, C., 2021. Integration of multi-scale remote sensing data for reindeer lichen fractional cover mapping in Eastern Canada. *Remote Sensing of Environment*, 267, 112731.

Irwin, L., N. C. Coops, M. Queinnec, et al., 2021. Single photon lidar signal attenuation under boreal forest conditions, *Remote Sensing Letters*, 12:10, 1049-1060.

Korpela, I. S., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. *Remote Sensing of Environment*, 112(10), 3891–3897.

Li, Q., B. Hu, J. Shang, and T. Remmel, 2025. Two-Stage Deep Learning Framework for Individual Tree Crown Detection and Delineation in Mixed-Wood Forests Using High-Resolution Light Detection and Ranging Data. *Remote Sensing*, 17(9), 1578.

Li, W., Guo, Q., Jakubowski, M. K., & Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering & Remote Sensing*, 78(1), 75–84.

Liu, C., H., Ilvesniemi, and C. J., Westman, C. J., 2000. Biomass of arboreal lichens and its vertical distribution in a boreal coniferous forest in central Finland, *Lichenologist* 32(5): 495–504

Mayor, S.J., J.A. Schaefer, D.C. Schneider, S.P. Mahoney. 2009. The spatial structure of habitat selection: A caribou's-eye-view. *Acta Oecologica*, 35:253-260.

McMullin, R. T., B.W. Lacey, and S.G. Newmaster, 2011. Estimating the biomass of woodland caribou forage lichens, *Can. J. For. Res*. 41: 1961–1969.

Ontario Ministry of Natural Resources and Forestry, 2021. Vegetation Sampling Network Protocol: Technical specifications for field plots. Ontario Ministry of Natural Resources and Forestry, Science and Research Branch, Peterborough, ON. Science and Research Technical Manual TM-10. 173 p. + appendices.

Palmroos, I., Norros, V., Keski-Saari, S., Mäyrä, J., Tanhuanpää, T., Kivinen, S., Pykälä, J., Kullberg, P., Kumpula, T., & Vihervaara, P. (2023). Remote sensing in mapping biodiversity–a case study of epiphytic lichen communities. Forest Ecology and Management, 538, 120993.

Parker K.L., Barboza P.S., and Stephenson T.R. 2005. Protein conservation in female caribou (Rangifer tarandus): effects of decreasing diet quality during the winter. *J. Mammal*. 86(3): 610–622.

Prodregosa, F., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*. Vol, 12, 2825–2830.

Rautiainen, M., Kuusinen, N., & Majasalmi, T., 2024. Remote sensing and spectroscopy of lichens. *Ecology and Evolution*, 14(3), e11110.

Razaque, A., Ben Haj Frej, M., Almi'ani, M., Alotaibi, M., & Alotaibi, B., 2021. Improved support vector machine enabled radial basis function and linear variants for remote sensing image classification. *Sensors*, 21(13), 4431

Silva, C. A., Hudak, A. T., Vierling, L. A., Loudermilk, E. L., O'Brien, J. J., Hiers, J. K., Jack, S. B., Gonzalez-Benecke, C., Lee, H., & Falkowski, M. J., 2016. Imputation of individual longleaf pine (Pinus palustris Mill.) tree attributes from field and LiDAR data. *Canadian Journal of Remote Sensing*, 42(5), 554–573.

Vega, C., Hamrouni, A., El Mokhtari, S., Morel, J., Bock, J., Renaud, J.-P., Bouvier, M., & Durrieu, S., 2014. PTrees: A point-based approach to forest tree extraction from lidar data. *International Journal of Applied Earth Observation and Geoinformation*, 33, 98–108.

Webber, Q.M.R. K.M. Ferraro, J.G. Hendrix, E. Vander Wal. 2022. What do caribou eat? A review of the literature on caribou diet. *Can. J. Zool.* 100:197-207

Węgrzyn, M. H., Fałowska, P., Kołodziejczyk, J., Alzayany, K., Wężyk, P., Zięba-Kulawik, K., Hawryło, P., Turowska, A., Grzesiak, B., & Lipnicki, L., 2021. Tree height as the main factor causing disappearance of the terricolous lichens in the lichen Scots pine forests. *Science of the Total Environment*, 771, 144834.

Węgrzyn, M. H., Kołodziejczyk, J., Fałowska, P., Wężyk, P., Zięba-Kulawik, K., Szostak, M., Turowska, A., Grzesiak, B., & Wietrzyk-Pełka, P., 2020. Influence of the environmental factors on the species composition of lichen Scots pine forests as a guide to maintain the community (Bory Tucholskie National Park, Poland). *Global Ecology and Conservation*, 22, e01017.

Wężyk, P., Hawryło, P., Szostak, M., Zięba-Kulawik, K., Winczek, M., Siedlarczyk, E., Kurzawiński, A., Rydzyk, J., Kmiecik, J., & Gilewski, W., 2019. Using LiDAR point clouds in determination of

the scots pine stands spatial structure meaning in the conservation of lichen communities in" Bory Tucholskie" National Park. Archiwum Fotogrametrii, Kartografii i Teledetekcji, 31, 85–103.

Yu, X., Hyyppä, J., Holopainen, M., & Vastaranta, M., 2010. Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes. *Remote Sensing*, 2(6), 1481–1495.

Zou, H., & Hastie, T., 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301–320.

Appendix A: guide to interpret 360° panoramic imagery for lichen information

1) Create a spreadsheet to Include the following parameters in the spreadsheet: The presence (1) or absence of lichen (0), the form of lichen, the colour of lichen, the photo quality, and where you see the lichen in the photo. It is good to make note of the specific habitat or weather in the photo set as well, ie. “bog land with snow coverage”

Table A1: Example of spreadsheet organization

Folder Name	Forest description: Springtime dense coniferous forest			
	Lichen? (0 for no lichen, 1 for lichen)	Type of lichen	Notes (what comments do you find important for your own, or others references?)	Image quality
File name				
File name				
File name				

2) Review the photos one by one to log the data needed.
3) Conduct quality control. For each folder of forest photos, pick 1 forest photo file at random and check to make sure the data that have been logged is actually consistent with the file in retrospect. If yes, move onto the next folder. If not, fix the data for the photo. Then, pick another two photos from the data set containing the erroneous photo data and check the quality of the data for those forest pictures. This is to ensure that there are no widespread inconsistencies.

TIPS:

Look **high and low!** There are epiphytic lichens (lichens that live on trees, including the trunks and branches), lithophytic lichens (those that grow on rocks) and those that sprawl on the ground. The ones that live and grow terrestrially like little bushes are the ones reindeer prefer to eat.

The **three types (body-plans) of lichen** are fruticose, foliose, and crustose. Crustose lichen forms plaque-like crusts on surfaces, while foliose lichens look like little leafs- some may even resemble ferns. Fruticose lichens are shrubby and take on a somewhat three-dimensional look and are the top delicious pick for reindeer.

Hanging lichens are classified as "fruticose", also known as shrubby lichen. They just grow so long, especially when out of reach of caribou, such that they lose that shrubby shape and look like green wigs!

It is fundamental to distinguish between lichen, fungi, snow, and moss. These can easily be confused for each other, especially at a distance.

(i) Mosses, like lichens, may cover large swathes of space. They are also a much **richer green**, when hydrated and alive, and yellow-brown on dead. Lichens typically come in softer "pastel" shades of green blue grey and green-yellow. Both can grow on rocks and trees and on the ground.

TABLE A2: DISTINGUISHING FEATURES BETWEEN LICHENS AND MOSSES

	Moss	Lichens
Where?	Rocks, live and dead wood, and on the ground	
Colour?	Vibrant greens, may take on yellow or brown colour if dead or dying. Never true blue.	Lighter, almost pastel shades of green, blue, grey, and green-yellow. Red and orange possible, but unlikely
Form?	More "2-D"	Some can remain close to their medium, ie. crust lichen, however others can grow into leaf or shrub like shapes (3-D)

(ii) There are **shelf fungi that look like white crustose lichen on trees yet they are not.** These two are especially easy to confuse from afar. Shelf fungi typically cluster together, with a more "triangular" or flat shape to each individual fruiting body. White crusting lichens on trees tend to take on a "rounder" shape, and may appear as standalone or evenly dispersed polka-dot.

Figure A1: White crustose lichen on bark of mature tree in Hearst forest, picture courtesy of Ontario Ministry of Natural Resources (Photo ID: VSN601161202110)

(iii) The **mycelium of certain fungi** can also look like lichen from long distances. If there is discolouration to the wood, but it looks like a varnish/a stain (lacking dimension), then it is unlikely to be lichen.

Figure 2A Blue staining on dead wood showing the mycelium of unidentified blue-fruiting fungi, photo taken in Mer Bleue, ON. Picture courtesy of Rachelle Li

(iv) Snow may settle into the bark of trees and cracks in the ground, which can resemble lichen. If it looks like white lines packed into the grooves of tree bark, then it probably is. When faced with a picture that contains heavy snow, even if YOU KNOW there is lichen there (but you cannot see it), do not put down what you do not see.

Special Note Regarding Hearst Forest: Lichens seem to be most prevalent in mature conifer (excluding larch forests) or mixed forests where they are usually found at lower heights. If you are unable to analyze a photo for whatever reason, specify the reason and DO NOT GUESS whether lichen is there or not. Only mark down what you feel sure of. Common reasons for undesirable image quality include brightness, water smudges, technical difficulties with the camera, and objects blocking your view. Be mindful of file names. Some files may be named out of numerical order, or written in strange ways, but do not change it in the spreadsheet. Following the exact names given to you ensures that the computer will be able to match your data with the proper forest photo