A Reproducible End-to-End Airborne LiDAR Workflow for Forest Structure Mapping

Zarrin Tasneem

Lassonde School of Engineering, York University, Toronto, Canada
zarrin@yorku.ca

Keywords: Airborne LiDAR, forest structure, canopy height model, Shannon entropy, canopy density, reproducible workflows

Abstract

Airborne LiDAR enables direct measurement of canopy height and three-dimensional forest structure, but many LiDAR-based studies remain difficult to reproduce due to ad hoc processing decisions and limited pipeline transparency (White et al., 2019).
This study presents a reproducible end-to-end workflow that transforms raw LAS point clouds into ecologically interpretable forest structure products. Using the Petawawa Research Forest (PRF) in Ontario as a case study (Natural Resources Canada, 2023; MacLean et al., 2019; Pickering, 2012), the pipeline integrates data ingestion and quality control, ground classification and height normalization using PMF and SMRF filters (Zhang et al., 2003; Pingel et al., 2013), canopy height model generation, canopy density and gap metrics, vertical structural complexity via Shannon entropy, and area-based plot- and stand-level summaries.
The workflow also supports rule-based species-group mapping and automated stand delineation for forest inventory applications. Sensitivity to key processing parameters is evaluated to support transparent and defensible long-term forest monitoring (Canadian Forest Service, 2005).

1. Introduction

Airborne LiDAR has become a foundational remote sensing technology for forest structure characterization (White et al., 2019). It enables direct measurement of canopy height, vertical distribution, and spatial heterogeneity at fine spatial resolution, supporting applications including forest inventory, habitat assessment, carbon accounting, and disturbance monitoring (Canadian Forest Service, 2005).

Despite widespread adoption, many LiDAR studies remain difficult to reproduce due to undocumented processing choices including ground classification parameters, smoothing, and rasterization workflows (White et al., 2019). Small parameter differences can propagate into large downstream differences in canopy metrics and ecological interpretation (Zhang et al., 2003; Pingel et al., 2013).

This study presents a reproducible end-to-end LiDAR workflow designed to transform raw point clouds into ecologically interpretable forest structure metrics and species-group products. The workflow is demonstrated using airborne LiDAR over the Petawawa Research Forest (PRF), a long-term experimental forest with extensive historical reference data (Natural Resources Canada, 2023; MacLean et al., 2019).

2. History of Petawawa Research Forest

The Petawawa Research Forest lies within the Ottawa River watershed on the unceded territory of the Algonquin Anishinaabe (University of Ottawa, n.d.). Established in 1918 by the Canadian Forest Service, PRF is one of Canada's longest-running experimental forest research sites (MacLean et al., 2019; Natural Resources Canada, 2023).

PRF has supported research in silviculture, forest genetics, fire ecology, and more recently remote sensing and geospatial monitoring. Its long-term plot network and historical management records make it an ideal site for reproducible forest analytics (Natural Resources Canada and Canadian Forest Service, 2022).

3. Materials and Methods

3.1 Study Area and Data Sources

The study was conducted within the Petawawa Research Forest (PRF), Ontario, Canada as shown in Figures 1, and 2, a long-term experimental forest characterized by mixed boreal forest stands, including coniferous, deciduous, and mixed wood compositions. The site provides a well-documented ecological baseline and extensive historical inventory data, making it suitable for evaluating reproducible remote sensing workflows.

Airborne LiDAR data were acquired over the PRF using a high-density laser scanning system, producing three-dimensional point clouds in LAS format (Natural Resources Canada, 2023; MacLean et al., 2019). The dataset includes multiple returns per pulse, enabling detailed characterization of vertical forest structure (White et al., 2019). Ancillary data, including base imagery, were used for contextual interpretation and validation of derived products .

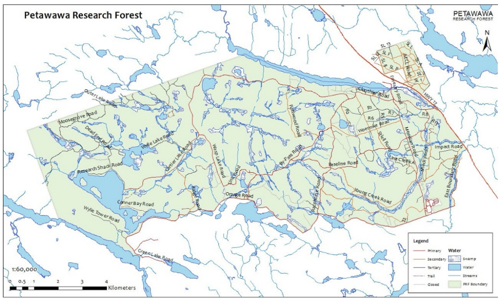

Figure 1. Boundary and road network of the Petawawa Research Forest, Ontario, Canada. Source: Natural Resources Canada, National Research Forests webpage. © *[2025]*, His Majesty the King in Right of Canada, as represented by the Minister of Natural Resources.

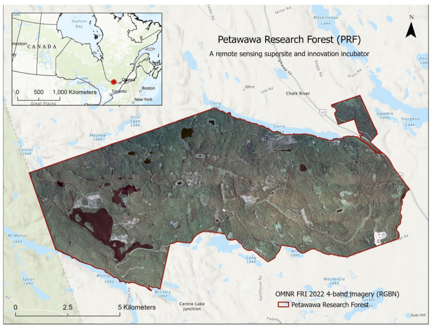

Figure 2. Location and extent of the Petawawa Research Forest on four-band forest inventory imagery. Source: National Forest Information System (NFIS) Open Data PRF web map. © *[2025]*, NFIS / contributing agencies.

3.2 Processing Architecture

The workflow was designed as a modular and reproducible pipeline that transforms raw LiDAR point clouds into ecologically interpretable forest structure metrics. As illustrated in Figure 3, the processing chain consists of four primary stages: (1) data ingestion and quality control, (2) ground classification and normalization, (3) raster and metric generation, and (4) spatial summarization.

To ensure reproducibility, each stage of the workflow was implemented using standardized parameter settings and sequential processing steps. The pipeline can be executed using a combination of open-source and commercial geospatial tools (e.g., PDAL, Python, ArcGIS/QGIS), with clearly defined inputs and outputs at each stage. This modular design allows individual components to be modified or extended without compromising the overall workflow integrity.

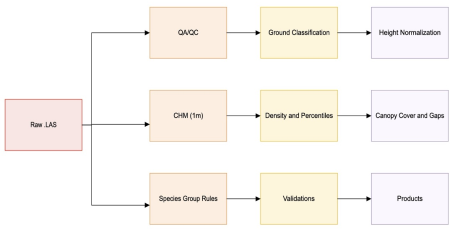

Figure 3. End-to-end LiDAR processing workflow.

3.3 QA/QC and Normalization

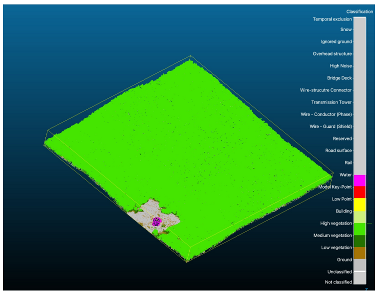

Figure 4. Classified point cloud used for QA/QC and normalization.

Ground classification was performed using PMF and SMRF filters (Zhang et al., 2003; Pingel et al., 2013). Height-above-ground (HAG) normalization was computed relative to the derived DTM.

3.4 Data Ingestion and Quality Control

Initial preprocessing involved ingestion of raw LAS files and verification of point cloud integrity. Quality control (QA/QC) procedures included checking for missing or overlapping tiles, coordinate system consistency, and spatial alignment across the dataset.

Particular attention was given to point density variability, which can arise from differences in flight line overlap, scan angle, and acquisition parameters. Spatial heterogeneity in point density can introduce bias in derived metrics, especially in raster-based products such as canopy height models (CHMs) and canopy cover estimates, where lower-density areas may under represent vegetation structure.

In addition to density variations, scan angle and flight line effects were evaluated. Off-nadir scan angles can result in geometric distortions and reduced penetration of laser pulses into dense canopy layers, potentially affecting both ground detection and vertical structure representation. Furthermore, edge effects between adjacent flight lines may introduce discontinuities in point distribution and intensity values, particularly in regions with limited overlap.

These artifacts can spread into downstream products if not identified and mitigated during the QA/QC stage. Basic intensity normalization considerations were also assessed, as LiDAR return intensity can vary due to sensor characteristics, range effects, and scan angle. Although intensity was not directly used as a primary metric in this study, inconsistencies in intensity values can serve as indicators of acquisition artifacts or calibration issues and were therefore visually inspected during quality control.

Noise filtering was applied to remove spurious returns, including isolated high and low points that are not representative of true terrain or vegetation structure. Visual inspection of the classified point cloud (Figure 4) was performed to verify the accuracy of ground and non-ground separation and to ensure that no systematic errors were introduced prior to normalization.

Establishing a rigorous QA/QC framework at this stage is important for ensuring the reliability of downstream analyses. Errors or inconsistencies in the input point cloud, such as density gaps, scan artifacts, or misclassified returns can spread through the workflow and significantly influence derived structural metrics.

By standardizing data ingestion and quality control procedures, the workflow ensures that subsequent processing steps are based on consistent and high-quality inputs, thereby supporting reproducibility and defensible forest structure mapping.

3.5 Ground Classification and Height Normalization

Ground classification was performed using two widely adopted filtering approaches: the Progressive Morphological Filter (PMF) and the Simple Morphological Filter (SMRF) (Pingel et al., 2013). Both methods are based on mathematical morphology, but differ in how they identify ground points and handle complex terrain conditions.

The PMF algorithm progressively increases the size of a structuring element (i.e., moving windows) to identify ground points by evaluating elevation differences between neighboring points (Pingel et al., 2013). At smaller window sizes, local variations are preserved, allowing detection of fine-scale terrain features. As the window size increases, larger non-ground objects such as trees and buildings are removed. This iterative approach is effective in relatively smooth terrain but can be sensitive to parameter selection, particularly in areas with abrupt elevation changes or dense vegetation cover.

In contrast, the SMRF algorithm introduces improvements in slope handling and edge preservation by incorporating slope-based thresholds and more adaptive filtering strategies (Pingel et al., 2013). SMRF is designed to better retain terrain discontinuities and reduce misclassification of ground points in complex environments. It is particularly effective in areas with steep slopes or heterogeneous canopy conditions, where traditional morphological filters may incorrectly classify low vegetation as ground or fail to capture sharp terrain features (Pingel et al., 2013).

Both PMF and SMRF were evaluated within the workflow to ensure accuracy of ground classification. Using multiple filtering approaches provides a means of assessing sensitivity to algorithm choice and helps mitigate systematic biases associated with any single method. This is especially important in forested environments such as the Petawawa Research Forest, where dense canopy cover can obscure ground returns, and variable terrain conditions (e.g., microtopography, slopes, and depressions) can complicate classification.

Following ground classification, a digital terrain model (DTM) was generated from the identified ground points using interpolation. Height-above-ground (HAG) normalization was then computed by subtracting the DTM elevation from each LiDAR return, resulting in a normalized point cloud where all elevations are referenced relative to the ground surface.

Accurate ground classification and normalization are important for reliable forest structure analysis. Errors in this step, such as misclassification of vegetation as ground or omission of true ground points can introduce systematic bias into canopy height estimates and propagate into all derived structural metrics. By incorporating both PMF and SMRF approaches and explicitly evaluating their behavior, the workflow enhances the reliability and reproducibility of the normalization process.

3.6 Canopy Height Model and Raster Derivation

A canopy height model (CHM) was generated from the normalized point cloud by extracting the maximum return height within each grid cell at a specified spatial resolution. The CHM provides a continuous raster representation of canopy height across the study area and serves as a foundational layer for subsequent structural analyses.

The selection of grid resolution represents an important tradeoff between spatial detail and noise. In this study, a fine-resolution grid (e.g., 1 m) was used to capture small-scale canopy variability and fine structural features such as gaps and edges. Finer resolutions preserve local heterogeneity but are more sensitive to noise and point density variability, whereas coarser resolutions (e.g., 5 m or greater) produce smoother surfaces at the expense of detail and may obscure ecologically relevant patterns. The chosen resolution reflects a balance between maintaining structural fidelity and ensuring computational efficiency.

The use of maximum return height per grid cell is a common approach for CHM generation, as it approximates the upper canopy surface and captures dominant vegetation height. Alternative approaches, such as percentile-based height metrics (e.g., 95th percentile), can reduce sensitivity to outliers and noise but may underestimate true canopy height in areas with sparse returns. In this workflow, the maximum return method was selected to preserve the full vertical extent of the canopy, particularly in structurally complex or uneven stands.

Additional raster products, including canopy cover and gap fraction, were derived using threshold-based methods applied to the normalized point cloud. Canopy cover was defined as the proportion of returns exceeding a specified height threshold (e.g., ¿2 m) within each grid cell, while gap fraction represents the proportion of returns below this threshold, indicating open or understory-dominated areas.

To mitigate high-frequency noise and local artifacts in the CHM, optional smoothing techniques (e.g., focal mean or Gaussian filtering) may be applied. However, excessive smoothing can blur sharp canopy features and reduce the ability to detect fine-scale structural variation. In this study, minimal or no smoothing was applied to retain high-resolution structural detail, with the understanding that downstream aggregation at larger spatial units can provide additional noise reduction.

Rasterization choices, including grid resolution, height extraction method, and smoothing have a direct influence on derived forest structure metrics. As such, these parameters were explicitly defined and standardized within the workflow to ensure reproducibility and to enable consistent comparisons across datasets and time periods.

3.7 Structural Metrics and Vertical Complexity

Canopy cover, gap fraction, and vertical structural complexity metrics were derived from normalized point clouds. Vertical structural complexity was quantified using Shannon entropy, which measures the distribution of LiDAR returns across the vertical strata:

$$H = -\sum p_i \log_2 p_i \tag{1}$$

where
H = Shannon entropy (dimensionless)
p_i = probability of LiDAR returns occurring within the ith vertical bin (dimensionless)

To compute entropy, the normalized point cloud was partitioned into a series of discrete vertical bins of equal height intervals spanning from ground level to the maximum canopy height within each analysis unit.

The frequency of LiDAR returns within each bin was normalized to form a probability distribution, which was then used to calculate entropy values. The choice of equal-height binning ensures consistency across the study area and facilitates comparison between different forest stands, although alternative binning strategies (e.g., adaptive or percentile-based bins) may be explored in future work.

Unlike traditional LiDAR-derived metrics, such as maximum height, mean height, or percentile-based measures, Shannon entropy captures the full vertical distribution of vegetation structure rather than focusing on a single summary statistic. This makes it particularly effective for distinguishing between structurally simple and complex forests. For example, two stands with similar canopy heights may exhibit very different entropy values if one consists of a single dominant layer while the other contains multiple vertical strata, including understory and mid-canopy vegetation.

From an ecological perspective, entropy provides a proxy for habitat complexity and structural diversity, which are key indicators of forest health, biodiversity, and successional stage. Higher entropy values are typically associated with multi-layered, heterogeneous canopies, often found in mature or mixed-species stands. In contrast, lower entropy values indicate more uniform vertical structure, such as even-aged plantations or recently disturbed areas.

This distinction is particularly important in mixed forest environments, where vertical heterogeneity cannot be adequately captured by height-based metrics alone. By incorporating entropy into the workflow, the analysis accounts for both vertical layering and the relative distribution of vegetation, providing a more comprehensive characterization of forest structure.

All structural metrics, including entropy, canopy cover, and height-based measures, were computed on a per-cell basis and subsequently aggregated to larger spatial units for analysis. The inclusion of entropy alongside conventional metrics imprtove the interpretability of LiDAR-derived products and supports more nuanced ecological assessments.

3.8 Area-Based Summarization and Stand-Level Metrics

To support ecological interpretation and forest inventory applications, point- and raster-based metrics were aggregated to plot- and stand-level units. Summary statistics such as mean canopy height, canopy cover, and entropy were computed within predefined spatial polygons.

These area-based approaches enable scaling from fine-resolution LiDAR measurements to management-relevant units, facilitating comparisons with existing forest inventory datasets and supporting operational decision-making.

3.9 Implementation and Reproducibility Framework

The LiDAR processing workflow was implemented as a series of modular and sequential processing steps, designed to ensure transparency, repeatability, and scalability. Each stage of the pipeline, from data ingestion to metric generation was executed using a combination of open-source and commercial geospatial tools, including PDAL for point cloud processing, Python for automation and scripting, and GIS platforms such as ArcGIS Pro, QGIS, QT Reader, and Cloud Compare for visualization and spatial analysis.

Raw LiDAR data were ingested in LAS format, which served as the primary input throughout the early stages of the workflow. Intermediate processing steps, including ground classification and normalization, were performed directly on point cloud data, while derived products such as canopy height models (CHMs), canopy cover, and entropy layers were exported as raster datasets (e.g., GeoTIFF format). This structured transition from point-based to raster-based representations enables efficient storage, analysis, and integration with other geospatial datasets.

To support reproducibility, the workflow was organized into discrete processing modules, each with clearly defined inputs, outputs, and parameter settings. Automation was achieved through Python-based scripting, allowing batch processing of multiple tiles and consistent application of processing parameters across the study area. All parameter values, such as filter window sizes, grid resolutions, and threshold values were explicitly defined and recorded to ensure that the workflow can be rerun with identical results.

In addition, the use of standardized file formats and processing sequences ensures compatibility across different software environments and facilitates reproducibility by other researchers. By structuring the workflow as a transparent and parameter-driven pipeline, this framework enables not only repeatable analysis within a single study, but also consistent application across multiple datasets and time periods, supporting long-term forest monitoring and comparative analysis.

3.10 Parameter Sensitivity and Reproducibility Considerations

The accuracy and interpretability of LiDAR-derived forest structure metrics are highly sensitive to key processing parameters, particularly those associated with ground classification, rasterization, and threshold-based metric derivation. In this workflow, parameter sensitivity was explicitly considered to ensure that results are both accurate and reproducible across different datasets and processing conditions.

One of the most influential parameters is the window size used in ground classification algorithms, particularly in PMF and SMRF filtering. Larger window sizes tend to produce smoother digital terrain models (DTMs) by generalizing local elevation variability, which can improve stability in flat or gently sloping terrain. However, excessive smoothing may obscure fine-scale microtopographic features and lead to misclassification of low vegetation as ground.

Conversely, smaller window sizes preserve local detail but may increase sensitivity to noise and result in fragmented or inconsistent terrain surfaces, especially in areas with dense canopy cover.

The choice of grid resolution in rasterization also has a significant impact on derived products such as canopy height models (CHMs) and canopy cover layers. Finer spatial resolutions (e.g., 1 m) capture detailed canopy heterogeneity and small-scale gaps but are more susceptible to noise and variations in point density. In contrast, coarser resolutions (e.g., 5 m) produce smoother and more generalized representations of forest structure, potentially masking ecologically meaningful variations and reducing the ability to detect fine-scale features.

Threshold-based parameters, including height thresholds used to define canopy cover and gap fraction, further influence structural metrics. For example, increasing the minimum height threshold for canopy cover (e.g., from 2 m to 5 m) will reduce the estimated canopy cover by excluding lower vegetation layers, while decreasing the threshold may include understory vegetation and inflate cover estimates. These choices directly affect ecological interpretation, particularly when distinguishing between canopy layers and understory structure.

The interaction between these parameters can produce compounded effects. For instance, a coarse grid resolution combined with a high canopy threshold may significantly underestimate structural complexity, whereas fine-resolution grids with lower thresholds may amplify noise and overestimate variability. As such, parameter selection must balance accuracy, ecological relevance, and computational efficiency.

To support reproducibility, all parameter values used in this study were explicitly defined and consistently applied across the workflow. While a full quantitative sensitivity analysis was beyond the scope of this study, qualitative evaluation of parameter effects highlights the importance of transparent parameterization in LiDAR processing pipelines. Future work may incorporate systematic sensitivity analyses or uncertainty quantification frameworks to further assess the influence of parameter choices on derived forest structure metrics.

4. Results

The proposed end-to-end LiDAR workflow successfully generated a suite of spatially explicit forest structure products, including canopy height models (CHMs), canopy cover, gap fraction, and vertical structural complexity layers. Visual inspection of the point cloud renderings (Figures 5–10) demonstrates that the preprocessing pipeline effectively preserved fine-scale structural detail while minimizing noise and classification artifacts.

Ground classification using PMF and SMRF filters produced a consistent digital terrain model (DTM), enabling reliable height-above-ground (HAG) normalization across the study area. The normalized point cloud exhibited clear vertical stratification of forest layers, which is important for downstream ecological interpretation.

The derived canopy height model captured expected spatial variability across the Petawawa Research Forest (PRF), with taller canopy structures corresponding to mature stands and lower heights associated with younger or disturbed areas. Canopy cover and gap fraction metrics further revealed heterogeneous horizontal structures, highlighting areas of dense canopy closure alongside localized canopy openings.

Vertical structural complexity, quantified using Shannon entropy, showed meaningful spatial gradients across the landscape. Higher entropy values corresponded to structurally complex, multi-layered stands, while lower values were associated with more homogeneous or even-aged forest conditions. These patterns are consistent with known forest dynamics and historical inventory data for PRF.

Importantly, the workflow produced spatially coherent and internally consistent outputs across all derived layers, indicating precision and accuracy of the processing pipeline. The reproducibility of the workflow ensures that identical inputs and parameter settings yield consistent results, supporting its applicability for long-term monitoring and repeatable forest analytics.

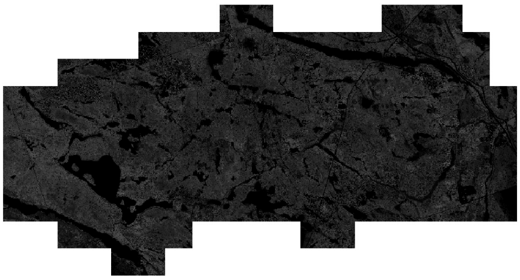

Figure 5. Point cloud visualization example.

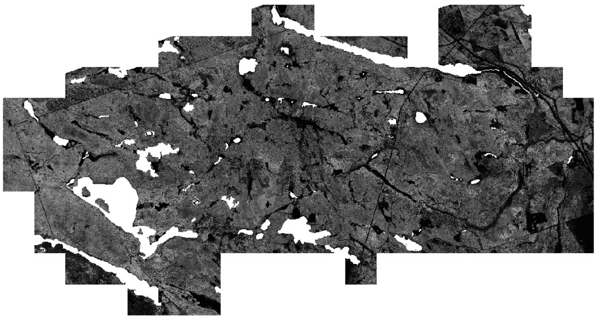

Figure 6. Point cloud visualization example.

Figure 7. Point cloud visualization example.

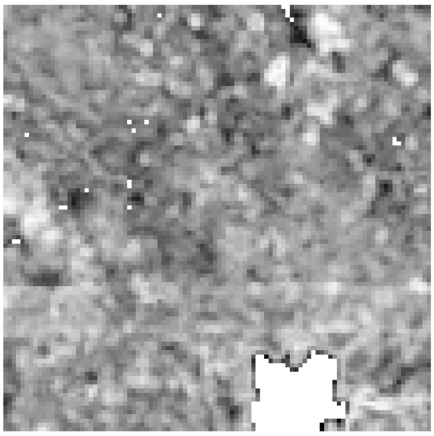

Figure 8. Point cloud visualization example.

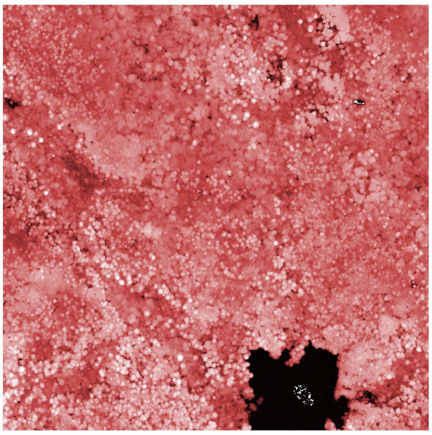

Figure 9. Point cloud visualization example.

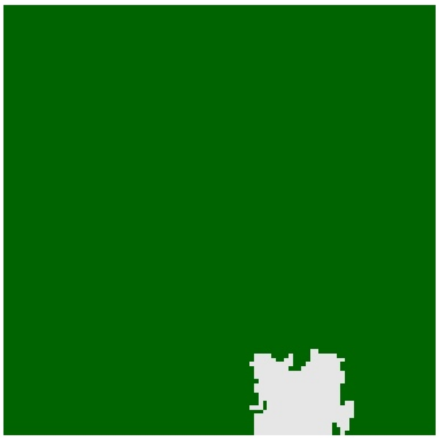

Figure 10. Point cloud visualization example.

The workflow produced spatially coherent canopy height, canopy cover, and vertical entropy layers consistent with historical PRF products.

5. Discussion

This study demonstrates that a fully reproducible airborne LiDAR workflow can generate ecologically meaningful forest structure metrics while maintaining transparency in processing decisions. By explicitly defining each step from ground classification to structural metric derivation, the workflow addresses an important gap in LiDAR-based forest studies, where undocumented parameter choices often limit reproducibility.

The results confirm that LiDAR-derived metrics such as canopy height, canopy cover, and vertical entropy are highly sensitive to preprocessing choices, particularly ground filtering and normalization. Even minor variations in filtering parameters can propagate into significant differences in derived products, reinforcing the importance of standardized and well-documented workflows.

The spatial patterns observed in canopy height and entropy align with ecological expectations for boreal forest systems, suggesting that the workflow preserves meaningful biophysical signals rather than introducing processing artifacts. The ability to capture both vertical and horizontal structural variability highlights the strength of LiDAR as a tool for comprehensive forest characterization.

However, several limitations remain. Cross-flight intensity variation and tile seam artifacts were observed, consistent with known challenges in airborne LiDAR processing. These effects can introduce discontinuities in derived metrics if not properly addressed. As noted in the paper, mitigation strategies such as overlap blending, radiometric normalization, and parameter harmonization are important for improving consistency across large datasets.

Another limitation is the dependence on rule-based approaches for certain classification steps, which may not generalize across diverse forest types or acquisition conditions. Future work could integrate machine learning or deep learning methods to enhance adaptability and improve classification accuracy, particularly for species-group mapping and stand delineation.

From an application perspective, the reproducible framework presented here provides a strong foundation for operational forest monitoring, including carbon accounting, habitat assessment, and disturbance detection. Its transparency and modularity also makes it well-suited for integration with emerging GeoAI pipelines and multi-sensor fusion approaches, which are increasingly important in large-scale environmental monitoring.

Overall, this work contributes to advancing LiDAR-based forest analytics by emphasizing reproducibility, parameter transparency, and ecological interpretability, positioning the workflow as a scalable solution for both research and operational forestry applications.

6. Future Research Direction

While this study establishes a reproducible and modular airborne LiDAR workflow for forest structure mapping, several avenues are there for advancing both methodological rigor and application scope.

A key direction is the integration of machine learning and deep learning approaches into the workflow. While the current framework depends on rule-based processing and physically interpretable metrics, data-driven models such as 3D convolutional neural networks, point-based architectures (e.g., PointNet++), and transformer-based models could enhance feature extraction and enable more reliable classification of forest structure and species groups. These approaches may also improve generalizability across different forest types and acquisition conditions.

Another important extension is the incorporation of multi-sensor data fusion. Combining airborne LiDAR with complementary datasets such as hyperspectral imagery, and thermal imagery can provide richer characterization of forest ecosystems. For example, LiDAR-derived structural metrics could be integrated with spectral information to improve species discrimination, biomass estimation, and health assessment, enabling more comprehensive ecological monitoring frameworks.

Future work should also focus on temporal analysis and 4D monitoring. Repeated LiDAR acquisitions offer the opportunity to quantify forest dynamics over time, including growth, disturbance, and recovery processes. Developing reproducible workflows for multi-temporal datasets will be important for detecting subtle structural changes and supporting long-term forest monitoring and climate-related studies.

From a methodological perspective, there is a need to explore parameter sensitivity and uncertainty quantification in greater depth. Although this study highlights the influence of processing parameters, systematic evaluation frameworks such as Monte Carlo simulations or Bayesian approaches could provide more formal characterization of uncertainty in LiDAR-derived metrics. This would improve confidence in downstream applications such as carbon accounting and habitat modeling.

Scalability is another important research direction. As LiDAR datasets continue to grow in size and coverage, implementing the workflow in cloud-native and high-performance computing (HPC) environments will be essential. Leveraging platforms such as Google Earth Engine, openEO, or distributed processing frameworks can enable efficient processing of regional to continental-scale datasets while maintaining reproducibility.

Finally, future research should explore the integration of this workflow into operational decision-support systems. Translating LiDAR-derived metrics into actionable insights for forest managers, policymakers, and conservation practitioners will require user-friendly tools, standardized outputs, and validation against field data. Bridging the gap between research-grade workflows and operational applications will be key to maximizing the impact of LiDAR-based forest analytics.

7. Conclusions

This study presents a reproducible, end-to-end airborne LiDAR workflow for forest structure mapping that transforms raw point cloud data into ecologically meaningful and spatially consistent products (Canadian Forest Service, 2005). By explicitly defining each stage of the processing pipeline, including data ingestion, quality control, ground classification, height normalization, and structural metric derivation, the workflow addresses a key limitation in many LiDAR-based studies: the lack of transparency and reproducibility in methodological choices.

The results demonstrate that the workflow can reliably generate canopy height, canopy cover, gap fraction, and vertical structural complexity metrics that are consistent with known forest conditions in the Petawawa Research Forest. The integration of Shannon entropy as a measure of vertical complexity further enhances the ecological interpretability of LiDAR-derived products, enabling more nuanced characterization of forest structure beyond traditional height-based metrics.

A key contribution of this work is the establishment of a standardized and modular processing framework that ensures consistent outputs across datasets and time periods. This reproducibility is essential for long-term forest monitoring, where comparability across acquisitions is important for detecting change, assessing disturbance, and supporting sustainable forest management.

Despite its strengths, the workflow remains sensitive to certain limitations, including cross-flight variability, tile seam artifacts, and parameter dependence in ground classification and normalization. Addressing these challenges through improved radiometric calibration, overlap blending, and adaptive parameterization will further importive the accuracy of LiDAR-based forest analytics.

Future work should focus on integrating machine learning approaches for automated classification and extending the framework to multi-sensor data fusion, including hyperspectral and radar datasets. Such integration would enable more comprehensive environmental monitoring and support the development of scalable, AI-driven forest intelligence systems.

This study demonstrates that reproducible LiDAR workflows are not only feasible but important for advancing forest structure mapping. By combining methodological transparency with ecologically meaningful outputs, the proposed framework provides a foundation for defensible, scalable, and future-ready forest monitoring applications.

Acknowledgements

The author gratefully acknowledges Leica Geosystems for its continued advancements in airborne LiDAR technology, which enable high-quality three-dimensional forest measurements. The author also acknowledges the American Society for Photogrammetry and Remote Sensing (ASPRS) for its sustained contributions to professional standards and the advancement of photogrammetry and remote sensing science.

References

Canadian Forest Service, 2005. Monitoring canada's national forest inventory. National Forest Information System (NFIS). Accessed 6 Jan 2026.

MacLean, D. A. et al., 2019. Petawawa Research Forest: 100 Years of Forest Research. *The Forestry Chronicle*, 95(4), —.

Natural Resources Canada, 2023. Petawawa Research Forest: A Century of Science, Innovation, and Collaboration. *Government of Canada Publications*. https://publications.gc.ca/collections/collection$_2023/rncan - nrcan/Fo147 - 2 - 2023 - 2 - eng.pdf$.

Natural Resources Canada, Canadian Forest Service, 2022. Petawawa research forest: A remote sensing supersite. National Forest Information System (NFIS) Open Data. Published 5 Oct 2022. Accessed 1 Jan 2025.

Pickering, L., 2012. The Petawawa Research Forest. *The Forestry Chronicle*, 88(1), 11–12.

Pingel, T. J., Clarke, K. C., McBride, W. A., 2013. An Improved Simple Morphological Filter for the Terrain Classification of Airborne LiDAR Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 77, 21–30.

University of Ottawa, n.d. A journey through indigenous presence on campus. University of Ottawa News. Accessed 6 Jan 2026.

White, J. C., Chen, H., Woods, M. E., Low, B., Nasonova, S., 2019. The Petawawa Research Forest: Establishment of a Remote Sensing Supersite. *The Forestry Chronicle*, 95(3), 149–156.

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., Zhang, C., 2003. A Progressive Morphological Filter for Removing Nonground Measurements from Airborne LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4), 872–882.