



FACULTY OF ENGINEERING AND SUSTAINABLE DEVELOPMENT
Department of Computer and Geospatial Sciences

Effects of Canopy Height Model Resolution on Individual Tree Segmentation

Implications for Identifying Scots Pine Trees with Old-
Growth Structural Characteristics

Alexandra Bärems
Clara Spik

2026

Master's Thesis, 30 Credits,
for the Degree of Master of Science in Engineering in Land Surveying and Geographic
Information Systems

Supervisor: Fredrik Zetterquist
Examiner: Julia Åhlen

Abstract

Pine trees with older structural characteristics are important for biodiversity because they provide valuable habitats for many species. These trees can indicate high conservation values and are therefore important to identify in forest management and nature conservation. However, identifying such trees through field inventory is time-consuming, particularly across large forest areas, which creates a need for efficient methods for large-scale mapping. The aim of this study is to investigate how spatial resolution in canopy height models (CHMs) derived from airborne Light Detection and Ranging (LiDAR) data affects rule-based segmentation of individual tree crowns and the identification of pine trees with older structural characteristics.

The study was conducted by segmenting individual tree crowns from two CHM datasets with different spatial resolutions: 0,5 m and 1 m. For each segment, structural variables including crown diameter, height growth and height were calculated. To account for variations in ecological conditions, the variables were analyzed within different strata based on elevation, soil moisture and topographic position. The crown segments were then classified according to the number of defined structural characteristics fulfilled, and the results were subsequently validated using field data.

The segmentation results showed clear differences between the two resolutions. The high-resolution CHM data improved the delineation of individual tree crowns, while the low-resolution CHM showed limitations mainly related to under-segmentation. The crown segmentation validation showed that the higher resolution reduced under-segmentation and improved the separation between neighboring trees, providing better conditions for the analysis of structural characteristics of individual tree crowns.

The classification showed that differences in segmentation also affected the identification of candidate trees. The high-resolution dataset identified more candidates fulfilling all classification criteria, whereas the low-resolution dataset produced a larger proportion of candidates overall. The results indicated greater potential for identifying forest areas with a high probability of containing structurally older Scots pine trees than for the precise identification of individual trees. The ecological stratification also contributed to the analysis of structural characteristics in relation to local environmental conditions. The combination of CHM data, rule-based analysis and ecological stratification shows potential to support the prioritization of field inventories and the identification of areas with potential conservation values in large-scale forest planning.

Keywords: *canopy height model (CHM), remote sensing, segmentation, Scots pine, old structural characteristics, conservation planning*

Sammanfattning

Tallar med äldre strukturer har stor betydelse för biologisk mångfald eftersom de utgör livsmiljöer som är viktiga för många arter. Dessa träd kan indikera höga naturvärden och är därför viktiga att identifiera inför exempelvis skogsplanering och naturvårdsåtgärder. Att identifiera sådana träd i fält är dock tidskrävande, särskilt i stora skogsområden, vilket skapar ett behov av effektiva metoder för storskaliga kartläggningar. Syftet med denna studie är att undersöka hur spatial upplösning i trädhöjdsraster (CHM) baserade på flygburen LiDAR-data påverkar regelbaserad segmentering av individuella trädkronor samt identifiering av tallar med äldre strukturella egenskaper.

Studien genomfördes genom segmentering av individuella trädkronor från CHM-data med två olika upplösningar: 0,5 m och 1 m. För varje segment beräknades strukturella variabler såsom krondiameter, tillväxt och höjd. För att ta hänsyn till variationer i ekologiska förhållanden analyserades variablerna inom olika strata baserade på höjd över havet, markfuktighet och topografiskt läge. Trädkronorna klassificerades därefter utifrån hur många av de definierade strukturella egenskaperna som uppfylldes och sedan validerades resultaten med hjälp av fältdata.

Segmenteringsresultatet visade tydliga skillnader mellan de två upplösningarna. Det högupplösta CHM-datasetet förbättrade avgränsningen av individuella trädkronor, medan det lågupplösta datasetet främst begränsades av en högre grad av undersegmentering. Valideringen av kronsegmenteringen visade att den högre upplösningen minskade undersegmentering och förbättrade separationen mellan närliggande träd, vilket gav bättre förutsättningar för analys av individuella trädkronors strukturella egenskaper.

Klassificeringen visade att skillnader i segmentering även påverkade identifieringen av kandidatträd. Det högupplösta datasetet resulterade i fler kandidater som uppfyllde samtliga kriterier, medan det lågupplösta datasetet gav en större andel kandidater totalt. Resultaten visade större potential för att identifiera skogsområden med hög sannolikhet för strukturellt äldre tallar, snarare än för exakt identifiering av enskilda träd. Den ekologiska stratifieringen bidrog samtidigt till att strukturella egenskaper kunde analyseras i relation till lokala miljöförhållanden. Kombinationen av CHM-data, regelbaserad analys och ekologisk stratifiering visar därmed potential för att stödja prioritering av fältinventeringar och identifiering av områden med potentiella naturvärden inom storskalig skogplanering.

Nyckelord: *Trädhöjdsraster, CHM, fjärranalys, segmentering, tall, äldre strukturella egenskaper, naturvård*

Preface

This thesis represents the culmination of our five years of studies for the Degree of Master of Science in Engineering in Land Surveying and Geographical Information Systems at the University of Gävle. The work was conducted during the spring of 2026 and marks the final stage of our academic journey.

Five years ago, we began this journey with limited knowledge and experience in many of the areas we have since explored and learned throughout our education. For this, we are deeply grateful to the many inspiring teachers and lecturers who have supported and guided us over the years.

We would also like to express our sincere gratitude to Skogforsk for initiating this thesis project and for providing valuable guidance and engaging discussions during the process. Their expertise and continuous support have been highly appreciated during the development of this work.

We are equally grateful to Billerud for contributing valuable expertise, participation in project meetings and support during the field inventory. Their practical insights and engagement have provided important perspectives for the study.

The opportunity to conduct this thesis in collaboration with Skogforsk and Billerud has provided valuable insights into the practical applications of remote sensing and forest analysis.

Finally, we would like to extend our sincere thanks to our supervisor, Fredrik Zetterquist, our program director at the University of Gävle, for invaluable guidance, encouragement and support during this thesis and throughout our years of study.

University of Gävle, May 2026

Alexandra Bärema

Clara Spik

Table of contents

Abstract	i
Sammanfattning	iii
Preface	v
1 Introduction	1
1.1 Purpose and research questions	2
1.2 Delimitations	3
1.3 Thesis outline	3
2 Theoretical background	4
2.1 Ecological value of old Scots pine trees	4
2.1.1 Older Scots pine trees and their structural attributes	4
2.2 Airborne LiDAR methods	6
2.3 CHM for individual tree analysis	7
2.4 Spatial resolution	8
2.5 Machine learning approaches in individual tree analysis	9
2.6 Individual tree detection	9
2.6.1 Segmentation	10
2.6.2 Watershed segmentation	10
2.6.3 Marker-controlled watershed segmentation	12
2.7 Topographic position index	12
2.8 Stratified identification of structural patterns	13
2.9 Model evaluation and performance metrics	14
2.9.1 Performance metrics	14
3 Method	16
3.1 Study area	16
3.2 Data and materials	17
3.3 Data preprocessing	19
3.4 Crown segmentation	20
3.4.1 Justification of preprocessing and segmentation settings	21
3.5 Classification of Scots pine trees with old-growth structural characteristics	27
3.5.1 Strata	29
3.5.2 Structural variables used for classification	32
3.6 Validation methodology	38
4 Results	39
4.1 Watershed-segmentation results	39
4.2 Classification results	41
4.2.1 Classification outcome	41
4.2.2 Model selectivity	44
4.2.3 Characteristics of classified trees	45
4.2.4 Influence of height growth filtering and environmental stratification	45
4.3 Validation of the results	47
5 Discussion	53
5.1 Main findings and interpretation	53

5.2	Methodological discussion	55
5.2.1	Parameter sensitivity and preprocessing	55
5.2.2	Influence of spatial variation and resolution	55
5.2.3	Uncertainty in height growth estimation	56
5.2.4	Classification approach and method selection.....	58
5.2.5	Influence of stratification and environmental variables	59
5.3	Limitations	60
5.3.1	Data and practical limitations	60
5.3.2	Methodological limitations	62
5.4	Suggestions for future research	63
5.5	Environmental and sustainability aspects.....	64
6	Conclusions.....	65
	References	67
	Appendix A – Summary statistics for tested height thresholds	A1
	Appendix B – Sensitivity analysis of thinning distance	B1
	Appendix C – Segmented crowns per tile.....	C1
	Appendix D – Percentile thresholds per strata for the high-resolution dataset	D1
	Appendix E – Percentile thresholds per strata for the low-resolution dataset	E1
	Appendix F – Distribution of candidate trees and strata area statistics (high-resolution) ..	F1
	Appendix G – Distribution of candidate trees and strata area statistics (low-resolution) ..	G1
	Appendix H – Field site photographs	H1
	Appendix I – Code availability	I1

1 Introduction

Biodiversity in forest landscapes has a strong connection to structures that develop over long periods of time (Lindenmayer et al., 2012; Siitonen, 2001). Old Scots pine trees are an example of such long developed forest structures, and they provide an important habitat for many different species (Edman et al., 2024). However, in today's managed forests, these types of structures are unusual as rotation periods are shorter than the time required for ecological processes to develop old-growth tree structures (Edman et al., 2024; Buness et al., 2025). These trees are therefore often used as indicators of high conservation values in nature conservation and forest planning, where information about their occurrence can support decisions related to conservation and forest management (Przybylski et al., 2021).

Nature conservation trees are characterized by different ecological and structural attributes, such as old age, slow growth, large or irregular crowns, cavities and the presence of species associated with high conservation values (Skogsstyrelsen, 2020). However, several of these characteristics require field observations and cannot be directly detected from canopy height model (CHM) data. Therefore, this study focuses on remotely detectable structural characteristics, including tree height, crown size and height growth, which may indicate old Scots pine trees with potential conservation values.

These structural characteristics are influenced not only by tree age but also by environmental conditions affecting tree growth. Therefore, methods that account for variation in site conditions are needed when interpreting CHM-derived structural indicators.

Field inventories of old Scots pine trees are time- and resource-consuming, particularly because these trees are relatively uncommon in managed forests (Edman et al., 2024). To facilitate forest planning and structural analysis of trees, remote sensing-based methods have become increasingly important (Vauhkonen et al., 2012). To conserve and manage old Scots pine trees, methods are needed that enable identification of trees at the individual level in the landscape (Vauhkonen et al., 2012). This in turn requires spatial information with sufficient spatial resolution to distinguish individual pines.

When analyzing trees at the individual level, height data in the form of a CHM is commonly used to represent the height of the vegetation above the ground. CHM data are an important basis for individual-based tree analysis where the spatial resolution plays a crucial role in how well the crown structures of the trees can be distinguished (Vauhkonen et al., 2012; Zhao et al., 2023). In this context, there is also a need to analyze how variations in the spatial resolution of CHM affect segmentation and the identification of structurally older Scots pine trees.

Rule-based segmentation methods, such as the watershed algorithm, are commonly used to identify individual trees in CHM data (Lindberg & Holmgren, 2017; Vauhkonen et al., 2012). The watershed algorithm is well suited for CHM data as it relies on elevation variations and local maxima to delineate individual tree crowns. However, the performance of crown segmentation in CHM data depends largely on the quality of the input data, one particularly important factor is the spatial resolution of the CHM raster. If the resolution is low, several trees risk being merged into one object, while a high resolution can cause a tree crown to be divided into several segments (Lindberg & Holmgren, 2017). Consequently, the choice of resolution can affect both over- and under-segmentation.

In recent years, machine learning methods, including Convolutional Neural Networks (CNNs), have become increasingly used in vegetation and forest remote sensing, including individual tree detection and crown segmentation (Kattenborn et al., 2021; Zhao et al., 2023). These methods can learn spatial patterns from image and raster data, and previous studies have shown that LiDAR-derived height information can be used for CNN-based individual tree analysis (Kattenborn et al., 2021; Hamraz et al., 2019). However, supervised machine learning approaches generally require representative training datasets and can be sensitive to variations in data quality and forest conditions, which can limit their applicability and transferability (Zhu et al., 2017; Kattenborn et al., 2021). Therefore, rule-based approaches remain relevant when the aim is to develop transparent methods based on interpretable structural variables.

This study was initiated in collaboration with Skogforsk and Billerud to investigate how spatial resolution in LiDAR-derived CHM data affects individual tree segmentation and whether such data can support the identification of Scots pine trees with old-growth structural characteristics.

1.1 Purpose and research questions

This study investigates how spatial resolution in CHM rasters, used as the sole data source, affects rule-based segmentation of individual trees. It also examines how the segmentation results influence the identification of Scots pine trees with old-growth structural characteristics. Finally, the study has an applied purpose, to assess whether the results can be used as decision support in forest planning. To achieve this, the following research questions are considered:

- How does spatial resolution influence the outcomes of rule-based segmentation of individual trees in CHM raster?

- How do segmentation results influence the ability to identify Scots pine trees with old-growth structural characteristics using a stratified approach that accounts for varying ecological conditions, and what potential do these results have as decision support in forest planning?

1.2 Delimitations

The study is limited to CHM rasters derived from laser scanning (Light Detection and Ranging, LiDAR) data for segmentation. The data used in this study consist of two CHM rasters with spatial resolutions of 1×1 meter and 0.5×0.5 meter, acquired in 2022 and 2024, respectively. Additional datasets, including a soil moisture map and a digital elevation model (DEM), are used to stratify trees based on environmental conditions, while UAV orthomosaic data and field inventory data are used for validation. The segmentation itself is based solely on CHM-derived variables.

The study area is delineated as a boreal, coniferous-dominated forest area in the Idre-Särna region. Therefore, the findings are primarily applicable to forests with similar structural conditions.

1.3 Thesis outline

This thesis is structured as follows. Chapter 2 presents the theoretical background and previous research related to remote sensing, CHM, tree segmentation and ecological structures in older Scots pine. Chapter 3 describes the study area, data and methods used in the study, including the justification for the selected parameters and methodological choices. Chapter 4 presents the results. Chapter 5 includes discussion, methodological considerations and suggestions for future work. Finally, Chapter 6 presents the conclusion.

2 Theoretical background

This chapter presents the theoretical background relevant to the study, including ecological aspects of old Scots pine trees and remote sensing methods used for their identification.

2.1 Ecological value of old Scots pine trees

A large part of the forest biodiversity depends on habitats consisting of living trees with cavities and other structural microhabitats that primarily develop in large and old trees (Edman et al., 2024; Lindenmayer et al., 2012; Skogsstyrelsen, 2025). In many forests these structures are uncommon because forestry and fragmentation affect the age structure and variation of the forest, leading to reduced occurrence of large and old trees and thus valuable forest environments (Edman et al., 2024; Linder & Östlund, 1998; Przybylski et al., 2021).

Buness et al. (2025) report that the rotation period in managed boreal forests in Sweden is approximately 100 years, referring to the period from stand establishment to harvest. Edman et al. (2024) also states that the rotation period in managed forests is shorter than the time required for these structures to develop. As a result, old stands of around 250 years are uncommon despite their importance for biodiversity (Buness et al., 2025), especially in boreal forests where many forest-dwelling species depend on long forest continuity and structures that develop over long periods of time (Siitonen, 2001).

However, field inventory of old Scots pine trees over large and remote areas is both time- and resource-consuming, and may require many sample plots to achieve reliable results (Bergen et al., 2009; Vierling et al., 2008). To support forest planning, nature conservation and decision-making, there is therefore a need for methods that enable a more large-scale and systematic mapping of such structures (Bergen et al., 2009; Puliti et al., 2020; Vierling et al., 2008).

Trees with characteristics associated with old age can be identified indirectly through structural attributes such as tree height and structural complexity, which can be derived from remote sensing data (Bergen et al., 2009; Kattenborn et al., 2021). Bergen et al. (2009) also mention that laser scanning has great potential and is increasingly being used to measure these characteristics over large areas in ecological analyses linked to biodiversity.

2.1.1 Older Scots pine trees and their structural attributes

Scots pine (*Pinus sylvestris* L) develops distinctive structural characteristics over time that can be used to distinguish older individuals from younger trees (Handegård et al., 2021). The study by Handegård et al. (2021) shows that older pine trees can often be

recognized by rough and flattened bark, drooping branches, irregular trunk shape and reduced height growth. Older pines also have broader and flatter crowns, which are linked to altered growth patterns characterized by reduced height growth and increased horizontal branch spread in aging trees (Oliver & Larson, 1996).

Although height growth decreases with age, total tree height remains an important structural feature, as it affects dominance and the vertical structure of the stand (Oliver & Larson, 1996). These structural characteristics develop through slow growth over a long period of time and are typical of older pine trees (Handegård et al., 2021). According to the Swedish Forest Agency (Skogsstyrelsen, 2020), forests with high conservation values often contain old trees with structural features that are important for biodiversity. Nature conservation trees are commonly identified based on several ecological and structural characteristics, including age-related features such as crown structure, slow growth, cavities, bark characteristics and associated species (Skogsstyrelsen, 2020). However, several of these indicators require field observations and cannot be directly identified from CHM data.

The development of these structural characteristics is influenced not only by tree age, but also site conditions that affect tree growth and forest development (Pretzsch et al., 2022; Spies, 2004). As a result, similar structural values may represent different ecological conditions depending on the surrounding environment, making it important to interpret these variables in relation to local site conditions rather than as absolute indicators of age. In this study, this is addressed through a stratified approach, in which trees are analyzed within environmentally comparable groups.

Structures developed in older pine trees are of great importance for biodiversity. Coarse bark, dead branches and cavities (Figure 1) create varying microhabitats that support a variety of species, including insects, lichens and fungi (Skogsstyrelsen, 2020). Research shows that many of these species are dependent on long-term continuity, because the structures that they rely on only develop in trees that have remained standing for long periods of time (Siitonen, 2001).

However, several characteristics associated with high conservation values, such as cavities and bark structures (Skogsstyrelsen, 2020), require field observations and cannot be directly detected from CHM data. Therefore, remote sensing analyses rely on structural characteristics derived from available spatial data.



Figure 1. Example of a Scots pine observed during the field inventory, showing structural characteristics associated with high conservation values, including dead branches and an altered crown structure.

Among the structural characteristics that can be derived from CHM data, attributes such as tree height, height growth and crown shape can be used as indirect indicators of older trees. These variables capture different aspects of tree development, including size, growth dynamics and crown architecture, which tend to change systematically with age. They are therefore relevant for identifying trees with structural characteristics associated with older individuals and form the basis for the variables used in this study.

2.2 Airborne LiDAR methods

LiDAR is an active remote sensing technique used to acquire three-dimensional information about the ground surface and overlying objects by measuring the time it takes for a laser pulse to be emitted, reflected by an object and returned to the sensor (Lefsky et al., 2002; Vosselman & Maas, 2010). Lefsky et al. (2002) describe how the measured distances are combined with information about the sensor's position and orientation, making it possible to calculate the three-dimensional coordinates of each laser return and create a georeferenced point cloud.

A laser pulse can be recorded as one or more returns depending on which objects it interacts with, for example vegetation at different height levels and the ground surface (Lefsky et al., 2002; Wehr & Lohr, 1999). This makes it possible to describe the vertical structure of vegetation (Dubayah et al., 2020; Lefsky et al., 2002). This is illustrated in Figure 2, which shows how a single laser pulse can generate multiple returns from vegetation at different height levels and from the ground surface.

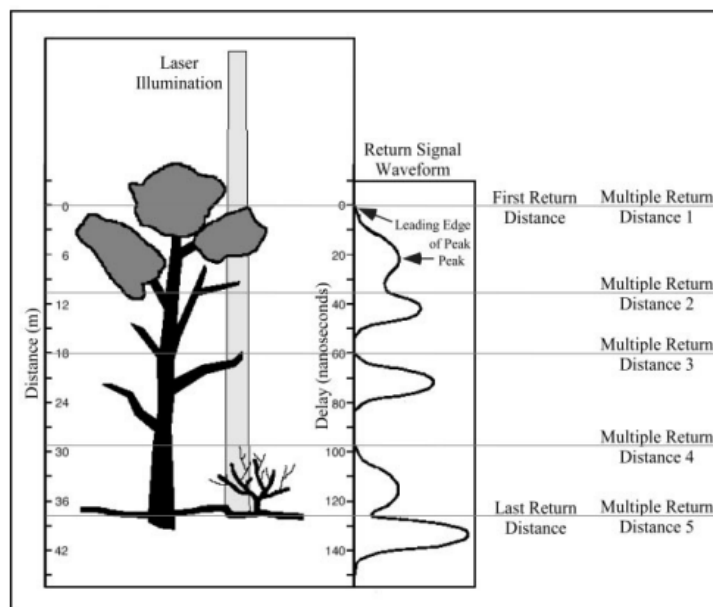


Figure 2. Schematic illustration of how a laser pulse can give rise to multiple returns from different height levels in the vegetation, from tree canopy to ground surface. The figure shows the principle behind first return, multiple returns and last return. Source: Lefsky et al. (2002).

Vosselman and Maas (2010) show that LiDAR data acquisition can be performed from different platforms, with airborne and terrestrial LiDAR being two main categories. In forest remote sensing airborne LiDAR (ALS) is widely used, traditionally from manned aircraft, while drone-based LiDAR (UAV LiDAR) is increasingly used to enable high-resolution analyses and more flexible data collection, especially in studies of smaller forest areas (Puliti et al., 2020; Wallace et al., 2012).

Airborne LiDAR from manned aircraft is commonly used for data collection over large geographical areas and is applied in forest inventories to estimate stand characteristics such as height and volume using measurements derived from LiDAR point clouds (Koch et al., 2006; Næsset, 2002). UAV LiDAR, on the other hand, makes it possible to collect data at lower flight altitudes, resulting in higher point density and a more detailed representation of the vegetation structure (Puliti et al., 2020; Wallace et al., 2012). This difference in data characteristics can influence the level of detail that can be captured in subsequent analyses.

Regardless of the platform used, LiDAR point clouds are commonly used as a basis for creating various elevation models that represent the height and vertical structure of the vegetation (Lefsky et al., 2002; Koch et al., 2006). These models are commonly used as a basis for further analysis of forest structure, such as the derivation of CHM, which are widely used for individual tree detection and segmentation in forest remote sensing (Popescu & Wynne, 2004; Vauhkonen et al., 2012).

2.3 CHM for individual tree analysis

Canopy Height Model (CHM) is a raster-based height model that represents vegetation height above the ground and is used to analyze forest vertical structure in forest remote sensing (Koch et al., 2006; Mielcarek et al., 2018). CHM rasters are usually created by subtracting a digital terrain model (DTM) from a digital surface model (DSM), resulting in a raster where each pixel represents the canopy height above the ground surface (Koch et al., 2006; Zhang et al., 2020).

According to Yu et al. (2006), CHM is used in forest remote sensing as a central tool for analyzing forest vertical structure. Airborne LiDAR generates a georeferenced point cloud from which DTM, DSM and CHM are calculated. Because CHM clearly highlights height variations in trees and tree crowns, the model is commonly used as a basis for identifying and delineating individual trees in forest analyses (Popescu & Wynne, 2004; Yu et al., 2006). This makes CHM particularly suitable for analyses where tree-level structural information is required.

When analyzing trees at the individual level, height data in the form of CHM is particularly useful, as tree height and crown structure are important characteristics for distinguishing individual trees (Vauhkonen et al., 2012). Local height maxima in a CHM raster can indicate treetops, and variations in height around these maxima can be linked to crown shape and extent. This enables the analysis of individual trees even without spectral information (Popescu & Wynne, 2004; Vauhkonen et al., 2012).

Because CHM represents vegetation height as a continuous raster surface, segmentation is required to delineate individual trees (Vauhkonen et al., 2012). The study by Yu et al. (2006) shows that CHM-based analyses can be affected by factors such as forest density, different tree species and spatial resolution, which can make the identification of individual trees more challenging. This means that the accuracy of individual tree detection is not only dependent on the method used, but also on data characteristics such as spatial resolution and the forest structure itself.

2.4 Spatial resolution

Spatial resolution describes the size of each pixel in a raster and indicates the smallest spatial detail that can be distinguished in the data (Lindberg & Holmgren, 2017). In CHM-based analyses, the spatial resolution of the input data may influence the accuracy of individual tree detection and crown delineation (Lucas et al., 2024). With sufficiently high spatial resolution, variations in crown structures can be distinguished more clearly, facilitating the segmentation of individual trees (Vauhkonen et al., 2012; Koch et al., 2006).

In contrast, at lower spatial resolution, height variations may be generalized over larger areas, which can contribute to several individual trees being merged into a single segment (Lucas et al., 2024; Koch et al., 2006). This illustrates that although CHM-based methods are suitable for individual-based analysis, the quality of the results is strongly influenced by data characteristics such as spatial resolution (Lucas et al., 2024).

The choice of spatial resolution therefore affects how well individual trees can be identified and delineated during segmentation. This makes spatial resolution an important factor to consider when CHM-based results are used as decision support in forest planning, since limitations in data quality can influence both the accuracy of the results and their practical applicability (Lindberg & Holmgren, 2017; Lucas et al., 2024). This highlights the importance of evaluating how different spatial resolutions influence segmentation outcomes, which is a central focus of this study.

2.5 Machine learning approaches in individual tree analysis

Machine learning methods, particularly Convolutional Neural Networks (CNNs), are increasingly used in vegetation and forest remote sensing for analysis of image and raster data (Kattenborn et al., 2021). CNN-based methods can identify spatial patterns by analyzing relationships between neighboring pixels, making them suitable for detecting structures and patterns in spatial datasets (Goodfellow et al., 2016; LeCun et al., 2015; Rawat & Wang, 2017).

In forest remote sensing, CNN-based approaches have been applied for individual tree detection and crown segmentation using LiDAR-derived height information (Hamraz et al., 2019; Zhao et al., 2023). Since CNN models can learn spatial patterns from raster data, they have shown potential for delineating complex tree crown structures (Kattenborn et al., 2021; Zhao et al., 2023).

However, supervised CNN-based approaches require representative training data with known reference information. The performance and transferability of these methods can therefore be limited by the availability of training data as well as variations in data quality, spatial resolution and forest conditions (Zhu et al., 2017; Kattenborn et al., 2021). When suitable training data are limited, rule-based approaches provide an alternative where tree detection, segmentation and subsequent classification are based on predefined and interpretable criteria.

2.6 Individual tree detection

Individual Tree Detection (ITD) refers to methods used to identify individual trees in remote sensing data (Hyypä et al., 2008). In CHM-based analyses, local maxima are commonly used to locate potential treetops (Hyypä et al., 2008). These treetop candidates can then be used as starting points for delineating crown extent using different segmentation methods. In this way, tree-level information such as tree height and crown size can be obtained (Dalponte & Coomes, 2016; Hyypä et al., 2008).

A commonly used method for identifying treetops in a CHM is Variable Window Filtering (VWF), which searches for local maxima within a moving search window. If a pixel represents the local maximum within its search window, it can be identified as a potential treetop (Chen et al., 2006; Popescu & Wynne, 2004).

The size of the moving search window is adjusted in VWF according to tree height, commonly based on the relationship between tree height and crown size (Popescu & Wynne, 2004). Because this relationship can vary between different forest types and stand structures, the parameters for VWF often need to be adjusted to the specific study area (Popescu & Wynne, 2004; Kaartinen et al., 2012). This is important because an inappropriate window size may lead to both over- and under-detection of treetops (Chen et al., 2006; Popescu & Wynne, 2004). VWF is therefore relevant for this study, where reliable treetop detection is required as input for the subsequent crown segmentation.

2.6.1 Segmentation

Segmentation means dividing an image into homogeneous regions with similar properties, such as intensity or structure, with the aim of grouping pixels into meaningful units (Gonzalez & Woods, 2018; Pal & Pal, 1993). In remote sensing, segmentation is a common step used to enable the identification and delineation of objects in spatial data, where groups of pixels are treated as contiguous objects rather than individual pixels (Blaschke et al., 2014; Pare et al., 2021).

In forest applications, segmentation is used to identify individual trees and tree crowns in LiDAR-based height data, where each segment can correspond to a single tree (Lefsky et al., 2002). In CHM-based individual tree analysis, treetop candidates identified from local maxima can be used to guide segmentation and delineate crown extent. Segmentation can be performed using various methods that are based on different principles for how objects are delineated in image and elevation data (Blaschke et al., 2014).

In individual based analysis of forests using CHM, segmentation is linked to several practical challenges (Lucas et al., 2024; Vauhkonen et al., 2012). The ability to identify and delineate individual trees is affected by the forest structural complexity, such as variations in crown size and tree density. In dense stands, trees may have overlapping crowns, which can make correct separation difficult even when relatively high-quality height data is used (Lucas et al., 2024).

2.6.2 Watershed segmentation

The watershed algorithm is an established method for image segmentation in remote sensing (Kakhani et al., 2019). The method is based on a topographic principle in which the image is interpreted as a topographic landscape where pixel values represent elevation differences. Local extreme points in the data are identified based on predefined criteria and used as starting points for the segmentation, allowing regions to grow until they meet and form segment boundaries along ridges in the data (Meyer & Beucher, 1990; Najman & Schmitt, 1996).

Local extreme points can be identified using different approaches. In forest remote sensing, a common approach is to apply moving windows whose size is adjusted according to tree height, so-called variable window filters (VWF), to estimate treetops in canopy height models (Lindberg & Holmgren, 2017; Vauhkonen et al., 2012).

The watershed algorithm is suitable for analysis of canopy height models and tree segmentation because LiDAR-based height data provides a representation of the vertical structure of trees and enables individual tree crown delineation. In watershed-based segmentation, the identified markers can be used to delineate individual tree crowns, where each segment is intended to correspond to an individual tree (Chen et al., 2006; Kakhani et al., 2019).

Previous studies have shown that watershed-based segmentation can be sensitive to noise and spatial resolution in height data, which may lead to different segmentation errors. A common problem is over-segmentation, where multiple local maxima within the same object cause it to be split into several segments (Meyer & Beucher, 1990). In forests where trees have complex crown structures, multiple local maxima may occur within the same tree crown. This increases the risk of over-segmentation, particularly at high spatial resolution (Lindberg & Holmgren, 2017). This means that the identification of local maxima, which is influenced by the choice of VWF parameters, directly affects the subsequent segmentation results.

Another well-known problem is under-segmentation, where several nearby trees are merged into a single segment. This usually occurs in dense stands or at low spatial resolution, where height differences between nearby trees are difficult to distinguish (Vauhkonen et al., 2012; Kaartinen et al., 2012). These errors are important to consider, as they affect both the number of trees detected and the delineation of individual crowns, which in turn may influence the reliability of individual tree detection.

The results of watershed-based segmentation are therefore strongly influenced by how the segmentation is controlled through the choice of parameters. These include how local maxima are identified, how sensitive the segmentation is to small variations in height data, and what criteria are used to determine whether adjacent segments should be considered separate or not. The choice of these parameters affects the number and delineation of segments and is therefore important to consider when interpreting the results (Kanda et al., 2004). This is particularly relevant for this study, where treetop detection using VWF is used as input for watershed-based segmentation, making the segmentation results dependent on the chosen detection parameters.

2.6.3 Marker-controlled watershed segmentation

To minimize the risk for over-segmentation, marker-controlled variants of the watershed algorithm have been developed (Chen et al., 2006; Kakhani et al., 2019). These methods use predefined markers to control the initiation of the segmentation process, providing object boundaries that better correspond to actual objects and reducing the number of segments (Meyer & Beucher, 1990; Kakhani et al., 2019). This makes marker-controlled approaches particularly relevant in forest environments where complex crown structures may otherwise increase the risk of over-segmentation.

In marker-controlled watershed segmentation, the markers are typically divided into inner markers, which represent the center of the objects, and background markers, which define the segmentation boundaries. The segmentation starts from these markers, and the regions grow until they meet, meaning that the placement of the markers largely determines the final object boundaries (Meyer & Beucher, 1990). This process is illustrated in Figure 3.

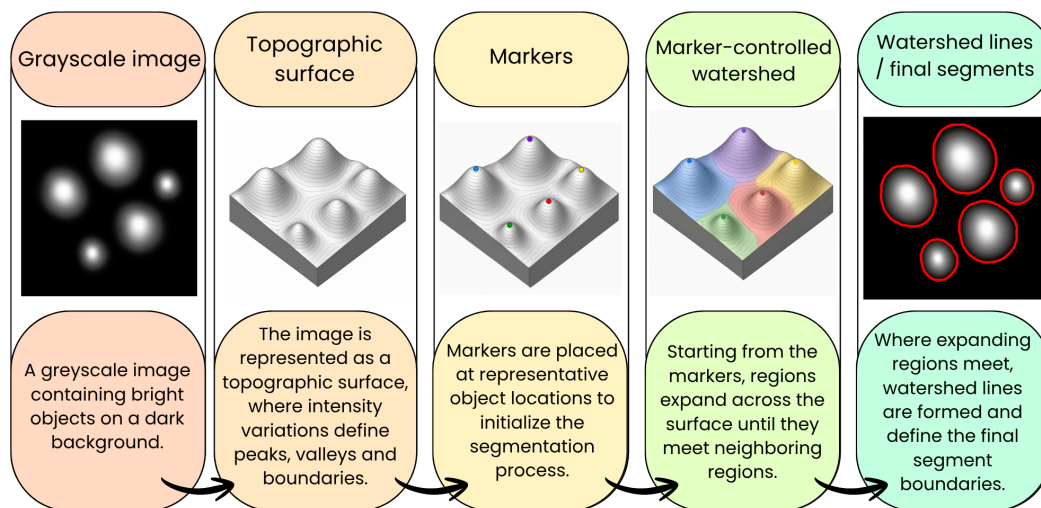


Figure 3. Schematic illustration of marker-controlled watershed segmentation. Source: Own illustration created in Canva, with AI-assisted image generation using ChatGPT.

2.7 Topographic position index

In addition to structural variables derived from LiDAR data, environmental factors are important for understanding variation in forest structure. One such factor is topography, which can influence local site conditions such as moisture availability and productivity.

The topographic position index (TPI) is a measurement used to describe the relative elevation of a point in relation to its surroundings, where the elevation value in a raster cell is compared to the average elevation within a defined neighborhood around the cell (Weiss, 2001; De Reu et al., 2013). Positive TPI-values represent locations that are higher than their surroundings, such as ridges, while negative values represent lower areas such as valleys. Values near zero correspond to flat areas or surfaces with a constant slope (Weiss, 2001).

Since TPI describes relative rather than absolute elevation, two points with the same elevation can have different TPI values depending on their local surroundings (Weiss, 2001; De Reu et al., 2013). This makes the index useful for identifying topographic features and for analyzing how the local terrain variation may influence ecological processes such as water accumulation, drainage and vegetation development (Weiss, 2001).

In forest applications, TPI has been shown to be useful for describing variation in stand structure (Szász et al., 2024). The study by Szász et al. (2024) showed correlations between TPI and tree height in LiDAR-based analyses, indicating that relative topographic position may influence the forest development and productivity. This suggests that topographic conditions may contribute to variations in forest structure and should therefore be considered when interpreting structural characteristics in forest analyses.

However, TPI is strongly dependent on the scale used in the analysis. The choice of neighborhood size is crucial in determining whether the index captures small-scale local variations or broader topographic patterns, meaning that the result can vary depending on the analysis radius used (Weiss, 2001; De Reu et al., 2013). De Reu et al. (2013) argue that this may lead to misleading classifications in heterogeneous landscapes, and that TPI should therefore be interpreted with caution and in relation to both the landscape characteristics and the purpose of the analysis.

2.8 Stratified identification of structural patterns

Structural characteristics of trees, as described above, are not only affected by age but also by site conditions. This implies that structural variables cannot be interpreted as absolute indicators of age without considering environmental context. To enable more accurate interpretation, analyses therefore need to account for variations in site conditions.

In ecological and statistical analysis, methods can be used to investigate whether certain phenomena are over- or underrepresented in specific contexts relative to chance expectations. This is done by comparing observed occurrences with expected occurrences based on the surrounding conditions, making it possible to analyze how organisms or structures are distributed in relation to environmental variation and to identify non-random patterns and associations between variables (Aarts et al., 2008; Tan et al., 2004).

One measure based on this principle is LIFT, which is used to measure the strength of association between variables by comparing how frequently variables occur together relative to what would be expected under independence (Han et al., 2012).

2.9 Model evaluation and performance metrics

The purpose of model evaluation in tree segmentation based on LiDAR data and CHM is to assess how well the segmentation results correspond to reference data from the field. The evaluation is usually carried out by comparison between detected trees with reference trees from field inventories (Kaartinen et al., 2012; Popescu et al., 2003). To quantify this agreement, automatically identified trees are compared with reference trees from field data, which allows the calculation of various performance measures (Kaartinen et al., 2012).

Performance metrics used to evaluate tree segmentation methods can be divided into different categories depending on what is to be analyzed. In the literature a distinction is made between object-based and pixel-based metrics (Kaartinen et al., 2012; Vauhkonen et al., 2012).

2.9.1 Performance metrics

In object-based evaluation, performance measures are based on the match between detected trees and reference trees from field data. Based on that match, the results are classified into correctly detected trees, omissions and commission errors, corresponding to missed and incorrectly detected trees, respectively (Kaartinen et al., 2012). These outcomes are divided into three categories (Alzubaidi et al., 2021).

- True Positive (TP) refers to an object that has been correctly identified by the model and is also present in the reference data.
- False Negative (FN) refers to an object that is present in the reference data but has not been detected by the model.
- False Positive (FP) refers to an object that has been identified by the model but is not present in the reference data.

Based on these outcomes, several performance measures can be calculated. Recall describes the proportion of reference trees that are correctly identified by the model

(Equation (1)), while precision describes the proportion of correctly identified trees in relation to all detected trees (Equation (2)). To combine precision and recall into a single measure, the F1 score is used (Equation (3)), which is the harmonic mean of precision and recall (Alzubaidi et al., 2021). According to Alzubaidi et al. (2021) and Minaee et al. (2022), these performance measures can be formulated as follows:

$$Recall = \frac{TP}{TP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$F1\ score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (3)$$

Performance metrics used in tree segmentation can also include pixel-based approaches, where the geometric overlap between predicted and reference segments is evaluated at the pixel level. A common metric is Intersection over Union (IoU), also known as the Jaccard index (Taha & Hanbury, 2015). However, the present study focused on object-based evaluation measures, as the primary aim was to assess the detection and classification of individual tree crowns rather than the exact geometric correspondence of crown boundaries.

Together, the concepts, methods and evaluation approaches presented in this chapter provide the theoretical foundation for the methodology used in this study, which is described in the following chapter.

3 Method

To address the aim of identifying trees with structural characteristics of older pine trees using LiDAR- derived structural variables, a structured methodological workflow was developed. This chapter first describes the study area, followed by an overview of the data and materials used in the study.

The subsequent sections present the workflows for data processing and crown segmentation, followed by parameter evaluations and sensitivity analyses used to justify the selected parameters. Finally, the method for classifying trees with older structures is presented, including the selection of variables and thresholds used in the classification.

3.1 Study area

The study area covers about 4652 hectares and is located between Särna and Idre in the northern part of Dalarna County (Figure 4). The area consists of a boreal forest landscape dominated by coniferous forest, mainly pine. The forest landscape shows variation in stand density and canopy cover. Tree height growth in the area is generally considered to be relatively low.

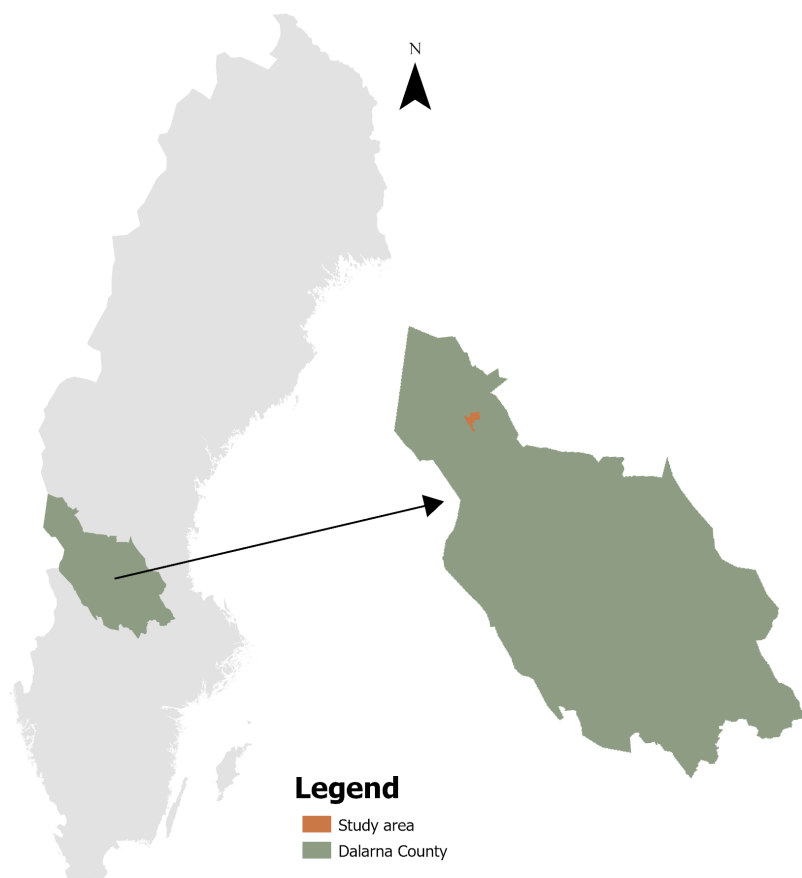


Figure 4. Location of the study area in northern Dalarna County, Sweden. The study area (orange) is located between Särna and Idre and is shown in relation to Dalarna County (green) and the rest of Sweden (grey).

3.2 Data and materials

The study used three CHM rasters derived from airborne LiDAR data. Two of the rasters were published by the Swedish Forest Agency (Skogsstyrelsen) and represent tree height models based on national airborne LiDAR acquired by the Swedish Mapping, Cadastral and Land Registration Authority (Lantmäteriet) in 2014 and 2022, respectively. These rasters have a lower spatial resolution with a cell size of 1 x 1 meter and a point density of approximately 1-2 points per m². The CHMs were derived from height-normalized point clouds, where the ground model was used to estimate vegetation height above ground level. The tree height rasters were generated using first-return LiDAR points and included processing steps to reduce gaps in the canopy surface (Lantmäteriet, 2022; Skogsstyrelsen, 2026b).

The third raster was obtained through Billerud and was derived from airborne LiDAR acquired in 2024. The dataset was produced by Arbonaut and has a higher spatial resolution with a cell size of 0.5 x 0.5 meter and a point density of approximately 20 points per m². Before CHM generation, the LiDAR point cloud was quality checked and height-normalized by Arbonaut. Ground points were set to zero while above-ground objects kept their height above the ground surface (personal communication with Arbonaut, May 20, 2026).

The comparison of spatial resolution was based on the 2022 and 2024 rasters (Figure 5), while the 2014 raster was used to estimate height growth. All CHM rasters used the same coordinate reference system, SWEREF 99 TM. In the remainder of this report, the two datasets used for comparison are referred to as the low-resolution CHM and the high-resolution CHM. All CHM datasets used in this study were made available through Skogforsk, which coordinated and provided access to the data used in this project. Although the comparison focuses on spatial resolution, the low- and high-resolution CHM datasets also differ in acquisition year, point density and data source, which were considered when interpreting the results.

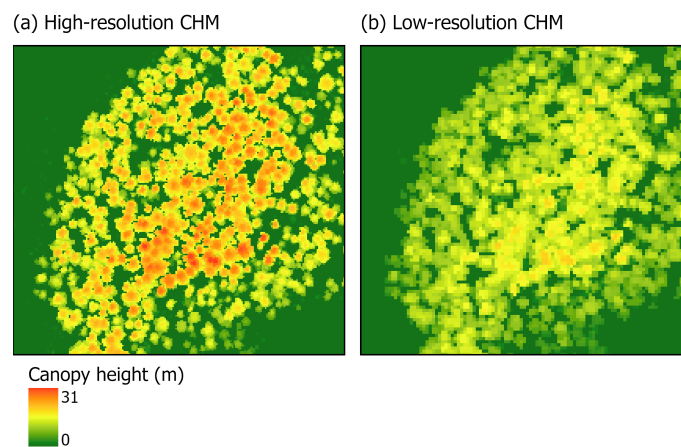


Figure 5. Representative subset of the high-resolution and low-resolution CHM, showing the same area in both datasets. The figure illustrates differences in spatial detail between the two resolutions.

Orthophotos produced by the Swedish Mapping, Cadastral and Land Registration Authority were downloaded as multiple GeoTIFF files. The orthophotos were then mosaicked into a single orthophoto and clipped to the exact extent of the study area. A digital elevation model (DEM) produced by the Swedish Mapping, Cadastral and Land Registration Authority was also used in the study. The soil moisture map was provided by the Swedish University of Agricultural Sciences.

To obtain validation data, a drone survey was conducted using a DJI Phantom 4 RTK equipped with an RGB camera. The collected drone imagery was processed into orthomosaics in Agisoft Metashape Professional. During the field survey, trees of high conservation value were identified and positioned using GNSS equipment and a tablet-based mapping application.

All pre-processing, segmentation and classification of the data were conducted in RStudio to ensure reproducibility and transparent workflow. ArcGIS Pro was used for GIS operations, such as visual inspection of results, map visualization and production of figures included in the report.

Table 1. Overview of the datasets used in the study, including data source, acquisition method, year, spatial resolution and coordinate reference system (CRS).

Data	Data source	Data type	Year	Spatial Resolution	CRS
CHM	Skogsstyrelsen	Airborne LiDAR	2014	1 x 1 m	SWEREF 99 TM
CHM	Skogsstyrelsen	Airborne LiDAR	2022	1 x 1 m	SWEREF 99 TM
CHM	Billerud	Airborne LiDAR	2024	0.5 x 0.5 m	SWEREF 99 TM
True Orthophoto	Lantmäteriet	Aerial Orthophoto	2022	0.5 x 0.5 m	SWEREF 99 TM
DEM	Lantmäteriet	Elevation Model	2019	2 x 2 m	SWEREF 99 TM
Soil moisture map	SLU	Raster map	2020	2 x 2 m	SWEREF 99 TM
UAV Orthomosaic	DJI Phantom 4 RTK	Orthomosaic (RGB)	2026	0.18 x 0.18 m	SWEREF 99 TM

ChatGPT (GPT-5 and GPT-5.5) was used during the study to support code development and debugging. The tool was also used as linguistic support during the writing process, including grammar checking, suggestions for academic phrasing and word choice, and assistance in refining translations from Swedish to English to achieve more natural scientific language. All generated outputs were reviewed, validated and revised by the authors before use in the study, who remained responsible for the content of the study (OpenAI, 2026).

3.3 Data preprocessing

This section describes the preprocessing applied to the CHM rasters prior to crown segmentation. Further justification of the selected preprocessing steps is provided in Section 3.4.1.

A quality check was performed by verifying the raster resolution in both RStudio and ArcGIS Pro to ensure that the CHM rasters used for segmentation and classification had the correct cell sizes (0.5 and 1.0 m). Although both of these CHM rasters were based on the horizontal coordinate reference system SWEREF 99 TM (EPSG:3006), minor differences were identified in how the coordinate reference system (CRS) was defined in the metadata (Table 2). To ensure consistency, the high-resolution CHM was defined as EPSG:3006 before further processing. The code also included validation checks to ensure that input data met the required specifications, contributing to robust and reproducible data handling.

The low-resolution CHM raster contained a color table that initially interpreted the elevation values as categorical classes. The color table was therefore removed so that the height values were instead treated as continuous numerical data. The low-resolution raster was also converted to meter, as the original height values were provided in decimeters.

Table 2. Summary of data specifications and preprocessing steps for the high- and low-resolution CHM datasets.

Data Specification	High-Resolution CHM	Low-Resolution CHM
Original resolution	0.5 x 0.5 m	1 x 1 m
Original data type	Floating point (32-bit)	Integer (16-bit)
Maximum height	Approx. 29 m	Approx. 263 dm
Original height unit	Meter	Decimeter
Unit conversion	None	Scaling ($\div 10$)
Final height unit	Meter	Meter
Original CRS	SWEREF99 TM (Custom)	SWEREF99 TM (EPSG:3006)
CRS adjustment	Defined as EPSG:3006	None
Final CRS	SWEREF99 TM (EPSG:3006)	SWEREF99 TM (EPSG:3006)

A shapefile representing the study area was imported and reprojected to match the CRS of the other datasets, SWEREF99 TM (EPSG:3006). The study area consisted of a single contiguous polygon covering approximately 4652 ha, and a 5 m buffer was created around it to clip the CHM rasters and reduce edge effects caused by partially represented tree crowns near the boundary.

To remove the influence of low vegetation, all CHM values under 3 meters were masked and set to NA. This ensured that low vegetation did not influence the smoothing or affect tree crown surfaces. The CHM rasters were smoothed using a 3 x 3 Gaussian filter to reduce small local height variations and create a more continuous surface.

Before crown segmentation, only pixels with a height of at least 3 m in both the smoothed and original CHM were retained for further analysis. Missing values were ignored during smoothing, and constant NA padding corresponding to a one-pixel border was applied around the raster edges prior to filtering to allow edge cells to be processed and reduce edge effects. The preprocessed CHM rasters were then exported and used as input for the subsequent watershed segmentation.

3.4 Crown segmentation

This section describes the workflow used for treetop detection and crown segmentation of the CHM rasters, including main processing steps and selected parameter settings. Further justification of the selected parameter values is provided in Section 3.4.1.

The workflow begins with treetop detection, followed by crown segmentation. Treetop detection was performed using variable window filtering (VWF), an established CHM-based method where the search window size is adjusted to tree height based on the relationship between tree height and crown size (Popescu & Wynne, 2004; Chen et al., 2006).

This approach was considered suitable because it allows the detection to better account for variation in crown dimension than fixed window size and therefore reduces the risk of over- and under-detection of treetops (Popescu & Wynne, 2004).

Three parameter sets were defined and evaluated: Small, Mid and Large, representing small, intermediate and large window sizes. Each set consisted of a global and a fine height-dependent window function. For both functions, the window size was calculated from tree height using a linear relationship defined by a slope and an intercept. The global function describes the overall relationship between tree height and window size, while the fine function allows additional adjustments of the window size at a smaller spatial scale. The fine function was prioritized in the final analysis. Among the tested parameter sets, the Mid parameter set was selected for the final workflow, with associated parameter values presented in Table 5.

After treetop detection, a thinning procedure was applied to merge adjacent local maxima and ensure that only one treetop represents each individual tree. A thinning distance of 0.5 meters was used for both CHM rasters. The detected treetops were then used as markers in the marker-controlled watershed segmentation. The CHM surface was interpreted as a height model, and each segment represented a delineated tree crown associated with a detected treetop.

To perform segmentation of the high-resolution CHM raster, the study area was divided into tiles of 1000 x 1000 m to manage the large data volume and reduce computational demands. In total, 120 tiles were generated. To reduce edge effects at tile boundaries, each tile was extended with an overlapping buffer of 10 meters prior to segmentation. The segmentation was then performed separately for each tile including the buffer zone.

After segmentation, only those treetops and crown polygons whose centroid were located within the original tile extent were retained. In this way, each crown appeared only once in the final dataset. The results from all tiles were merged into a single continuous dataset covering the entire study area. In contrast, the segmentation of the low-resolution CHM raster was performed as a continuous analysis without partitioning into tiles, as the lower data volume did not impose the same computational constraints.

The resulting segments were saved as polygon features (shapefile format) for further analysis. These segments were then used as the basis for extracting structural variables and performing the classification of tree crowns.

3.4.1 Justification of preprocessing and segmentation settings

This section presents the parameter evaluations and sensitivity analyses performed to justify preprocessing and segmentation settings. The purpose was to identify robust parameter values for CHM preprocessing, treetop detection and crown segmentation before applying the workflow on the full study area.

Delimitation and buffer

Table 3 summarizes how different buffer distances (0, 5 and 10 m) affect the elevation data. When the buffer increased from 0 to 5 and 10 m, the area included in the analysis increased but the mean height and spread are largely unchanged for both low- and high-resolution CHM.

Table 3. Summary of the effect of the different buffer distances (0,5, and 10 m) on area, mean height, and standard deviation (SD) of CHM pixel height values for the low- and high- resolution CHM datasets.

Buffer (m)	Resolution (m)	Area (ha)	Mean height (m)	SD (m)
0	1	4645.3	4.576	5.467
5	1	4665.7	4.556	5.463
10	1	4686.0	4.537	5.460
0	0.5	4626.6	4.094	5.287
5	0.5	4646.3	4.089	5.286
10	0.5	4666.0	4.084	5.284

Table 4 shows how the buffer zone differs from the inner area. The average height of the inner zone is stable regardless of the buffer size and the edge zone is dominated by very low values.

Table 4. Comparison of mean height and standard deviation (SD) of CGM pixel height values between the inner area and the buffer zone for different buffer distances.

Buffer (m)	Inner zone		Edge zone	
	Mean height (m)	SD (m)	Mean height (m)	SD (m)
0	4.575	5.467	—	—
5	4.576	5.467	0.371	1.992
10	4.556	5.463	0.067	0.828

— Edge zone not defined for 0 m buffer.

A 5 m buffer was added around the study area to reduce edge effects, such as incorrect or abnormally low height values occurring near the boundaries where tree crowns are only partially covered by the data. This buffer size was considered a reasonable trade-off between minimizing edge effects and avoiding unnecessary expansion of the analysis area.

Selection of CHM smoothing method

CHM smoothing is commonly used as a preprocessing step in individual based treetop detection to stabilize the canopy height surface and reduce small-scale variations that may cause false local maxima (Pitkänen et al., 2004). In this study, both a mean filter and a Gaussian filter were evaluated. The mean filter calculates a simple mean value of all the pixels within the kernel, meaning that all pixels are assigned the same weight. This results in a general smoothing of the height surface but may also reduce local variations and thereby reduce potential local maxima (Gonzalez & Woods, 2018).

A Gaussian smoothing filter was applied using a normalized 3×3 Gaussian kernel $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} / 16$. The Gaussian filter is based on a Gaussian function in which the weight decreases with increasing distance from the center of the kernel. Pixels near the center are assigned a greater weight than pixels further from the center, which reduces noise while better preserving structures in the canopy height surface (Witkin, 1984).

To evaluate which smoothing method was most suitable, both a 3×3 mean filter and a 3×3 Gaussian filter were tested on a representative subset of the study area. The results were compared quantitatively based on the number of identified crowns and qualitatively through a visual assessment of the resulting crown segments in relation to the CHM.

The Gaussian smoothing resulted in a slightly higher number of identified crown segments within the subset compared to the mean filter. A visual comparison indicated that these segments corresponded more closely to local maxima in the canopy height surface while the mean filter in some cases smoothed small height variations so that potential maxima were not identified (Figure 6). Based on this comparison, Gaussian smoothing was selected for this study.

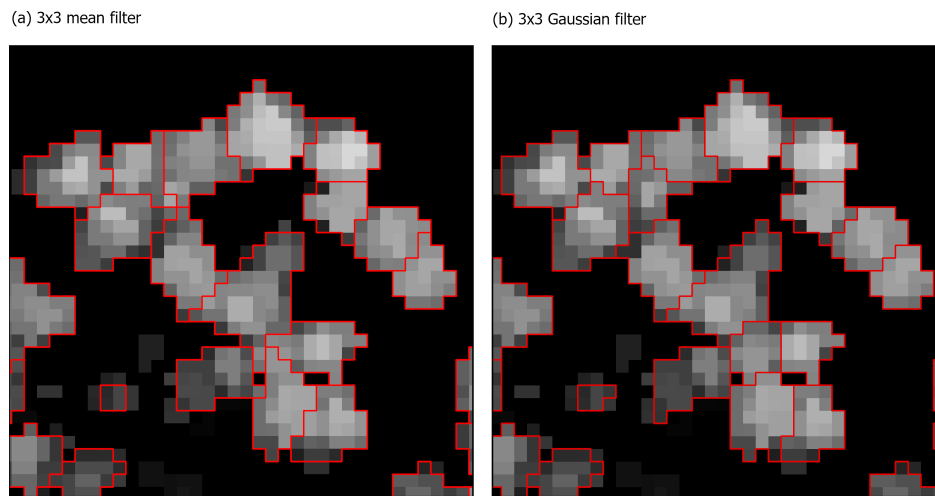


Figure 6. Comparison of smoothing applied to the low-resolution CHM using (a) a 3×3 mean filter and (b) a 3×3 Gaussian filter. The Gaussian filter preserves canopy maxima more effectively than the mean filter.

Masking of low vegetation

The height distribution of CHM values within the study area was analyzed to select a threshold-value to exclude low vegetation. Histograms for the height-values were created for both the high-resolution and the low-resolution CHM-data. A threshold of 3 meters was selected (see red dashed line in Figure 7), since it lies above the large number of low-height values and therefore delineates low vegetation without affecting tree canopies. The pattern was similar for both spatial resolutions, which indicates that the threshold provides a consistent result for both high- and low-

resolution CHM. The histograms were complemented by summary statistics for the tested thresholds, which are reported in Appendix A

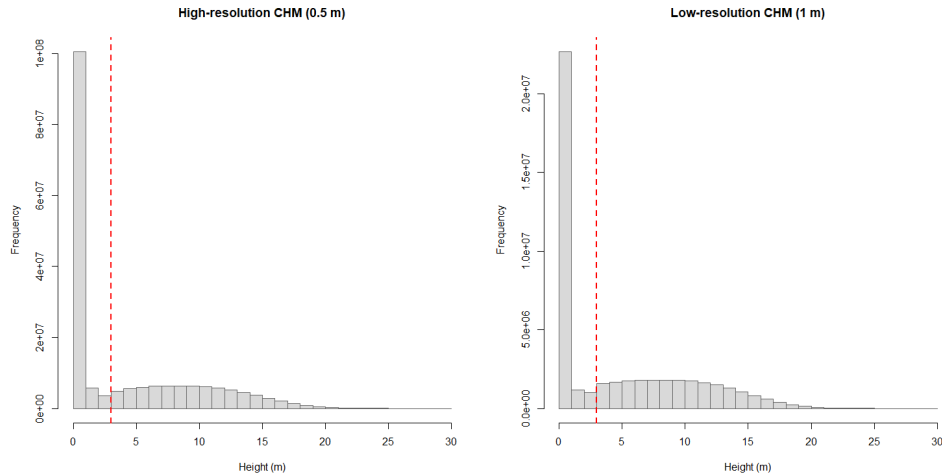


Figure 7. Height distributions of CHM values for the high-resolution (0.5 m) and low-resolution (1.0 m) datasets within the study area. The red dashed line marks the selected threshold of 3 m used to separate low vegetation from tree canopies.

Tree top detection using variable window filtering (VWF)

As described in Section 3.4, three parameter sets were evaluated for the variable window filtering: Small, Mid and Large. The purpose of the sensitivity analysis was to assess how strongly treetop detection was affected by the selected parameter set and to ensure that the result was stable within the tested range. To reduce computational time, the analysis was performed on a 200 x 200 m subset of the study area.

Table 5 shows the tested parameter sets for each CHM-resolution, together with the corresponding number of identified treetops and calculated tree density. The results show that the differences between the parameter sets were relatively small, especially between Mid and Large. Therefore, the Mid parameter was selected for the final analysis for both resolutions.

Table 5. Tested parameter sets for variable window filtering (VWF) and corresponding treetop detection results for both CHM resolutions. Slope and intercept parameters refer to the linear variable window function.

Dataset	Set	Global slope	Global intercept	Fine slope	Fine intercept	Treetops	Trees/ha
0.5 m	Small	0.040	0.60	0.030	0.50	1599	399.75
0.5 m	Mid	0.045	0.90	0.035	0.80	1557	389.25
0.5 m	Large	0.050	1.00	0.040	0.80	1556	389.00
1.0 m	Small	0.040	0.60	0.030	0.50	1427	356.75
1.0 m	Mid	0.045	0.90	0.035	0.80	1427	356.75
1.0 m	Large	0.050	1.00	0.040	0.80	1427	356.75

Thinning distance

Multiple local maxima can occur within the same tree crown when detecting treetops from a CHM. To reduce duplicate detection, some form of thinning or merging of neighboring peaks is usually applied to retain only one representative peak per tree.

Previous studies in individual-based tree analysis using LiDAR show that post-processing of treetops, such as separation or merging of neighboring detections is strongly dependent on the spatial resolution of the data and the forest stand structure, and that no general standard parameter values exist (Kaartinen et al., 2012; Lindberg & Holmgren, 2017).

Therefore, the thinning distance was evaluated empirically for the current datasets. A sensitivity analysis was conducted using distances between 0.5 and 1.1 m, where the lower bound corresponded to the pixel size of the high-resolution CHM. The analysis was conducted separately for the high-resolution and low-resolution CHM data.

The results showed a modest decrease in the number of detected treetops between thinning distances of 0.5 and 0.8 m for the high-resolution CHM (2.8%) and a clearer reduction was observed at 1.1 m. For the low-resolution CHM, the number of detected treetops barely changed across the different thinning distances. The low sensitivity for different thinning distances is likely related to the raster resolution (1×1 m), where the tested thinning distances fall within a similar spatial range relative to the pixel size and reduce the effect of the thinning.

A visual comparison of the detected treetops for the high-resolution CHM confirmed these patterns (Figure 8). Only small differences were observed between thinning distances of 0.5 and 0.8 m, although the 0.5 m distance identified some additional local maxima that corresponded to local maxima in the CHM. At 1.1 m, a clearer reduction in the number of detected treetops was observed.

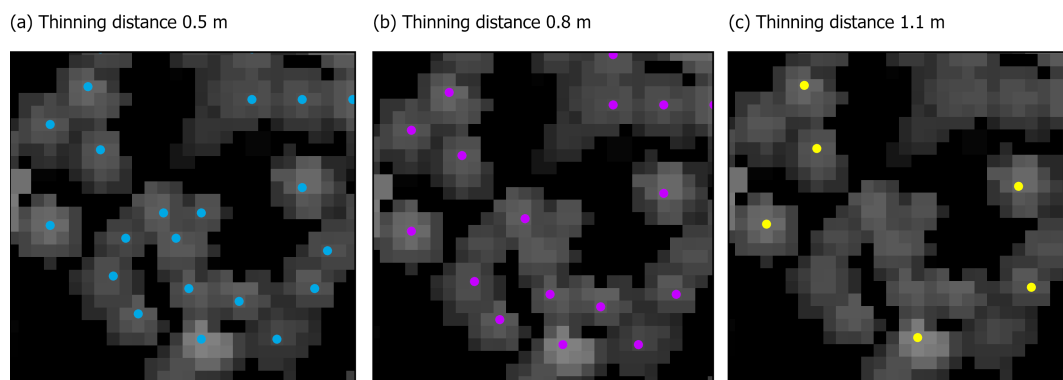


Figure 8. Comparison of treetop detection results using different thinning distances. The background represents the CHM, and the colored points represent detected treetops. Panels show thinning distances of (a) 0.5 m, (b) 0.8 m, and (c) 1.1 m.

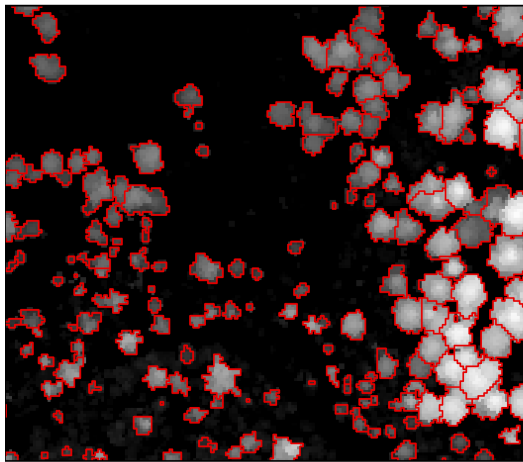
Based on both the quantitative sensitivity analysis and the visual assessment, a thinning distance of 0.5 m was selected. The detailed results of the thinning distance sensitivity analysis are presented in Appendix B

Evaluation of threshold applications relative to smoothing

Two different preprocessing strategies for the CHM were compared to determine at which stage low height values should be thresholded in relation to smoothing and how this affected the segmentation results. In the first strategy, the height threshold was applied after smoothing (ST), while in the second strategy it was applied before smoothing (TS). For each strategy, the entire segmentation process was subsequently carried out, and the results were visually compared (Figure 9). The comparison showed that applying the threshold after smoothing resulted in more coherent crown shapes and a more complete crown delineation.

This is because smoothing affects each pixel's height value by incorporating height values from surrounding pixels. Pixels that are just below the threshold value in the original CHM raster may therefore have a slightly higher value after smoothing. If the threshold is applied before smoothing, these pixels are removed immediately, which may cause some segments to be removed. Therefore, the ST approach was selected for further analysis.

(a) Threshold after smoothing (ST)



(b) Threshold before smoothing (TS)

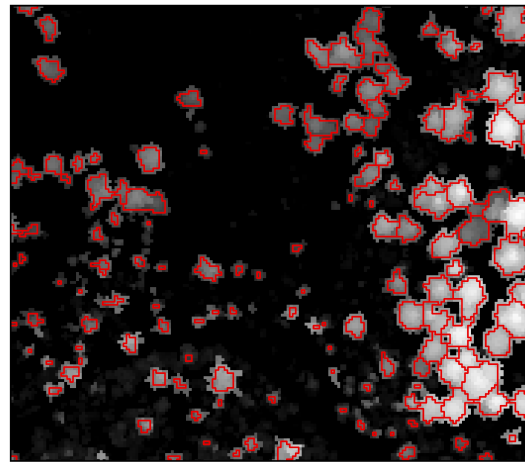


Figure 9. Visual comparison of crown segmentation results using two preprocessing strategies in high-resolution CHM data. (a) Threshold applied after smoothing. (b) Threshold applied before smoothing.

The smoothing caused height values to be affected by surrounding cells. This led to individual pixels exceeding the threshold even though their original height was lower. To ensure that the crown segmentation was based on actual height structures, an additional condition was therefore introduced where a pixel was only included if it was ≥ 3 meters in both the smoothed and the original CHM raster. The stricter condition constrains crown expansion and results in more spatially consistent crown boundaries (Figure 10).

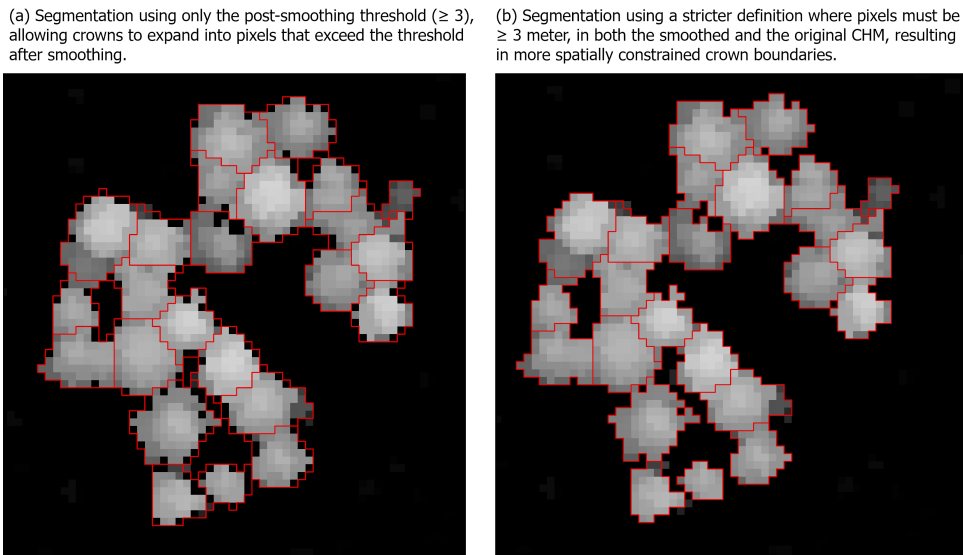


Figure 10. Comparison of crown segmentation using two threshold definitions. (a) Segmentation using only the post-smoothing threshold (≥ 3 m). (b) Segmentation where pixels are required to exceed the threshold in both the smoothed and the original CHM.

3.5 Classification of Scots pine trees with old-growth structural characteristics

The classification was based on the segmented objects produced by the crown segmentation workflow, which were interpreted as potential tree crowns. Since the segmentation was based solely on height variation in the CHM, some segments may also represent other elevated objects rather than actual tree crowns. Filtering criteria were therefore included to reduce the inclusion of segments unlikely to represent individual tree crowns. The stratification procedure is described in more detail in Section 3.5.1, while the structural variables used for classification are further explained in Section 3.5.2.

During the development of the classification approach, LIFT was evaluated as an initial method for identifying combinations of structural characteristics that occur more frequently than expected. Tree height, crown diameter and height growth were initially included as classification variables, while flatness was later tested as an additional variable to represent crown shape. However, the resulting combinations proved too restrictive for the current dataset and only a limited number of trees were identified. The method required several variables to show strong associations simultaneously, which substantially reduced the number of retained crown segments. LIFT was therefore not used in the final classification.

Based on these results, a stratified percentile-based approach was instead applied to account for environmental variation by evaluating structural characteristics within strata representing similar site conditions, rather than applying the same thresholds across the entire study area. The strata were defined based on elevation, soil moisture and TPI.

To assign environmental conditions to each crown polygon, the segmented crown polygons were linked to raster data representing elevation and soil moisture. The mean elevation value was extracted for each crown polygon using zonal statistics based on the DEM raster. These values were then divided into three tertile-based classes, meaning the data were split into three equal groups (low, medium, high).

Soil moisture was linked to the crown polygons by calculating the most frequently occurring class within each polygon and assigning it to the respective crown. The soil moisture data were classified into three classes: dry-mesic, mesic-moist and moist-wet. Topographic position was calculated using TPI, based on a circular neighborhood with a radius of 40 meters. The mean TPI value was calculated for each crown polygon and then divided into three tertile-based classes (low, medium, high).

The elevation, soil moisture and TPI classes were then combined to define the final strata used in the classification, resulting in a total of 27 strata. Structural variables describing individual tree characteristics were calculated for each crown polygon, including height, height growth, crown diameter and flatness.

Height growth was calculated as the difference in the 95th percentile (P95) height between two measurement occasions and then converted to annual growth by dividing by the number of years between acquisitions: 2014-2024 for the high-resolution CHM and 2014-2022 for the low-resolution CHM. To reduce the influence of unrealistic values, crown polygons with negative total growth or total growth exceeding 6 m during the study period were excluded by setting these values to NA. The 6 m threshold was used as an initial filtering step to remove clearly unrealistic outliers in the growth estimates before the subsequent classification procedure. For the low-resolution dataset, this corresponds to excluding annual growth values above 0.75 m year⁻¹ and for the high-resolution dataset above 0.60 m year⁻¹.

Crown diameter was calculated from the polygon area by assuming a circular crown shape. Crown polygons with an area exceeding 130 m² were flagged as potentially under-segmented, as unusually large crown areas may indicate that multiple neighboring trees were merged into a single crown segment.

Flatness was calculated as the difference between the 95th percentile of height values within the crown and the mean height, normalized by the 95th percentile. Crowns with invalid or negative values were excluded from further analysis. Although flatness was initially considered as a classification variable within the percentile-based approach, its distribution showed limited sensitivity to changes in percentile thresholds and was therefore retained only as a fixed quality filter to remove crown segments with unrealistic crown shapes rather than as an indicator of old-growth characteristics.

In addition to the stratified percentile-based thresholds, fixed biological thresholds were first applied to exclude crown polygons that did not meet basic criteria. These thresholds included minimum height of 12 m, a minimum crown diameter of 4.5 m, a maximum annual growth of 0.06 m year⁻¹, and a minimum flatness value of 0.08. Only crown polygons that met these criteria were retained for further analysis. These thresholds were based on a combination of previous studies, empirical data distributions, and sensitivity analyses described in Section 3.5.2.

The selection and evaluation of the classification thresholds were also supported by discussions with forestry experts at Skogforsk to ensure ecological relevance and practical applicability.

Thereafter, group-specific percentile-based thresholds were calculated for each stratum based on height, crown diameter and height growth. For height and crown diameter, the 75th percentile was used as threshold while the 15th percentile was used for height growth. The final classification therefore combined percentile-based and fixed thresholds. The percentile-based thresholds enabled relative comparisons within strata, while the fixed thresholds ensured that all retained crown polygons met minimum biological and structural requirements.

For the remaining crowns, the structural variables were compared with the group-specific percentile-based thresholds for each stratum. For each variable that met its respective criterion, the crown was assigned one point. The total score was then used to classify each crown polygon into categories based on how many criteria were fulfilled: A_full (all three criteria fulfilled), B_partial (two out of three criteria fulfilled), and Non_candidate (zero or one criterion fulfilled).

Segments classified as A_full had to meet both the stratified percentile-based and fixed thresholds for tree height, crown diameter and growth. Segments classified as B_partial had to fulfill two out of three percentile-based thresholds while still meeting the fixed threshold for the remaining variable. This ensured that all included crown polygons fulfilled the minimum structural requirements even when not all percentile criteria were met.

3.5.1 Strata

The previously defined strata were used to account for local variations in site conditions and topography by grouping tree crowns with similar soil moisture, elevation and topographic position. The purpose was to enable comparisons between trees growing under similar environmental conditions and to allow the percentile-based thresholds to be calculated relative to ecologically comparable groups.

Soil moisture is an important factor influencing tree growth, survival and the structural development of forest stands. Previous studies have shown that variations

in soil moisture can explain differences in tree height, crown diameter and mortality, which in turn affect how structural characteristics develop over time (Nordin et al., 2023). Soil moisture was therefore included as a stratification variable (Figure 11).

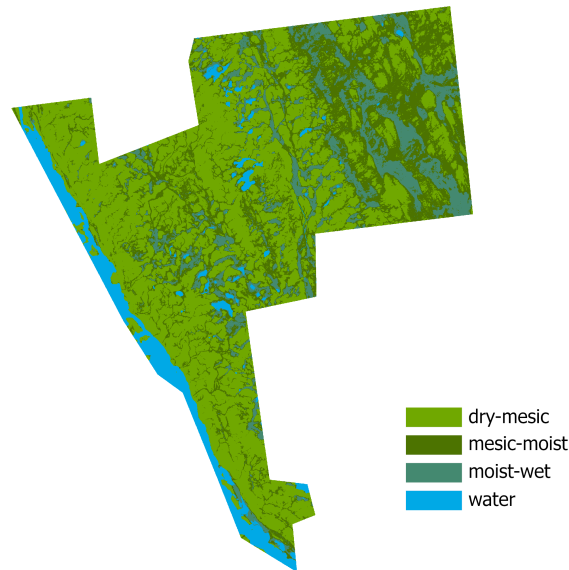


Figure 11. Soil moisture map of the study area shows the different soil moisture classes. Data source: Swedish University of Agricultural Sciences (SLU).

Forest structure is influenced by both tree age and site factors such as climate and resource availability (Pretzsch et al., 2022). Since these factors may vary with elevation, tree growth and structural characteristics can also differ across elevation gradients (Körner, 2007). Elevation was therefore included as a stratification variable (Figure 12).

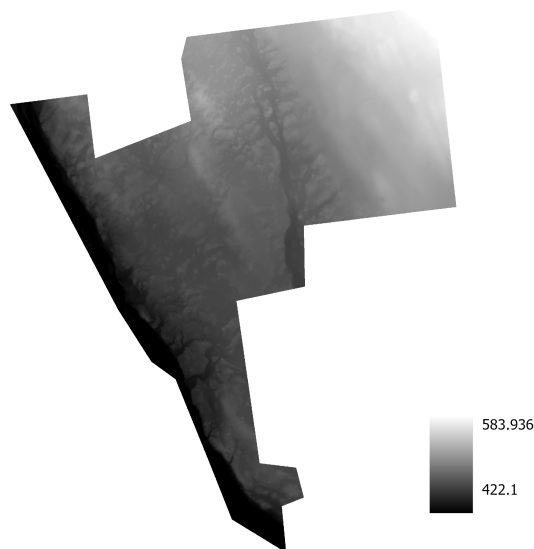


Figure 12. Digital elevation model (DEM) of the study area showing variation in elevation (m above sea level). Data source: Lantmäteriet.

Topographic position was represented using TPI, which describes the relative elevation of a point in relation to its surrounding neighborhood (Weiss, 2001; De Reu et al., 2013). To determine an appropriate spatial scale for the analysis, the TPI radius was evaluated for both low- and high-resolution datasets using radii of 10, 20, 30, 40 and 50 m.

In both datasets, the proportion of crowns changing TPI class decreased progressively with increasing radius, from approximately 22-23% between 10 and 20 m to about 8.5-8.6% between 40 and 50 m (Table 6). At the same time, the correlation between consecutive radii increased, with the relationship between 40 and 50 m being very high ($r \approx 0.99$).

Table 6. Changes in TPI classification and correlation between consecutive radii for low- and high-resolution datasets.

Radius (m)	Class changes low (%)	Class changes high (%)	Correlation low (r vs previous)	Correlation high (r vs previous)
20	22.0	23.1	0.917	0.909
30	14.5	14.8	0.962	0.961
40	10.8	10.9	0.979	0.979
50	8.5	8.6	0.987	0.987

This pattern was also visually confirmed (Figure 13), where a radius of 10 m captured small local variations, while the strong similarity between 40 and 50 m indicated that only minor changes occurred beyond this scale. Based on these results, which showed a similar stabilization pattern in both datasets, a radius of 40 m was selected for the TPI calculation.

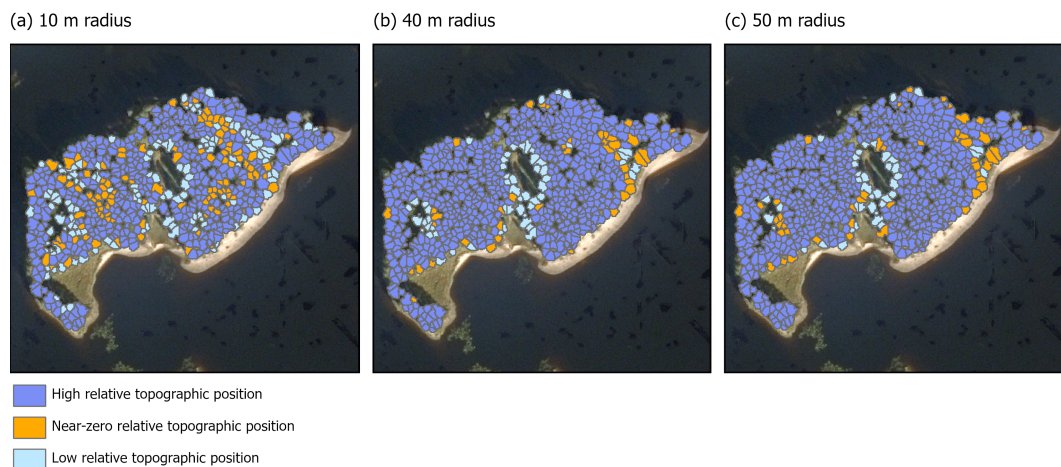


Figure 13. Illustration of the spatial effect of increasing the TPI radius from 10 to 40 m in the high-resolution CHM, while the similarity between 40 and 50 m visually confirms that increasing the radius beyond 40 m results in only minor changes.

The cut-off values for the elevation- and TPI-based strata were calculated separately for the low- and high-resolution datasets by dividing the observed distributions into tertiles (Table 7). The 33rd and the 67th percentiles were used to define three relative classes for elevation and topographic position. These classes were then combined with the soil moisture classes to create the final strata used in the classification. The resulting cut-off values were saved to ensure reproducible stratification across analyses and spatial resolutions.

Table 7. Tertile-based cut-off values used to define elevation and TPI strata for low- and high-resolution datasets. The lower and upper cut-offs correspond to the 33rd and 67th percentiles of the observed distributions.

Resolution	Elevation lower cut-off (m)	Elevation upper cut-off (m)	TPI lower cut-off	TPI upper cut-off
Low	468.449	499.508	-0.129	0.172
High	468.354	499.442	-0.013	0.186

3.5.2 Structural variables used for classification

Based on the structural characteristics described in the theoretical background, a set of measurable variables was selected to identify Scots pine trees with old structural characteristics. These include tree height, crown diameter, height growth, and a measure of crown flatness. Together, these variables describe characteristics that often change as pine trees age (Handegård et al., 2021; Oliver & Larsson, 1996).

Percentile-based measures of tree height are commonly used in LiDAR-based forest analysis (Næsset, 2002). The 95th percentile (P95) was used to represent the upper part of the vertical crown structure while reducing sensitivity to extreme values compared to using the maximum height. Height growth was based on differences in P95 values between measurement occasions and converted to annual growth. Unrealistic growth values were excluded to reduce the influence of segmentation inconsistencies and outliers in the CHM data.

In the classification step, fixed thresholds were applied, including a minimum height of 12 meters and a maximum annual growth of 0.06 m year⁻¹, to exclude unrealistic values and focus on trees with more developed structural characteristics. The growth threshold was determined based on a sensitivity analysis of how the fixed and percentile-based thresholds interacted (Table 10).

In both datasets, all trees within the 75th percentile for height were taller than 12 m. This indicates that the fixed threshold did not further constrain the percentile-based threshold selection but was applied as an additional constraint to ensure that trees fulfilling other criteria, such as crown diameter and growth, also exceeded a minimum level of structural development.

Crown diameter was included as a structural variable, as it is commonly used in analyses of forest structure and remote sensing and describes the horizontal spread of the crown (Rautiainen & Stenberg, 2005). Crown size is not random but is related to the tree's dimensions, such as height and trunk diameter (Hasenauer, 1997). Studies show that the average crown diameter for Scots pine is around 5.4 meters (Buchacher & Ledermann, 2020) while maximum crown diameters for free-standing trees can reach approximately 10-13 meters (Buchacher & Ledermann, 2020; Hasenauer, 1997) which corresponds to a crown area of approximately 80-130 m² assuming a circular crown projection.

The distribution of crown diameter shows that most values are concentrated between approximately 2 and 5 m with a gradual decrease towards larger diameters (Figure 14). The 75th percentile threshold of crown diameter ranged between approximately 4.0 and 4.8 m across the different strata.

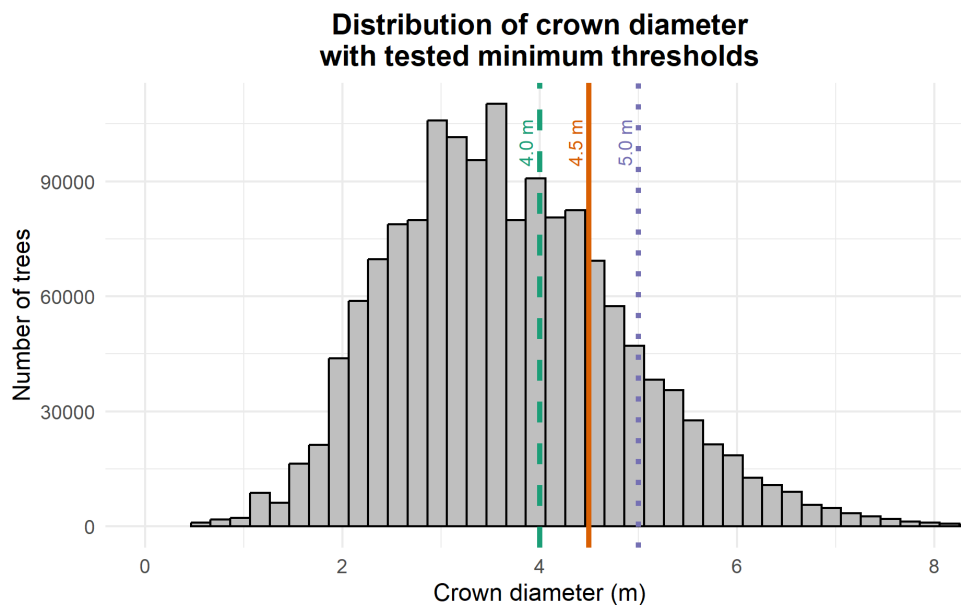


Figure 14. Distribution of crown diameter with tested minimum thresholds (4.0, 4.5 and 5.0 m). The threshold is shown relative to the overall distribution and illustrates how increasing the minimum threshold progressively excludes trees from the upper percentile-based selection.

As shown in Table 8, a minimum threshold of 4.5 m resulted in a limited reduction of the trees selected within the 75th percentile, whereas increasing the threshold to 5 m excluded a substantially larger proportion. Further analysis (Table 9) showed that a considerable share of the trees excluded when increasing the threshold from 4.5 m to 5 m still fulfilled the other classification criteria.

This indicates that a threshold of 5 m would be overly restrictive, as it removes many trees with structural characteristics associated with older trees. A threshold of 4.5 m was therefore selected as a more balanced lower limit, excluding the smallest crowns while retaining the majority of relevant candidates identified by the percentile-based selection.

Table 8. Effect of different minimum crown diameter thresholds on the 75th percentile-based selection.

Threshold (m)	Trees retained	Trees removed	Proportion retained
4.0	369.706	0	1.000
4.5	336.781	32.925	0.911
5.0	211.230	158.476	0.571

Table 9. Number and proportion of trees excluded when increasing the minimum crown diameter threshold from 4.5 m to 5 m within the 75th percentile selected trees. The table also shows how many of the excluded trees fulfilled other classification criteria.

Metric	Trees	Proportion (%)
Total removed (4.5-5 m)	125.551	100
Passing height criterion	31.390	25.0
Passing growth criterion	30.601	24.4

When the same thresholds were applied to the low-resolution dataset, the minimum threshold had no effect as all trees above the 75th percentile already exceeded 5 m. This confirms that the percentile-based threshold alone was sufficient in this case.

These results highlight the importance of evaluating threshold values in relation to the specific dataset rather than relying on predefined threshold values. While crown diameter is strongly related to tree size and growth (Buchacher & Ledermann, 2020), the analysis shows that appropriate threshold levels depend on the underlying data distribution. The selected thresholds were therefore based on empirical observations, to ensure that the classification reflects the structural variation present in the data.

Crown area was not used as an analysis variable, but as an aid to identify unreasonable crown segments. Based on the reference values above, crown areas that exceed 130 m² may indicate under-segmentation, particularly in the low-resolution dataset. Crown area calculated from segmented crown polygons was therefore used to flag crowns larger than 130m² as potentially under-segmented.

Percentiles

Percentile-based thresholds were used to identify trees that deviated from the local distribution within each stratum. The purpose was to enable a relative comparison between trees growing under similar environmental conditions, in addition to the fixed thresholds used to maintain biological plausibility.

The 25th and the 75th percentiles were initially selected based on the group-wise percentile tables generated for all analyzed variables within each stratum presented in Appendix D and Appendix E. These levels represented the lower and upper quartiles of the local distribution and were used as the starting point for parameter calibration (Figure 15). For variables where high values represent older structural characteristics, such as tree height and crown diameter, the 75th percentile was used. For annual growth, where low values are characteristic of older trees, the 25th percentile was initially applied.

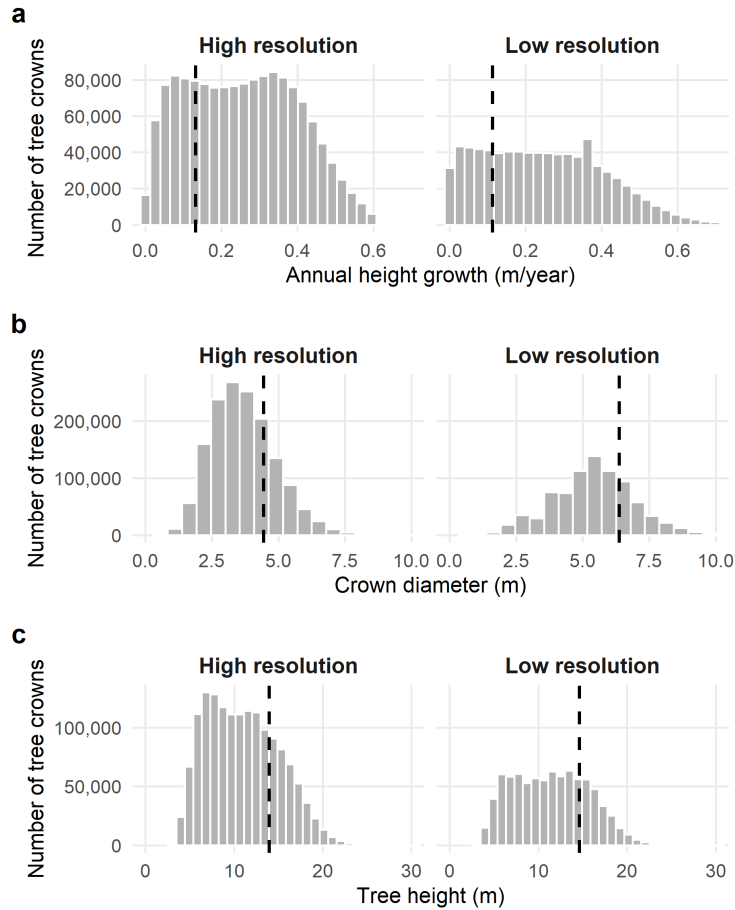


Figure 15. Distribution of (a) annual growth, (b) crown diameter, and (c) tree height for the high- and low-resolution datasets. Dashed lines represent the initial selected percentile thresholds used for classification: the 25th percentile for annual growth and the 75th percentile for crown diameter and tree height.

After the initial classification, the selected crown segments were examined more closely, several trees classified with low growth still showed growth rates that were too high to represent the reduced growth expected in older trees. This showed the need for further adjustment of the growth threshold used in the final classification.

A more detailed sensitivity analysis was therefore conducted for the growth variable (Table 10), in which both the percentile-based and the fixed threshold were evaluated. Lower and higher percentiles were tested together with stricter fixed thresholds to assess how the relative and absolute definitions of low growth interacted.

Table 10. Sensitivity analysis of the annual growth variable using percentile thresholds from the 15th to the 35th percentile in combination with fixed thresholds of 0.06 and 0.08 m year⁻¹. The table shows the number of crown segments selected by both criteria, as well as the number excluded by either the percentile-based or the fixed threshold alone.

Percentile	Fixed threshold	Selected by both	Excluded by fixed threshold	Excluded by percentile
p15	0.06	111,867	105,318	25,551
p15	0.08	145,573	71,612	55,189
p20	0.06	122,848	166,696	14,570
p20	0.08	163,807	125,737	36,955
p25	0.06	126,338	235,623	11,080
p25	0.08	177,722	184,239	23,040
p30	0.06	129,380	304,951	8,038
p30	0.08	184,219	250,112	16,543
p35	0.06	132,103	374,571	5,315
p35	0.08	187,640	319,034	13,122

The analysis showed that less restrictive combinations increased the number of selected crown segments, while stricter combinations reduced the final selection. Across the tested combinations, the fixed threshold had a stronger influence on the outcome than the percentile-based threshold. However, the percentile criterion was retained to ensure that annual growth was still evaluated relative to trees growing under similar ecological conditions within each stratum.

Since the aim was to identify trees with particularly low annual growth, the combination of the 15th percentile and fixed threshold of 0.06 m year⁻¹ was selected as the most appropriate.

To evaluate the stability of the thresholds for tree height and crown diameter, a sensitivity analysis was conducted comparing the selected 75th percentile with the neighboring 70th and 80th percentiles (Table 11). The number of crown segments within the 70-75 and 75-80 percentile intervals was almost identical for both the low- and high-resolution datasets. This indicates that the 75th percentile represented a balanced midpoint between less and more restrictive selections and was therefore retained.

Table 11. Sensitivity analysis of the 75th percentile for tree height and crown diameter in low- and high-resolution data. The table shows the number of crown segments meeting the 70th, 75th and 80th percentiles, as well as the number of segments within the adjacent percentile intervals 70-75 and 75-80.

Resolution	Variable	P70	P75	P80	70–75	75–80
Low	Height	248,731	208,351	167,319	40,380	41,032
High	Height	450,493	375,486	300,333	75,007	75,153
Low	Crown diameter	245,851	206,473	176,873	39,378	38,600
High	Crown diameter	447,362	374,600	298,329	72,762	76,271

The flatness variable was initially evaluated as a percentile-based criterion. Although a large proportion of crown segments passed the fixed biological threshold, many were excluded when applying the relative threshold at the 25th percentile. A targeted sensitivity analysis between the 25th and 30th percentiles showed that only a limited proportion of additional crown segments were included when increasing the percentile level (Table 12).

This indicates that the flatness variable became overly restrictive as a percentile-based criterion and that minor adjustments in the percentile level did not provide any clear improvement. The variable was therefore not included as a percentile-based classification criterion and was instead retained as a fixed quality filter to exclude structurally unrealistic or excessively flat crowns.

Table 12. Summary of the analysis for the flatness variable between the 25th and 30th percentile in low- and high-resolution data. The table shows the number of crown segments that passed the fixed biological flatness threshold, failed the p25 percentile threshold, but would be included when applying the p30 threshold. Additional share represents the percentage increase in retained crown segments when using p30 instead of p25.

Resolution	Total analyzed crowns	Pass fixed threshold	Failing p25 threshold	Included at p30	Additional share
High	1,501,934	1,460,788	1,127,392	80,799	7.2 %
Low	824,867	788,463	617,189	41,215	6.7 %

3.6 Validation methodology

This section describes the validation method used in the study. The validation was divided into two parts: validation of the crown segmentation and validation of the classification.

First, the crown segmentation was evaluated through visual comparison between segmented crown polygons and an orthomosaic generated from UAV imagery within selected test areas. The UAV survey was conducted at a flight altitude of 65 m with 80% forward overlap and 60% side overlap between images using GNSS-supported positioning.

In addition, the segmentation results were manually evaluated by visually comparing the segmented crowns with the orthomosaic and classifying the crown segments as correctly segmented, under-segmented, over-segmented or missed. Only crowns with fully visible segments were included in the evaluation. Under-segmentation referred to cases where a single segment covered multiple visually distinguishable tree crowns, while over-segmentation referred to cases where multiple segments represented a single crown. Missed crowns were visually identifiable crowns lacking a corresponding segment. The proportion of correctly segmented crowns was then calculated for each evaluation area and dataset.

The second part of the validation focused on classified candidate trees, which were compared with trees of high conservation value registered during field inventory. The inventory was conducted together with personnel with experience in identifying trees of high conservation value. Field inventories were carried out within selected sample plots, where trees identified as having conservation value were positioned using GNSS equipment and a tablet-based mapping application. For each recorded tree, structural characteristics such as fire scars, dead branches, crown shape and diameter at breast height (DBH) were documented in a field protocol.

The comparison was used to evaluate the agreement between the classified candidates and trees of high conservation value identified during the field inventory. Inventoried trees located within a classified candidate crown were considered correctly identified (true positives), while inventoried trees not overlapping a candidate crown were considered missed detections (false negatives). Candidate crowns not corresponding to inventoried trees were considered false positives. Precision, recall and F1-Score were then calculated based on these outcomes.

As a complementary validation, candidate density was compared with the density of inventoried trees of high conservation value within each inventory area. This was done to evaluate whether areas with a higher density of model-identified candidates also corresponded to areas with a higher density of inventoried trees, rather than only assessing individual tree-level matches.

4 Results

This section presents the results of the segmentation and classification analyses conducted in the study. First, the results of the watershed segmentation are presented, followed by the results of the classification. Finally, the results of the validation are presented.

4.1 Watershed-segmentation results

The first step in the watershed segmentation was the detection of local treetops in each CHM. This resulted in 1,650,154 identified treetops in the high-resolution CHM and 854,102 in the low-resolution CHM (Figure 16), showing a clear difference between the two datasets.

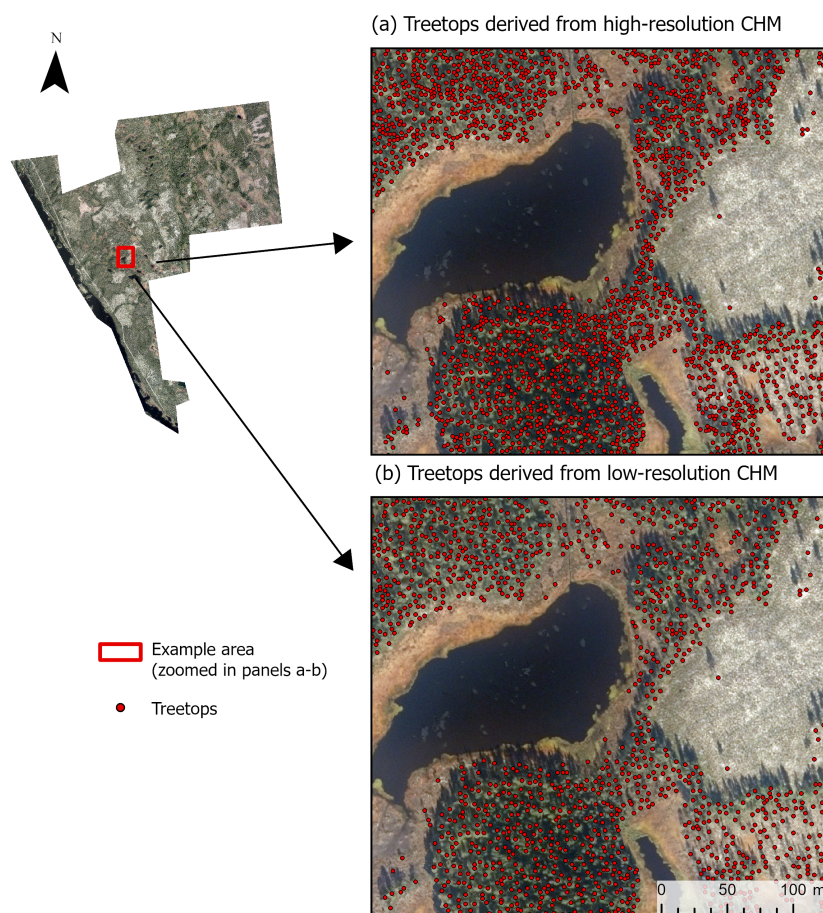


Figure 16. Identified treetops based on (a) the high-resolution CHM and (b) the low-resolution CHM in a sample area within the study area (shown in red on the overview map). Red points represent detected treetops. Background map data and orthophoto: © Lantmäteriet.

Based on these detected treetops, the marker-controlled watershed segmentation resulted in 1,650,112 segmented tree crowns for the high-resolution raster and 853,976 for the low-resolution raster. This corresponds to a tree density of approximately 355 trees per hectare for the high-resolution CHM and 184 trees per hectare for the low-resolution CHM.

To handle the large data volume, the high-resolution CHM required substantial computational resources and was therefore divided into tiles of 1000 x 1000 m. Out of a total of 120 tiles, 68 contained identified tree crowns, which shows that a large part of the study area consists of water, clear-cuts and open areas. The number of identified tree crowns per tile varied greatly, ranging from 96 in the least dense tile to 54,400 in the densest tile.

The high-resolution CHM identified almost twice as many individual tree segments as the low-resolution CHM. Differences between the segmentation results are further illustrated in Figure 17, where the same area is shown for both rasters. In the high-resolution CHM, the canopy appears more detailed and more individual tree crowns can be distinguished, whereas the crowns in the low-resolution appear larger and more merged.

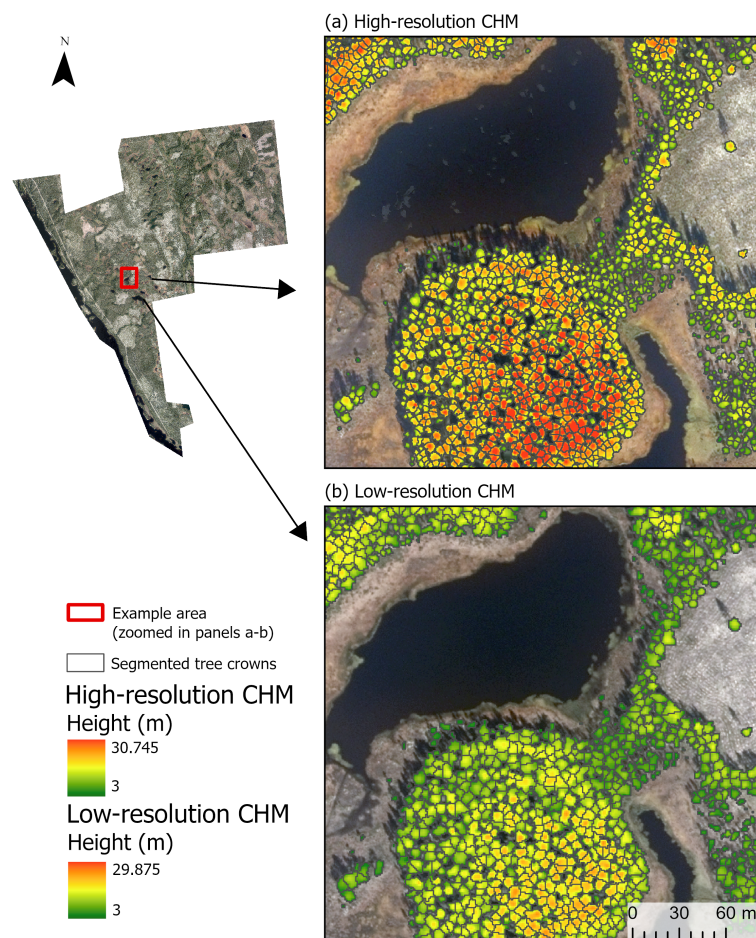


Figure 17. Comparison of segmentation results based on (a) the high-resolution CHM and (b) the low-resolution CHM in a sample area within the study area (shown in red on the overview map). Colors represent tree height (m), and the polygons show the segmented tree crowns. Background map data and orthophoto: © Lantmäteriet.

4.2 Classification results

The classification was applied to the segmented tree crowns. The results are presented below, focusing on the distribution of the classified candidates, differences between datasets, and factors influencing the classification outcome.

4.2.1 Classification outcome

The classification results showed clear differences between the datasets (Table 13). In the high-resolution data, 32,302 candidates were identified in category A_full, fulfilling all criteria, and 10,928 candidates in category B_partial, fulfilling two of the criteria. The remaining 1,462,668 were classified as non-candidates, as they satisfied only one or none of the criteria. Together, categories A and B constituted 2.87% of the entire dataset.

In the low-resolution data, 15,830 candidates were identified in category A_full, fulfilling all criteria, and 20,013 candidates in category B_partial, fulfilling two of the criteria. The remaining 800,577 were classified as non-candidates, as they satisfied only one or none of the criteria. Together, categories A and B constituted 4.29% of the entire dataset.

Table 13. Class distribution for the low- and high-resolution datasets. Category A refers to candidates that satisfy all three criteria, while B refers to candidates that satisfy two of the criteria. The table also shows the number of non-candidates and the proportion of A and B (%).

Dataset	A_full	B_partial	Non_candidate	A + B (%)
High	32,302	10,928	1,462,688	2.87
Low	15,830	20,013	800,577	4.29

The proportion of candidates was higher in the low-resolution dataset compared with the high-resolution dataset. The low-resolution dataset also showed larger and more merged crown segments, as. Observed in the segmentation results. Figure 18 illustrates how the classified candidate crowns were distributed within the example area for the two CHM resolutions. Differences in the density and spatial extent of candidate crowns can be observed between the datasets.

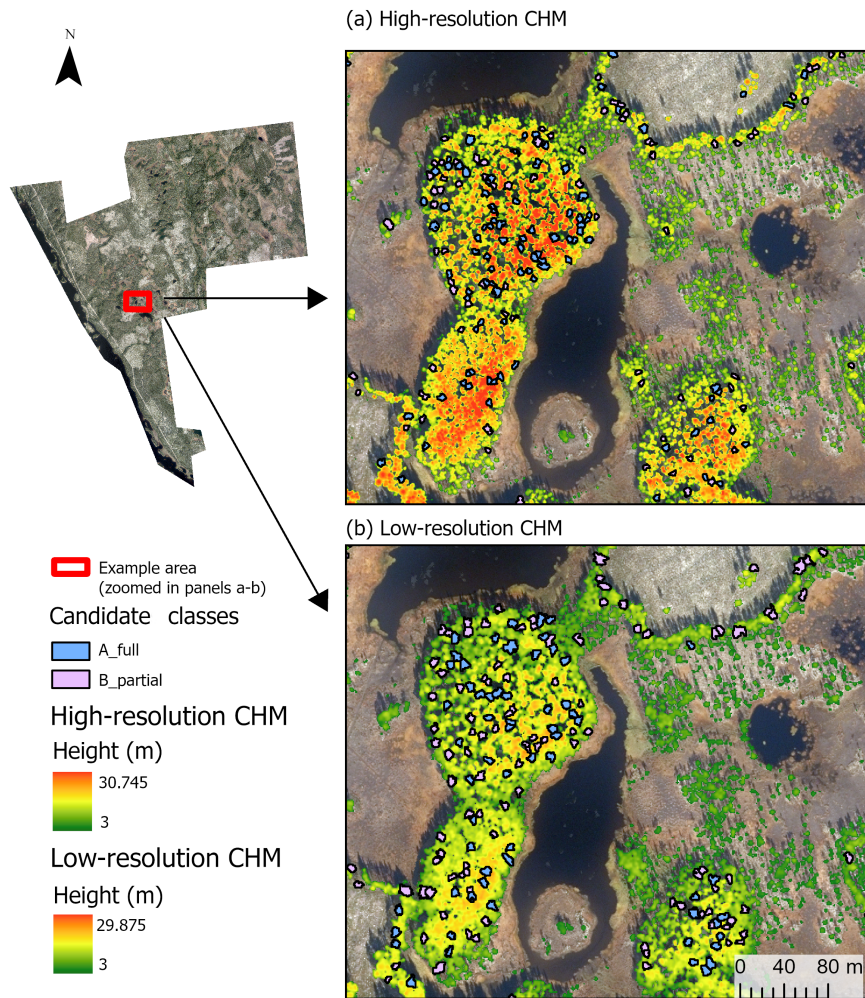


Figure 18. Example area showing the distribution of classified crowns in the (a) high-resolution CHM and (b) low-resolution CHM. Candidate crowns classified as A_full fulfilled all three criteria, while B_partial fulfilled two criteria. Background map data and orthophoto: © Lantmäteriet.

In the low-resolution dataset, 51 objects were flagged as potentially under-segmented crowns. Of these, 11 belonged to category A_full, 5 to category B_partial, and 35 to the non-candidate class (Table 14). In the high-resolution dataset, no crowns were flagged as potentially under-segmented by the applied filtering approach.

Table 14. Number of flagged and non-flagged objects in the low-resolution dataset. Flagged objects represent crowns with indications of under-segmentation. The table shows the distribution across categories A_full, B_partial, and non-candidates in the low-resolution dataset.

Class	Not flagged	Flagged
A_full	15,819	11
B_partial	20,008	5
Non_candidate	800,542	35

The absolute number of candidate trees varied between the strata in both datasets. In the high-resolution dataset (Figure 19), the highest absolute number of candidates occurred in the EH_mesic-moist_TM stratum, with 6,573 candidate trees. In the low-resolution dataset (Figure 20), the highest absolute number occurred in the EM_dry_mesic_TH stratum, with 4,524 candidate trees. However, EH_mesic-moist_TM, which had the highest count in the high-resolution dataset, also ranked high in the low-resolution dataset, with 4,363 candidates.

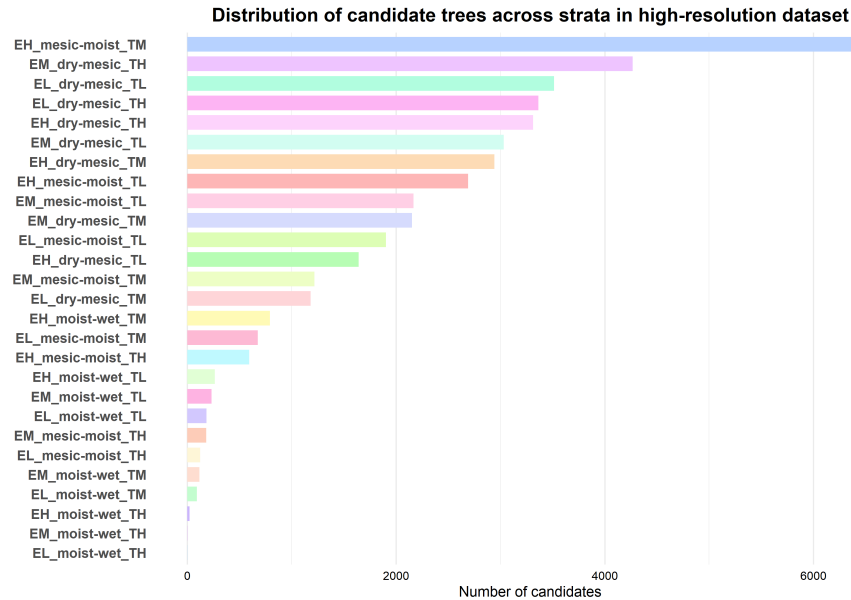


Figure 19. Distribution of candidate trees across the 27 strata in the high-resolution dataset. The strata are based on combinations of elevation, soil moisture, and TPI classes. The x-axis shows the number of candidate trees identified within each stratum.

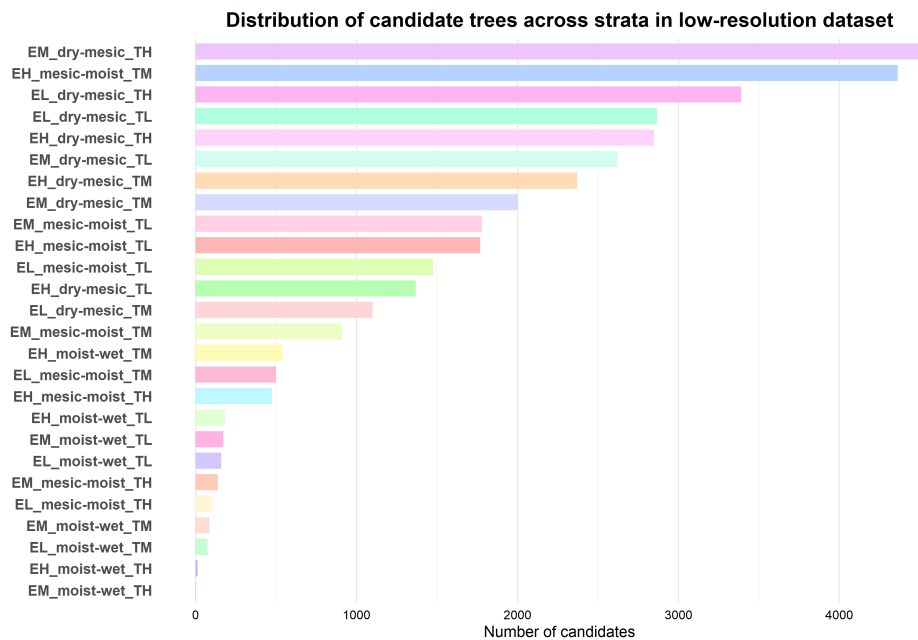


Figure 20. Distribution of candidate trees across the 27 strata in the low-resolution dataset. The strata are based on combinations of elevation, soil moisture, and TPI classes. The x-axis shows the number of candidate trees identified within each stratum.

The results indicate that candidate distribution patterns differed depending on the combinations of elevation, soil moisture, and TPI. For example, the EH_mesic-moist_TM stratum contained the highest absolute number of candidates in the high-resolution dataset, whereas other mesic-moist combinations contained substantially fewer candidates. Likewise, both dry-mesic and mesic-moist strata were represented among the most candidate-rich groups across different combinations of elevation and TPI classes, suggesting that no single environmental variable alone explained the observed distribution patterns.

However, the strata differed in total area, meaning that absolute candidate counts do not fully describe relative candidate occurrence. Larger strata may contain more candidates partly because they cover a greater area. Therefore, candidate density, expressed as candidates per hectare, was also considered to provide a more comparable measure across strata.

Overall, the density values generally supported the patterns observed in the absolute counts, as several strata with high candidate counts also showed relatively high candidate densities. Although EH strata did not consistently contain the highest absolute candidate counts, several EH strata showed comparatively high candidate densities in both datasets, particularly in the high-resolution dataset. Exact values for candidate count, strata area, and candidate density are provided in Appendix F and Appendix G

These results show that candidate occurrence varied between environmental strata and that candidate distribution was influenced by combinations of site conditions rather than a single environmental variable.

4.2.2 Model selectivity

The score distribution shown in Table 15 illustrates how many objects satisfied zero to three criteria in each dataset. In both datasets, most objects had a score of 0, meaning that no criteria were satisfied. The high-resolution dataset contained more objects that satisfied all three criteria, while the low-resolution dataset had more objects that satisfied two criteria.

Table 15. Distribution of satisfied criteria for the low- and high-resolution datasets. The score represents the number of criteria satisfied by each object, ranging from 0 to 3.

Dataset	0	1	2	3
High	1,455,157	777	10,928	32,302
Low	768,074	11,993	20,013	15,830

4.2.3 Characteristics of classified trees

The classified trees showed some differences in height, diameter and height growth (Table 16). In both datasets, category A_full had a higher average height than category B_partial. Category A_full also had a larger average diameter than category B_partial. Category B_partial had a slightly higher average growth rate than category A_full in both datasets.

Table 16. Mean height, standard deviation (SD) of height, mean crown diameter and mean annual height growth for categories A and B in the low- and high-resolution datasets.

Dataset	Class	Height (m)	Height SD (m)	Crown diameter (m)	Growth (m year-1)
High	A_full	16.4	2.24	5.81	0.0349
High	B_partial	13.9	1.77	5.48	0.0394
Low	A_full	16.7	2.17	7.57	0.0277
Low	B_partial	15.4	2.18	6.44	0.0306

Overall, trees classified as A_full were characterized by greater height, larger crown diameters and lower annual height growth compared with B_partial within both datasets. This low-resolution dataset showed larger average crown. Diameters compared with the high-resolution dataset.

4.2.4 Influence of height growth filtering and environmental stratification

As part of the classification process, annual growth values were calculated for the segmented tree crowns in both CHM datasets. Valid growth values were obtained for the majority of crowns in both datasets, while crowns with negative or unrealistic growth values were assigned missing values after filtering and therefore could not satisfy the growth criterion during classification. This filtering step influenced the final classification outcome and should be considered when interpreting the number of crowns fulfilling all criteria.

In addition to growth filtering, stratification was applied to account for variation in environmental conditions across the study area, as described in Section 3.5. This resulted in 27 distinct strata based on combinations of elevation, soil moisture and TPI classes. Figure 21 shows the spatial distribution of these strata, illustrating that the study area consisted of a heterogeneous landscape where different combinations of elevation, soil moisture and topographic conditions occurred. This variation shows that trees included in the classification were located under different environmental conditions.

The stratified classification allowed percentile-based thresholds to be calculated separately within each environmental group rather than applying a single threshold across the entire study area. As a result, candidate trees were evaluated relative to trees growing under similar site conditions, allowing structural characteristics to be interpreted within their environmental context.

Strata groups

- EH_dry-mesic_TH
- EH_dry-mesic_TL
- EH_dry-mesic_TM
- EH_mesic-moist_TH
- EH_mesic-moist_TL
- EH_mesic-moist_TM
- EH_moist-wet_TH
- EH_moist-wet_TL
- EH_moist-wet_TM
- EL_dry-mesic_TH
- EL_dry-mesic_TL
- EL_dry-mesic_TM
- EL_mesic-moist_TH
- EL_mesic-moist_TL
- EL_mesic-moist_TM
- EL_moist-wet_TH
- EL_moist-wet_TL
- EL_moist-wet_TM
- EM_dry-mesic_TH
- EM_dry-mesic_TL
- EM_dry-mesic_TM
- EM_mesic-moist_TH
- EM_mesic-moist_TL
- EM_mesic-moist_TM
- EM_moist-wet_TH
- EM_moist-wet_TL
- EM_moist-wet_TM

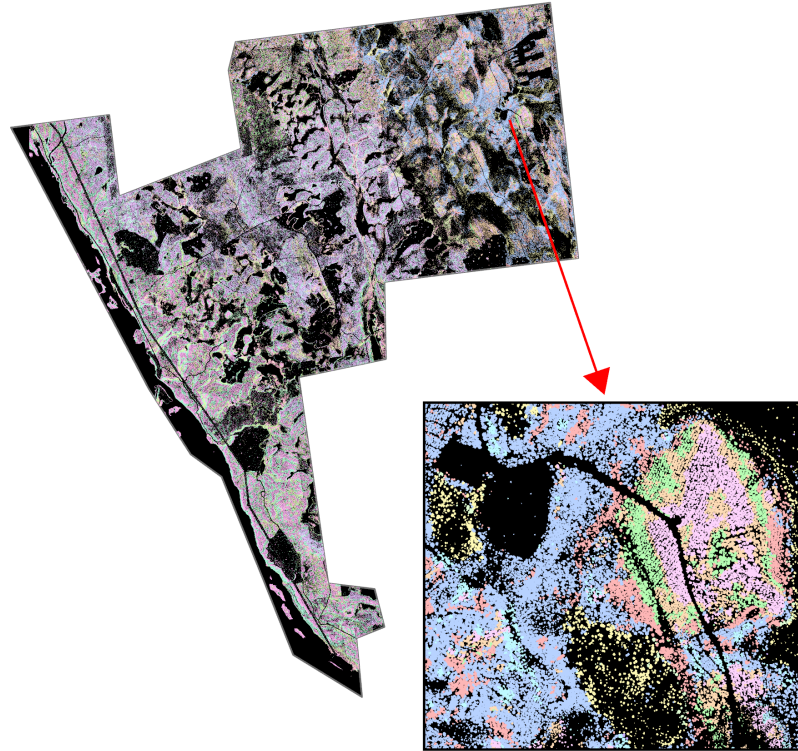


Figure 21. Spatial distribution of the 27 strata used in the classification, based on combinations of elevation (E), soil moisture, and TPI classes (T). Elevation and TPI classes are represented as low (L), medium (M), and high (H). Shown for the high-resolution dataset. Different colors represent different strata.

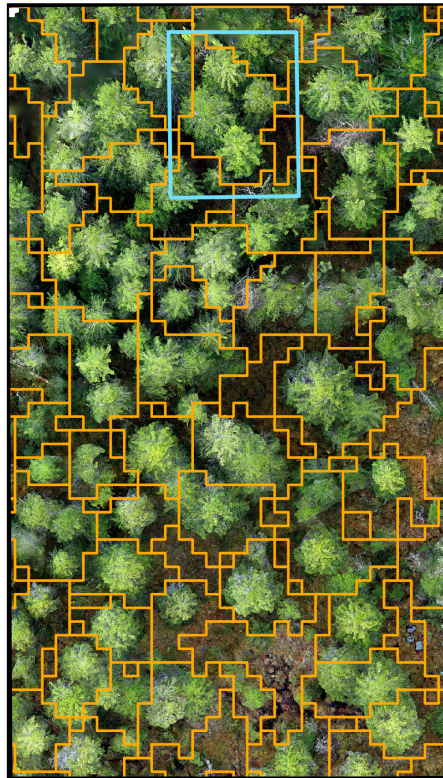
4.3 Validation of the results

Figure 22 shows examples of crown segmentation results from the high- and low-resolution CHM datasets overlaid on the UAV orthomosaic. Visual differences between the datasets can be observed, where the high-resolution CHM generally produced segments that followed individual crown boundaries more closely. In contrast, the low-resolution CHM more often produced larger segments that covered multiple crowns.

(a) High-resolution segmentation



(b) Low-resolution segmentation



(c) Detailed example, high-resolution



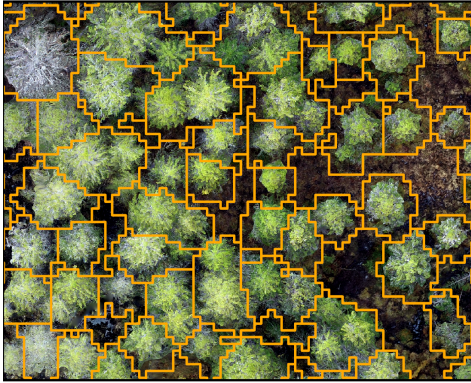
(d) Detailed example, low-resolution



Figure 22. Crown segmentation results from the (a) high-resolution CHM and (b) low-resolution CHM within the same sample area, overlaid on an orthomosaic generated from UAV imagery. Orange polygons represent segmented tree crowns. Blue rectangles indicate the detailed example areas shown in panels (c) and (d), where differences in crown delineation between the two spatial resolutions can be observed more clearly.

In addition to the evaluation area shown in Figure 22, three smaller evaluation areas were included to further assess the crown segmentation results under varying forest conditions. These areas are presented in Figure 23.

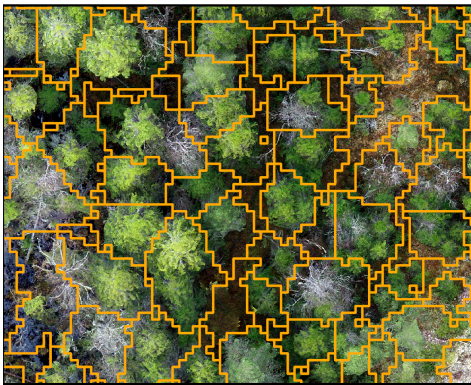
Evaluation area 2, High-resolution



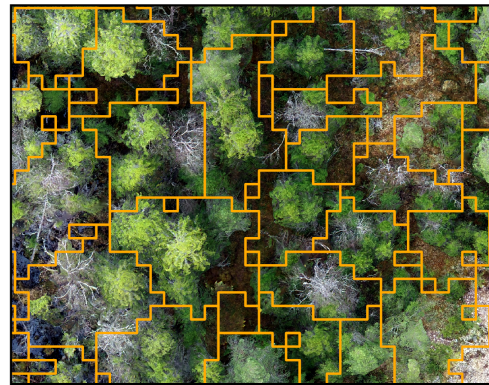
Evaluation area 2, Low-resolution



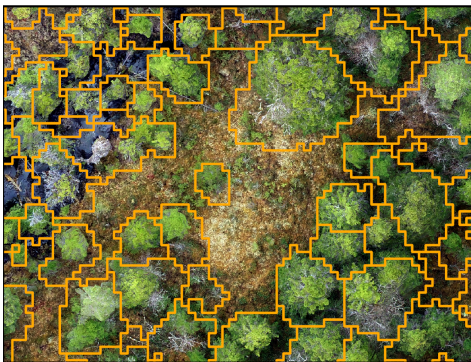
Evaluation area 3, High-resolution



Evaluation area 3, Low-resolution



Evaluation area 4, High-resolution



Evaluation area 4, Low-resolution

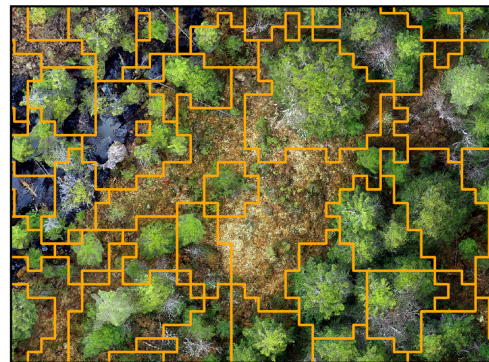


Figure 23. Additional evaluation areas used for assessment of crown segmentation results. The figure shows segmented crown polygons derived from the high-resolution and low-resolution CHM datasets overlaid on an orthomosaic imagery. The areas were selected to represent varying forest structures with different structural complexities.

Manual evaluation of the crown segmentation showed clear differences in segmentation performance between the high- and low-resolution CHM datasets across the evaluation areas (Table 17). In all evaluation areas, the high-resolution CHM resulted in a higher proportion of correctly segmented crowns compared with the low-resolution CHM.

The high-resolution dataset generally produced crown segments that corresponded more closely to individual trees, while the low-resolution dataset more frequently resulted in under-segmented crowns where neighboring trees were merged into single segments. Under-segmentation was the dominant segmentation error across all evaluation areas, particularly in the low-resolution dataset and in more structurally complex forest conditions.

Since the two CHM resolutions produced different numbers of crown segments within the evaluation areas, the total number of evaluated crowns differed between the datasets.

Table 17. Results from the manual evaluation of crown segmentation for the high- and low-resolution CHM datasets within the same evaluation areas. Evaluation area 1 is shown in panels (a-b) in Figure 22, while additional evaluation areas (2–4) for both resolutions are presented in Figure 23.

Area	Dataset	Correct	Underseg.	Overseg.	Missed	Total	Accuracy (%)
1	High	45	20	0	0	65	69.2
1	Low	10	35	0	0	45	22.2
2	High	27	16	0	0	43	62.8
2	Low	4	23	0	0	27	14.8
3	High	22	21	0	0	43	51.2
3	Low	2	21	0	0	23	8.7
4	High	14	18	0	1	33	42.4
4	Low	6	17	0	2	25	24

Overall, the results showed differences in segmentation performance between the two CHM resolutions. The high-resolution CHM resulted in a larger number of delineated crown segments and higher segmentation accuracy compared with the low-resolution CHM. Lower segmentation accuracies were observed in the more structurally complex evaluation areas, particularly in the low-resolution dataset.

Due to the substantially higher segmentation accuracy observed for the high-resolution CHM, the following presentation of the classification validation focuses on the high-resolution dataset.

Figure 24 illustrates the spatial distribution of classified candidate crowns and inventoried trees of high conservation value within the four field inventory areas used for validation. Candidate crowns classified as A_full or B_partial were identified in all inventory areas, although their abundance and distribution varied between sites. The inventoried trees were unevenly distributed across the areas, with some sites containing more observations than others.

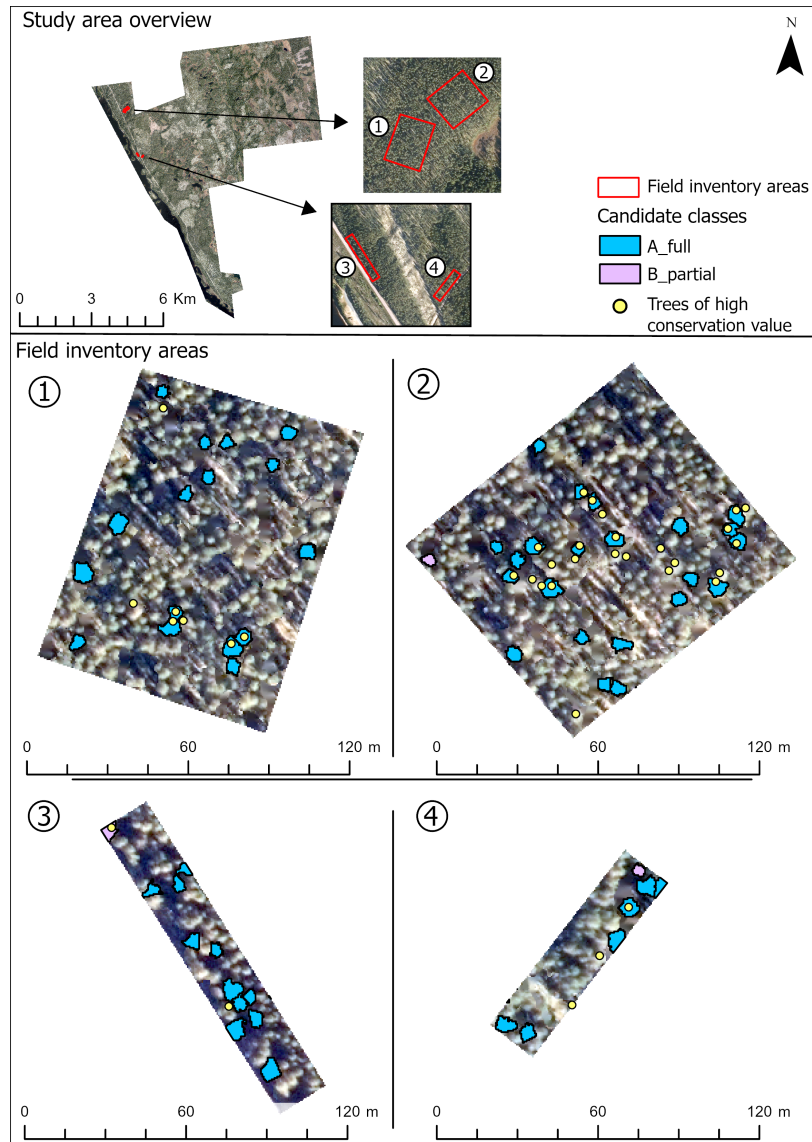


Figure 24. Overview of the field inventory areas used for validation. The upper panel shows the location of four inventoried areas within the study area. The lower panels show the distribution of classified candidate crowns and inventoried trees of high conservation value within each field inventory area. Background map data and orthophoto: © Lantmäteriet.

The classification performance varied between the four inventory areas (Table 18). Recall ranged from 0.333 to 0.571, with the highest value observed in area 2 and the lowest in area 4. Precision was generally low across all areas, ranging from 0.083 to 0.478, indicating that a relatively large proportion of the classified candidate crowns did not correspond to inventoried trees of high conservation value.

The F1-Score reflected this pattern, with values between 0.142 and 0.455, where area 1 showed the highest overall balance between recall and precision. When aggregating all inventory areas, the model achieved an overall recall of 0.472, a precision of 0.293, and an F1-Score of 0.362.

Table 18. Validation results for the four inventory areas showing true positives (TP), false positives (FP), false negatives (FN), recall, precision and F1-score.

Inventory area	TP	FP	FN	Recall	Precision	F1-Score
1	11	12	13	0.458	0.478	0.468
2	4	12	3	0.571	0.250	0.348
3	1	11	1	0.500	0.083	0.142
4	1	6	2	0.333	0.143	0.200
Total	17	41	19	0.472	0.293	0.362

The relatively low precision values should be interpreted with caution, as the field inventory data only included trees identified as having high conservation value. Consequently, some classified candidate crowns that did not match an inventoried tree may still represent trees with old structural characteristics that were not recorded during field inventory.

To complement the individual tree-level validation, a density-based comparison was therefore performed (Table 19). The number of classified candidate crowns per hectare was compared with the number of inventoried trees of high conservation value per hectare within each inventory area. This comparison was used to evaluate whether areas with a higher concentration of model-identified candidates also corresponded to areas with a higher density of inventoried high conservation value trees.

Table 19. Density-based comparison between inventoried trees of high conservation value and classified candidate crowns within the four inventory areas. Values are presented as number of trees/candidates per hectare.

Inventory area	Area size (ha)	Inventoried trees / ha	Candidates / ha
1	1.0	7	16
2	1.0	24	23
3	0.25	8	48
4	0.16	18.75	43.75

The density-based comparison showed that inventory areas with higher densities of inventoried trees generally also contained higher densities of classified candidate crowns. Inventory area 1 had the lowest density of both inventoried trees and classified candidates, while areas 2-4 showed higher values for both variables.

Inventory area 2 showed particularly similar densities of inventoried trees and classified candidates. In contrast, inventory areas 3 and 4 showed higher densities of classified candidates compared with inventoried trees.

This suggests that the model identifies a broader set of trees with structural characteristics associated with older forest conditions than those recorded as trees of high conservation value during field inventory.

As a complementary validation, a DVIS layer representing predicted forest conservation values was visually compared with the spatial distribution of classified candidate crowns. DVIS is a machine-learning-based dataset produced by the Swedish Forest Agency, where higher values indicate a stronger likelihood of forest conservation value (Skogsstyrelsen, 2026a). Figure 25 shows example areas with high and low DVIS values together with the classified candidate crowns.

The observed correspondence between candidate distribution and DVIS patterns further supports the potential use of the model for highlighting areas with structurally old trees and potentially higher conservation values.

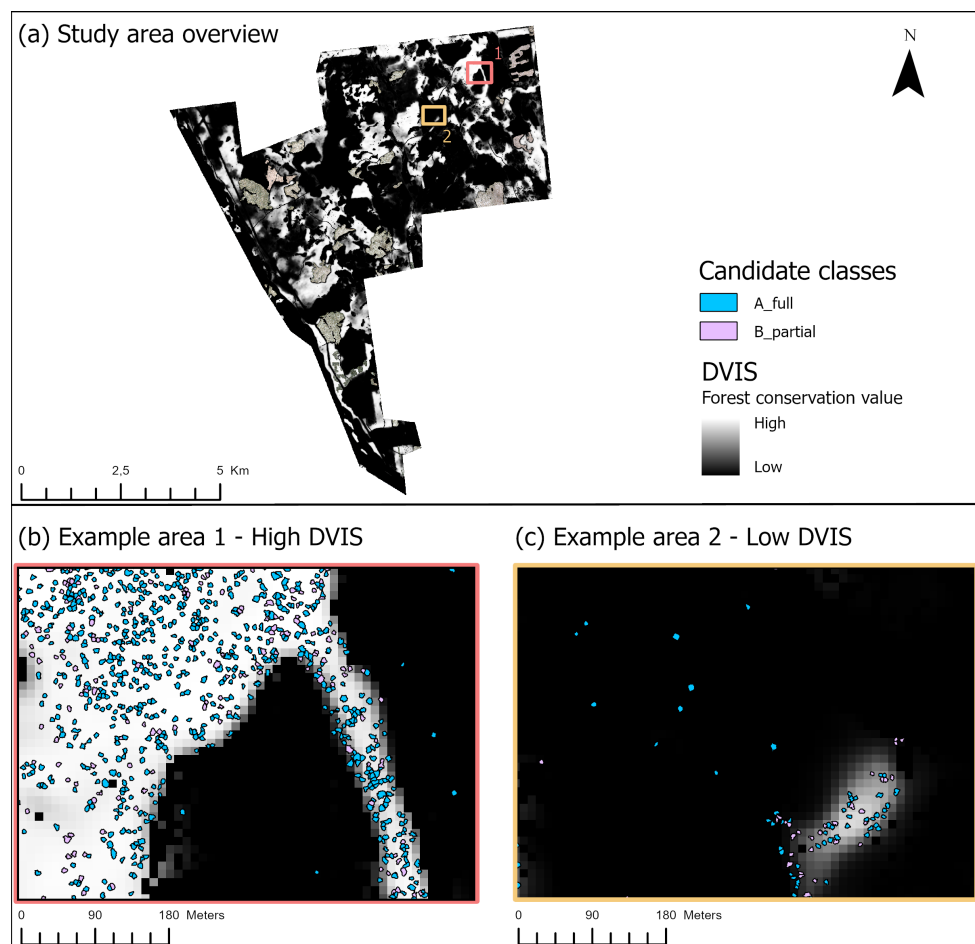


Figure 25. Comparison between the spatial distribution of classified candidate crowns and the DVIS layer. (a) Study area overview showing the location of the example areas. (b) Example area with high DVIS values. (c) Example area with low DVIS values. DVIS data: © Skogsstyrelsen.

5 Discussion

This chapter discusses the results of the study, as well as the strengths and limitations identified in the analysis. The discussion is divided into three parts: main findings, methodology and limitations. Finally, the study's relevance in relation to environmental and sustainable development aspects is addressed, and potential directions for future work are outlined.

5.1 Main findings and interpretation

The results demonstrated that spatial resolution strongly influenced both the segmentation of individual tree crowns and the subsequent classification of Scots pine trees with old-growth structural characteristics. The high-resolution CHM generally improved crown delineation and enabled a more detailed representation of crown structure, while the low-resolution CHM produced a more generalized canopy representation with greater under-segmentation. These findings directly address the first research question and demonstrate that spatial resolution substantially affects both segmentation accuracy and the interpretation of structural crown characteristics in CHM-based forest analysis.

The manual segmentation evaluation further illustrated these differences between the datasets, showing substantially higher segmentation accuracy for the high-resolution dataset and identifying under-segmentation as the dominant error type, particularly in structurally complex forest areas. This suggests that higher spatial resolution is especially important in heterogeneous forest environments where crown overlap and variation in canopy height increase the difficulty of separating individual trees. The results therefore indicate that spatial resolution influences not only the number of detected crowns, but also how forest structure is represented and interpreted in CHM-based analyses.

The differences in segmentation between the two datasets also influenced the classification results. A clear difference was that the low-resolution dataset resulted in a higher proportion of classified candidates compared to the high-resolution dataset. This may be because the lower resolution provides a more generalized representation of the canopy surface and a higher degree of under-segmentation, which reduces the small-scale variations, increasing the likelihood that a crown fulfills multiple classification criteria. In the high-resolution dataset, more detail and structural variation were preserved within each crown segment, which may have resulted in fewer crowns fulfilling all criteria simultaneously.

At the same time, the high-resolution dataset contained a substantially larger number of segments classified as A_full, fulfilling all three criteria. This may indicate that higher spatial resolution enables a more detailed representation of crown structure and therefore improves the conditions for identifying trees with distinct structural characteristics. The differences between the datasets therefore appear to be related not only to the number of identified candidates, but also to how crown structure was represented and interpreted during the classification process.

Only a limited number of potentially under-segmented trees were identified and flagged in the low-resolution data. However, the subsequent validation showed that under-segmentation occurred in both datasets. This also suggests that the selected threshold of 130 m², which was based on literature values for maximum crown size, may have limited the detection of under-segmented crowns and resulted in several remaining unflagged crown segments that were under-segmented. Since the study area was characterized by generally small and slow-growing trees, the applied threshold may have been too high in relation to the local crown sizes.

The validation results further showed that precision values were generally low, while recall values were comparatively higher. This likely reflects the difference between the model objective and the field inventory data. The model was designed to identify trees with structural characteristics associated with structurally old trees, whereas the field inventory focused specifically on trees of high conservation value. For example, trees may have been identified as having high conservation value based on features such as fire scars or cavities, even if they lacked the larger crown dimensions or structural characteristics targeted by the model. Consequently, classified candidate crowns may represent structurally old trees that were not recorded as trees of high conservation value during the inventory.

Although precision values were low, the comparatively higher recall values indicate that a substantial proportion of inventoried trees of high conservation value were still captured by the classification. This supports the potential of the workflow as a screening or prioritizing tool, where the objective is to identify forest areas with higher likelihood of containing structurally old trees rather than to achieve exact detection of every individual tree.

The density-based validation and visual comparison with the DVIS layer further supported this interpretation by indicating a spatial correspondence between classified candidate crowns and areas with higher conservation values. The results therefore support the potential of the workflow as decision support for identifying forest areas likely to contain structurally old trees.

To better understand these results, the methodological choices and their influence on the outcome are discussed in the following section.

5.2 Methodological discussion

The methodological choices in this study affected the results at different steps in the workflow. This section explains how parameter selection, spatial resolution and data handling influenced individual tree detection and classification and discusses the strengths and limitations of the method.

5.2.1 Parameter sensitivity and preprocessing

To assess the robustness of the method, several parameters and preprocessing steps were tested to improve treetop detection and crown segmentation. The sensitivity analyses showed that buffer size and the VWF filter parameters produced stable results within the tested intervals, whereas other preprocessing choices, such as smoothing method, thinning distance and order of thresholding and smoothing had a greater influence on the segmentation outcome.

This highlights the importance of carefully evaluating preprocessing steps and parameter settings in CHM-based individual tree detection, as they directly influence both the number of detected treetops and the spatial delineation of crown segments. The systematic evaluation of preprocessing steps and parameter settings therefore strengthened the robustness and reliability of the applied workflow. The results therefore demonstrate that preprocessing should not be regarded as a purely technical step, but as a methodological component that can substantially influence both segmentation outcomes and the ecological interpretation of the final classification.

The influence of preprocessing choices was further illustrated by the analysis of thinning distance. Differences between thinning distances of 0.5, 0.8 and 1.1 meters produced only minor differences in the low-resolution CHM. This reflects the raster spatial resolution of 1 m, meaning that the tested thinning distances fall within the size of a single pixel and therefore have limited influence on the canopy representation. In contrast, the high-resolution CHM showed greater sensitivity to the thinning distance, likely because the finer spatial resolution preserved more small-scale structural variation in the canopy surface.

5.2.2 Influence of spatial variation and resolution

Higher spatial resolution allows small variations in the elevation surface to be represented in greater detail, increasing the likelihood that multiple local maxima are detected within the same tree crown. As a result, the thinning distance becomes more important for merging nearby local maxima belonging to the same tree in the high-resolution CHM. The sensitivity analysis therefore primarily focused on the high-resolution data, highlighting that the role of the thinning parameter is closely linked to the spatial resolution of the CHM.

In addition to spatial resolution, variation within the study area also influenced the results. The number of detected treetops varied substantially between the tiles, indicating the presence of both dense forest stands and more open areas. Such structural variation affects segmentation performance, as overlapping crowns in dense stands make it more difficult to distinguish individual trees. This suggests that segmentation performance is influenced by local forest structure, since the same methodological settings may perform differently in dense forest stands compared with more open forest environments.

The observed differences between the two CHMs further emphasize the importance of data quality and consistency throughout the analysis, as spatial resolution influences how canopy structure is represented. In the high-resolution CHM, finer spatial detail improved the separation of closely spaced trees. In contrast, the smoother canopy representation in the low-resolution CHM may increase the risk of under-segmentation, where multiple trees may be represented by a single crown segment. This pattern is consistent with previous studies showing that spatial resolution affects the detection of individual trees in LiDAR-based analyses (Vauhkonen et al., 2012). However, it should be noted that the compared datasets also differed in other aspects, such as acquisition year, point density and data source, which may have contributed to the observed differences.

5.2.3 Uncertainty in height growth estimation

The estimation of height growth introduced some uncertainty in the results, as both negative and unrealistic values were identified. Negative growth occurred when the height extracted for a crown polygon was lower in the later CHM raster than in earlier datasets.

These deviations can be explained by a combination of real-world changes, methodological limitations and differences between datasets. In some cases, the original tree may have been felled or fallen between measurement occasions, which means that the crown segment no longer represents the same tree. However, visual and statistical examination of the results also revealed cases where the spatial extent of the crown segments did not correspond between time points (Figure 26). This highlights the importance of combining quantitative analysis with visual inspection, as several of the identified deviations would otherwise have been difficult to detect and interpret correctly.

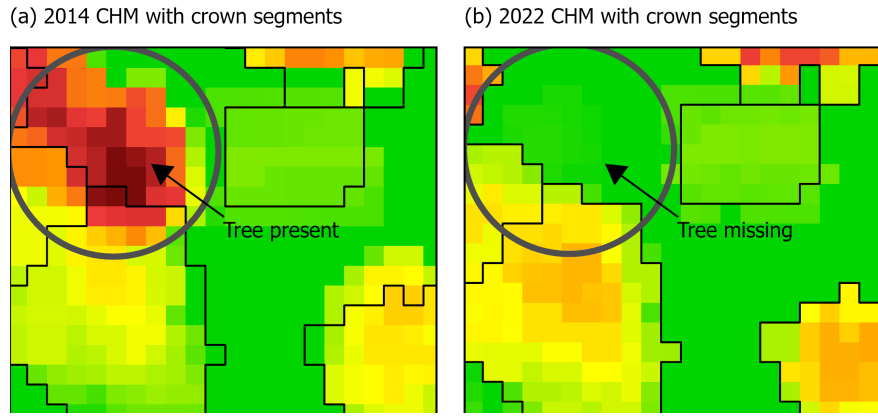


Figure 26. Example of mismatched crown segments between two different times. The black outlines show crown segments derived from the 2022 CHM, overlaid on both the 2014 CHM (a) and the 2022 CHM (b). The circled area shows a tree present in 2014 (a) that is no longer present in 2022 (b). The crown segment in 2022 still covers this area, which means that height values from 2014 can be extracted from a tree that no longer exists.

This issue is particularly relevant for the high-resolution CHM, where no earlier dataset with comparable spatial resolution was available. Therefore, height growth was estimated by combining crown polygons derived from the 2024 high-resolution CHM with height values extracted from the 2014 low-resolution CHM. As a result, differences in spatial detail further increased the likelihood of segment mismatch and contributed to unrealistic growth values.

In addition, height was initially extracted as the maximum value within each segment, which increased sensitivity to outliers and mismatches. As shown in Figure 27, this could lead to cases where the height from 2014 originated from a different tree than the one represented in later datasets, resulting in artificially high or negative growth values.

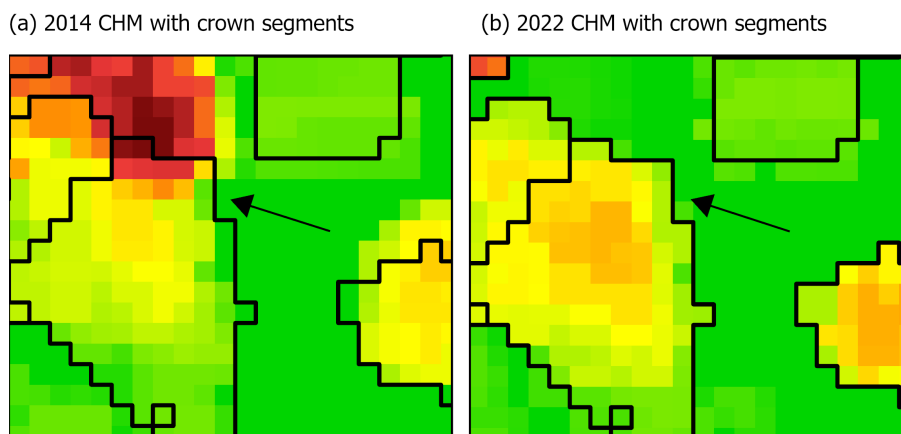


Figure 27. Example showing how maximum height extraction can be affected by overlapping neighboring crowns. The black outlines show crown segments derived from the 2022 CHM, overlaid on both the 2014 CHM (a) and the 2022 CHM (b). The arrow indicates an area where high height values from a neighboring crown overlap the crown segment in the 2014 CHM (a). In the 2022 CHM (b), the neighboring tree is no longer present, meaning that this overlap does not occur. This resulted in a higher extracted maximum height in 2014 and could therefore produce an artificially low or negative height growth estimate.

To reduce the impact of these sources of error, the 95th percentile of height values within each crown segment was used instead of the maximum value, as this approach is less sensitive to extreme values and inconsistencies in segment delineation. Furthermore, negative total growth values and total growth values exceeding 6 m over the study period were excluded from the final analysis, as these were considered unlikely to represent actual tree development.

Overall, the results demonstrate that height growth estimates are highly sensitive to both segment consistency and differences in spatial resolution between datasets. Since height growth was one of the variables used in the classification, these uncertainties may also have influenced which trees were ultimately identified as candidates with structural characteristics associated with older Scots pine trees.

This highlights the importance of carefully considering methodological choices when using height growth as a variable in classification, as uncertainties in growth estimates may directly influence the classification results. In addition, the generally low growth conditions in the study area make interpretation more challenging, as small changes in tree height over time may be difficult to distinguish from variations introduced by data resolution, segmentation and processing methods.

5.2.4 Classification approach and method selection

The choice of classification approach had an important influence on the final classification outcome. LIFT was initially considered because it could identify combinations of structural variables that co-occur more frequently than expected. However, in this study the method was too restrictive and resulted in very few identified trees. This indicates that methods requiring several strong variable associations simultaneously may be less suitable when the aim is to identify a broader range of trees with structural characteristics associated with older Scots pine.

The percentile-based approach provided a more flexible alternative. By defining thresholds within environmental strata, the classification could account for local variation in growing conditions rather than applying the same threshold values across the entire study area. This was particularly important because structural characteristics such as height, crown diameter and height growth may vary depending on site conditions.

Although flatness was initially considered as a complementary classification variable, it contributed less to the percentile-based classification than expected. The variable showed limited variation between percentile thresholds, which reduced its usefulness for identifying relative structural differences between trees. However, flatness still served an important function as a quality filter. Very low values were observed for segments that probably did not represent tree crowns, such as roof structures or other non-vegetation objects that were captured during the segmentation. Flatness was

therefore retained as a fixed threshold to exclude segments with unrealistically flat vertical structures.

The combination of stratified percentile-based and fixed thresholds helped balance ecological interpretation with structural realism. The percentile thresholds enabled relative comparisons within strata and allowed trees to be evaluated in relation to local site conditions. In contrast, the fixed thresholds acted as biological, structural and quality-related constraints, ensuring that the selected segments met minimum structural requirements. This combination likely increased the robustness of the classification compared to relying solely on either absolute or relative thresholds.

At the same time, the use of fixed thresholds may also have excluded some trees with less typical characteristics that could still hold ecological value. The classification should therefore be interpreted as a simplified representation of structural patterns associated with older Scots pine trees rather than a complete identification of all ecologically valuable individuals. This reflects a broader challenge in remote sensing-based ecological classification, where complex ecological qualities often must be approximated using a limited set of measurable structural variables.

5.2.5 Influence of stratification and environmental variables

In addition to the choice of classification method, the results were also influenced by how the data were stratified based on environmental variables. By stratifying the data according to conditions such as elevation soil moisture and topographic position, threshold values could be adapted to local variation. By evaluating trees within environmentally similar conditions, the classification could better account for natural variation in productivity and stand structure.

This can be seen as a strength of the method, as it reduces the risk of trees in less productive environments being underestimated and trees in more favorable environments being overestimated due to natural differences in height growth and structure. Stratification therefore played an important role in enabling the classification to account for varying structural conditions within the study area, which was central to the second research question regarding the identification of structurally old trees under different environmental conditions.

However, stratification also contributed to a simplification of continuous environmental gradients. Variables such as elevation and topographic position were divided into a limited number of classes, which improves clarity and reproducibility but may result in similar trees being assigned to different strata if they are located near class boundaries. In addition, the selection of environmental variables represents a trade-off between ecological detail and the complexity of the method, as including more variables would increase complexity and potentially lead to smaller groups with fewer observations.

The use of TPI as an environmental variable requires consideration in forest environments. Previous research has shown that the measure is sensitive to both scale and landscape complexity (De Reu et al., 2013), meaning that different radii capture different types of topographic variation and that similar TPI values may represent different topographic conditions depending on the surrounding landscape. TPI was therefore combined with the other stratification variables rather than used as an independent indicator, reducing the influence of this sensitivity while still accounting for topographic variation within the study area.

The selected radius of 40 m appeared to provide a balance between capturing broader environmental patterns and avoiding excessive influence from small local topographic variation. Smaller radii likely reflected fine-scale terrain variability to a greater extent, whereas larger radii resulted in only minor changes in the TPI classification. However, the selected radius remains a methodological choice that may influence how topographic variation is represented.

Finally, the cut-off values for elevation- and TPI-based strata were calculated separately for the low- and high-resolution datasets. This allowed the stratification to reflect the distribution and variability within each dataset, strengthening the internal relevance of the classification. However, because the thresholds were derived independently for each dataset, the resulting classifications were relative to the conditions within each resolution rather than based on identical absolute thresholds. This should be considered when comparing the results between spatial resolutions.

5.3 Limitations

This study has several limitations related to data availability, methodological choices and validation procedures. These limitations are discussed below.

5.3.1 Data and practical limitations

Because the analysis relied solely on CHM-derived variables, the study was limited to structural representations of tree crowns rather than direct ecological observations. This affects how the results should be interpreted, since ecologically valuable trees may contain important biological structures that are not detectable in CHM data.

The use of CHM also contributed to limitations during the classification of structurally old trees and relating the classified candidates to trees of high conservation value. Trees of high conservation value may be identified as such solely because of structures such as fire scars, cavities or dead branches, which cannot be derived from CHM data. As a result, some trees of high conservation value may not exhibit the structural characteristics targeted by the classification, while trees showing these characteristics may not necessarily be identified as trees of high conservation value in the field inventory. This highlights an important distinction between structurally derived remote sensing indicators and field-based ecological assessments of conservation value.

Another limitation related to the use of CHM is that segmentation is based on height variation rather than object type. As a result, elevated non-vegetation objects such as roof surfaces may also be detected and included in the segmentation. This limitation motivated the use of additional filtering steps, to reduce the inclusion of segments unlikely to represent tree crowns.

Due to the large data volumes, some adjustments to the workflow were necessary to manage memory usage and processing time efficiently. This included dividing the data into smaller tiles and processing them step by step. To improve computational efficiency, zonal statistics were used instead of direct extraction of raster values, since direct extraction proved to be considerably more computationally demanding. These adjustments enabled the analyses to be completed without substantially affecting the results, although they may in some cases have influenced the level of detail. At the same time, this provided valuable insights into how large spatial datasets can be managed efficiently.

Finally, determining appropriate classification thresholds was challenging. In the absence of field-based reference data, thresholds were based on the current data distributions and previous studies. However, structural characteristics and growth patterns may vary between study areas depending on local environmental conditions. In the present study area, tree height growth was generally low, meaning that trees with ecological value may not necessarily correspond to the structural dimensions reported in other studies.

Although the workflow itself is reproducible, the results highlight the importance of adapting fixed threshold values to local forest conditions within each study area. Since the study area consisted of boreal forests characterized by relatively low productivity and slow tree growth, the results may not be directly transferable to forest environments with substantially different conditions or growth dynamics. A more robust approach would therefore have included field visits to identify trees with known structural characteristics and use these observations as a basis for selecting locally relevant fixed threshold values.

5.3.2 Methodological limitations

The classification and validation approaches also involved methodological limitations. The classification is based on a scoring system that combines several criteria, which simplifies the complex structural characteristics of the trees. This may lead to trees that have similar characteristics being classified differently depending on small variations in the variables, especially when values are close to the threshold limits.

These uncertainties are further influenced by the segmentation, as the classification relies on delineated crown polygons. Incorrectly delimited crowns can therefore lead to uncertain values for the structural variables and thereby affect the classification. Uncertainties introduced during the segmentation step may propagate to the final classification results.

The data obtained during field inventory showed that most of the field-inventoried trees that were not classified by the model did not meet the height growth criterion, indicating that the height growth variable had a major impact on the classification result. As discussed in Section 5.2.3, the growth for high-resolution CHM from 2024 was calculated using height values extracted from the low-resolution CHM from 2014, as no older dataset with equivalent resolution was available.

Differences in spatial resolution between these datasets may therefore have affected the agreement between crown segments and elevation values. This in turn may have contributed to some trees being excluded from the classification due to the height growth variable, even though they simultaneously exhibited other structural characteristics linked to older pines. This suggests that differences in spatial resolution between the datasets used for growth calculation may have affected the identification of trees with older structural characteristics.

The validation approach also involved methodological limitations related to the reference data. The field inventory focused on trees identified as having high conservation value, whereas the classification aimed to identify trees with older structural characteristics derived from crown and height variables. The validation therefore did not represent a direct comparison between modeled candidates and inventoried trees. Consequently, the classified candidates considered false positive may still represent trees with older structural characteristics that were not registered as high conservation value trees during the inventory.

5.4 Suggestions for future research

Based on the results and the limitations of the study, there are several possible ways to further develop the work. Automating both segmentation and classification using machine learning was considered as an initial direction for this study. However, due to limitations in time, available reference data, and the need to first evaluate the potential of the selected variables, a rule-based approach was chosen instead.

Since several uncertainties were linked to crown delineation and height growth estimation, future research could build on the workflow developed in this study by investigating machine learning-based methods for both crown segmentation and tree classification. Such approaches could potentially improve both segmentation consistency and the identification of structurally complex trees.

In particular, convolutional neural networks (CNNs) and instance segmentation methods such as Mask R-CNN may be suitable for identifying complex spatial patterns in CHM data and delineating individual tree crowns in heterogeneous forest environments (Kattenborn et al., 2021; Zhao et al., 2023).

An advantage of this type of method is that it can utilize structural information from height data, such as a CHM, to identify trees based on geometric and spatial patterns rather than manually defined thresholds (Hamraz et al., 2019). This means that the rule-based classification used in this study could potentially be replaced or supplemented by a data-driven method, which could contribute to a more generalizable and scalable classification.

Future studies could also investigate the integration of multispectral or aerial image data together with LiDAR-derived CHM data. Combining structural and spectral information may improve the ability to distinguish tree species, crown conditions and other characteristics associated with older Scots pine trees. Such data integration could further increase the potential of machine learning-based approaches for identifying ecologically valuable trees in heterogeneous forest environments.

Access to reliable reference data facilitates the development and evaluation of machine learning-based methods. Through the field validation carried out, these data may support future methodological development, including the evaluation of more data-driven approaches.

5.5 Environmental and sustainability aspects

The study is relevant to the environment and sustainable development as it aims to identify Scots pine trees with characteristics associated with older forest structures which often have high ecological value and are important habitats for many species. By identifying these trees, the risk of them being felled can be reduced, which contributes to preserving biodiversity. The results can be used as a basis for forest management planning and nature conservation, where ecological aspects need to be considered at an early stage. By using spatial analyses and remote sensing data, the need for extensive field inventories can be reduced, which in turn can contribute to more efficient planning and use of resources.

The study may contribute to more sustainable forestry where production is balanced against nature conservation. The study is therefore linked to the UN Sustainable Development Goal 15: Life on Land, which aims, among other things, to conserve forests and reduce the loss of biodiversity (Globala målen, 2024). Above all, the study contributes to the ecological dimension of sustainable development by supporting the conservation of biodiversity. At the same time, improved decision-making can also have indirect effects on social and economic sustainability, for example by facilitating decisions and trade-offs between production and nature conservation, as well as by improving the efficiency of forestry planning.

In the long term, the method can contribute to a more proactive approach where older trees and other valuable structures are identified at an early stage, before forestry operations are carried out. This can reduce the risk of biologically valuable environments being negatively affected and facilitate the integration of nature conservation in forestry planning. In summary, the study shows that spatial analyses and remote sensing methods can serve as support for more sustainable forestry, where both production and conservation of biodiversity are considered.

6 Conclusions

This thesis investigated how spatial resolution in CHM data influences rule-based segmentation of individual tree crowns and how these segmentation results affect the identification of Scots pine trees with characteristics associated with older forest structures using a stratified classification approach accounting for local environmental conditions.

The validation demonstrated that the high-resolution CHM produced a higher proportion of correctly segmented crowns and less under-segmentation compared with the low-resolution CHM. The improved separation of nearby trees provided better conditions for the classification of trees with characteristics associated with older Scots pine trees.

Although the classification results showed limitations in identifying individual trees with high certainty, some inventoried trees were still captured by the classification. The method nevertheless demonstrated greater potential for identifying forest stands and parts of forest stands likely to contain structurally old trees. Rather than functioning as a precise detection tool for single trees, the model was more effective in highlighting forest stands with a higher probability of containing structurally old Scots pine trees and other ecologically valuable forest structures. This suggests that the approach may serve as a valuable support tool for prioritizing field inventories and identifying forest areas of potential conservation interest in large-scale forest planning.

The stratified approach enabled structural variables to be interpreted relative to local environmental conditions. The distribution of candidate trees across strata showed that structurally old tree characteristics occurred under different combinations of elevation, soil moisture and topographic conditions. Since the study area consisted of heterogeneous forest environments, applying the same thresholds across the entire area would not have reflected the ecological variability between sites. The stratified approach therefore reduced the risk of evaluating trees under unequal ecological conditions, where structurally smaller or slower-growing trees in less productive environments could still be identified based on their relative characteristics.

The study shows the potential of combining LiDAR-derived CHM data with rule-based analysis for ecological forest assessment at larger spatial scales. The results highlight how both spatial resolution and local environmental variation influence the ability to detect and interpret forest structural characteristics.

By integrating crown segmentation, structural analysis and ecological stratification, the study contributes to the development of methods that can support more efficient assessment of forest environments with potential ecological value.

Although a rule-based approach was used in this study, future research could build on the results by developing machine learning-based methods for both segmentation and classification. Such approaches could enable more automated and generalizable identification of trees with old-growth structural characteristics, particularly in heterogeneous forest environments. The field evaluation data collected during the study can support the development and evaluation of these methods in future work.

References

- Aarts, G., MacKenzie, M., McConnell, B., Fedak, M., & Matthiopoulos, J. (2008). Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, 31(1), 140–160. <https://doi.org/10.1111/j.2007.0906-7590.05236.x>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., Nelson, R. F., Parker, G. G., & Radeloff, V. C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research: Biogeosciences*, 114(G2). <https://doi.org/10.1029/2008JG000883>
- Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., & Tiede, D. (2014). Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 180–191. <https://doi.org/10.1016/j.isprsjprs.2013.09.014>
- Buchacher, R., & Ledermann, T. (2020). Interregional Crown Width Models for Individual Trees Growing in Pure and Mixed Stands in Austria. *Forests*, 11, 114. <https://doi.org/10.3390/f11010114>
- Buness, V., Sundqvist, M. K., Tuhin Ali, S., Annighöfer, P., Miguel Aragon, C., Lanzrein, I., Metcalfe, D. B., Nilsson, M.-C., & Gundale, M. J. (2025). Resource quantity and heterogeneity drive successional plant diversity in managed and unmanaged boreal forests. *Ecography*, 2025(7). <https://doi.org/10.1111/ecog.07676>
- Chen, Q., Baldocchi, D., Gong, P., & Kelly, M. (2006). Isolating Individual Trees in a Savanna Woodland Using Small Footprint Lidar Data. *Photogrammetric Engineering & Remote Sensing*, 72(8), 923–932. <https://doi.org/10.14358/PERS.72.8.923>

- Dalponte, M., & 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. <https://doi.org/10.1111/2041-210X.12575>
- De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., & Crombé, P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, 186, 39–49. <https://doi.org/10.1016/j.geomorph.2012.12.015>
- Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1, 100002. <https://doi.org/10.1016/j.srs.2020.100002>
- Edman, M., Eriksson, A.-M., Carlsson, F., & Rydkvist, T. (2024). Veteranising Scots pine trees by initiating tree hollowing: Inoculation with the fungal keystone species *Porodaedalia pini*. *Fungal Ecology*, 72, 101375. <https://doi.org/10.1016/j.funeco.2024.101375>
- Globala målen. (2024, December 3). *Mål 15: Ekosystem och biologisk mångfald*. <https://globalamalen.se/om-globala-malen/mal-15-ekosystem-och-biologisk-mangfald/>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson.
- Hamraz, H., Jacobs, N. B., Contreras, M. A., & Clark, C. H. (2019). Deep learning for conifer/deciduous classification of airborne LiDAR 3D point clouds representing individual trees. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158, 219–230. <https://doi.org/10.1016/j.isprsjprs.2019.10.011>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Handegard, E., Gjerde, I., Bollandsås, O. M., & Storaunet, K. O. (2021). Identifying old Norway spruce and Scots pine trees by morphological traits and site characteristics. *Scandinavian Journal of Forest Research*, 36(7–8), 550–562. <https://doi.org/10.1080/02827581.2021.1996628>

- Hasenauer, H. (1997). Dimensional relationships of open-grown trees in Austria. *Forest Ecology and Management*, 96(3), 197–206.
[https://doi.org/10.1016/S0378-1127\(97\)00057-1](https://doi.org/10.1016/S0378-1127(97)00057-1)
- Hyypä, J., Hyypä, H., Leckie, D., Gougeon, F., Yu, X., & Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29(5), 1339–1366. <https://doi.org/10.1080/01431160701736489>
- Kaartinen, H., Hyypä, J., Yu, X., Vastaranta, M., Hyypä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B., & Wu, J.-C. (2012). An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing*, 4(4), 950–974.
<https://doi.org/10.3390/rs4040950>
- Kakhani, N., Mokhtarzade, M., & Valadan Zouj, M. J. (2019). Classification of very high-resolution remote sensing images by applying a new edge-based marker-controlled watershed segmentation method. *Signal, Image and Video Processing*, 13(7), 1319–1327. <https://doi.org/10.1007/s11760-019-01477-6>
- Kanda, F., Kubo, M., & Muramoto, K. (2004). Watershed segmentation and classification of tree species using high resolution forest imagery. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 6, 3822–3825 vol.6. <https://doi.org/10.1109/IGARSS.2004.1369956>
- Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 173, 24–49.
<https://doi.org/10.1016/j.isprsjprs.2020.12.010>
- Koch, B., Heyder, U., & Weinacker, H. (2006). Detection of Individual Tree Crowns in Airborne Lidar Data. *Photogrammetric Engineering & Remote Sensing*, 72(4), 357–363. <https://doi.org/10.14358/PERS.72.4.357>
- Körner, C. (2007). The use of ‘altitude’ in ecological research. *Trends in Ecology & Evolution*, 22(11), 569–574. <https://doi.org/10.1016/j.tree.2007.09.006>

- Lantmäteriet. (2022). *Kvalitetsbeskrivning laserdata*. Lantmäteriet.
https://www.lantmateriet.se/globalassets/geodata/geodataprodukter/hojddata/kvalitetsbeskrivning_laserdata.pdf
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *BioScience*, 52(1), 19–30.
[https://doi.org/10.1641/0006-3568\(2002\)052\[0019:LRSFES\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2)
- Lindberg, E., & Holmgren, J. (2017). Individual Tree Crown Methods for 3D Data from Remote Sensing. *Current Forestry Reports*, 3(1), 19–31.
<https://doi.org/10.1007/s40725-017-0051-6>
- Lindenmayer, D., Laurance, W., & Franklin, J. (2012). Global Decline in Large Old Trees. *Science (New York, N.Y.)*, 338, 1305–1306.
<https://doi.org/10.1126/science.1231070>
- Linder, P., & Östlund, L. (1998). Structural changes in three mid-boreal Swedish forest landscapes, 1885–1996. *Biological Conservation*, 85(1–2), 9–19.
[https://doi.org/10.1016/S0006-3207\(97\)00168-7](https://doi.org/10.1016/S0006-3207(97)00168-7)
- Lucas, M., Pukrop, M., Beckschäfer, P., & Waske, B. (2024). Individual tree detection and crown delineation in the Harz National Park from 2009 to 2022 using mask R–CNN and aerial imagery. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 13, 100071.
<https://doi.org/10.1016/j.ophoto.2024.100071>
- Meyer, F., & Beucher, S. (1990). Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1, 21–46.
[https://doi.org/10.1016/1047-3203\(90\)90014-M](https://doi.org/10.1016/1047-3203(90)90014-M)
- Mielcarek, M., Stereńczak, K., & Khosravipour, A. (2018). Testing and evaluating different LiDAR-derived canopy height model generation methods for tree height estimation. *International Journal of Applied Earth Observation and Geoinformation*, 71, 132–143. <https://doi.org/10.1016/j.jag.2018.05.002>

- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2022). Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3523–3542. <https://doi.org/10.1109/TPAMI.2021.3059968>
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80(1), 88–99. [https://doi.org/10.1016/S0034-4257\(01\)00290-5](https://doi.org/10.1016/S0034-4257(01)00290-5)
- Najman, L., & Schmitt, M. (1996). Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12), 1163–1173. <https://doi.org/10.1109/34.546254>
- Nordin, P., Olofsson, E., & Hjelm, K. (2023). Within-site adaptation: Growth and mortality of Norway spruce, Scots pine and Silver birch seedlings in different planting positions across a soil moisture gradient. *Silva Fennica*, 57(3). <https://doi.org/10.14214/sf.23004>
- Oliver, C. D., & Larson, B. C. (1996). *Forest stand dynamics* (Updated ed.). Wiley.
- OpenAI. (2026, May 15). *Introducing GPT-5*. <https://openai.com/index/introducing-gpt-5/>
- Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition*, 26(9), 1277–1294. [https://doi.org/10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J)
- Pare, S., Mittal, H., Sajid, M., Bansal, J. C., Saxena, A., Jan, T., Pedrycz, W., & Prasad, M. (2021). Remote Sensing Imagery Segmentation: A Hybrid Approach. *Remote Sensing*, 13(22). <https://doi.org/10.3390/rs13224604>
- Pitkänen, J., Maltamo, M., Hyypä, J., & Yu, X. (2004). Adaptive methods for individual tree detection on airborne laser based canopy height model. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(8/W2), 187–191. https://www.researchgate.net/publication/228876216_Adaptive_methods_for_individual_tree_detection_on_airborne_laser_based_canopy_height_mode

- Popescu, S. C., & Wynne, R. H. (2004). Seeing the Trees in the forest. *Photogrammetric Engineering & Remote Sensing*, 70(5), 589–604.
<https://doi.org/10.14358/PERS.70.5.589>
- Popescu, S. C., Wynne, R. H., & Nelson, R. F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, 29(5), 564–577.
<https://doi.org/10.5589/m03-027>
- Pretzsch, H., del Río, M., Grote, R., Klemmt, H.-J., Ordóñez, C., & Oviedo, F. B. (2022). Tracing drought effects from the tree to the stand growth in temperate and Mediterranean forests: Insights and consequences for forest ecology and management. *European Journal of Forest Research*, 141(4), 727–751.
<https://doi.org/10.1007/s10342-022-01451-x>
- Przybylski, P., Mohytych, V., Rutkowski, P., Tereba, A., Tyburski, Ł., & Fyalkowska, K. (2021). Relationships between Some Biodiversity Indicators and Crown Damage of *Pinus sylvestris* L. in Natural Old Growth Pine Forests. *Sustainability*, 13(3). <https://doi.org/10.3390/su13031239>
- Puliti, S., Dash, J. P., Watt, M. S., Breidenbach, J., & Pearse, G. D. (2020). A comparison of UAV laser scanning, photogrammetry and airborne laser scanning for precision inventory of small-forest properties. *Forestry: An International Journal of Forest Research*, 93(1), 150–162.
<https://doi.org/10.1093/forestry/cpz057>
- Rautiainen, M., & Stenberg, P. (2005). Simplified tree crown model using standard forest mensuration data for Scots pine. *Agricultural and Forest Meteorology*, 128(1), 123–129. <https://doi.org/10.1016/j.agrformet.2004.09.002>
- Rawat, W., & Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, 29, 2352–2449.
https://doi.org/10.1162/NECO_a_00990
- Siitonen, J. (2001). Forest Management, Coarse Woody Debris and Saprophytic Organisms: Fennoscandian Boreal Forests as an Example. *Ecological Bulletins*, (49), 11–41. <https://www.jstor.org/stable/20113262>
- Skogsstyrelsen. (2026a, March 20). *Digitala värden i skog*.
<https://www.skogsstyrelsen.se/e-tjanster-och-kartor/karttjanster/digitala-varden-i-skog/>

- Skogsstyrelsen. (2020). *Levande träd och buskar med naturvärden* (Målbilder För God Miljöhänsyn). <https://www.skogsstyrelsen.se/globalassets/mer-om-skog/malbilder-for-god-miljohansyn/malbilder-trad-och-buskar-med-naturvarden-samt-dod-ved/levande-buskar-och-trad-med-naturvarden--exempel-2020.pdf>
- Skogsstyrelsen. (2026b). Trädhöjd Laserdata skog—Produktbeskrivning. <https://www.skogsstyrelsen.se/globalassets/sjalvservice/karttjanster/geodata/tjanster/produktbeskrivningar/raster-tradhojd-laserdata-skog---produktbeskrivning.pdf>
- Skogsstyrelsen. (2025). *Värden att bevara*. Skogsstyrelsen. <https://www.skogsstyrelsen.se/aga-skog/skydda-skog/varden-att-bevara/>
- Spies, T. A. (2004). Ecological Concepts and Diversity of Old-Growth Forests. *Journal of Forestry*, 102(3), 14–20. <https://doi.org/10.1093/jof/102.3.14>
- Szász, B., Heil, B., Kovács, G., Heilig, D., Veperdi, G., Mészáros, D., Illés, G., & Czimmer, K. (2024). Investigation of the Relationship between Topographic and Forest Stand Characteristics Using Aerial Laser Scanning and Field Survey Data. *Forests*, 15(9), 1546. <https://doi.org/10.3390/f15091546>
- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, 15. <https://doi.org/10.1186/s12880-015-0068-x>
- Tan, P.-N., Kumar, V., & Srivastava, J. (2004). Selecting the right objective measure for association analysis. *Information Systems, Knowledge Discovery and Data Mining (KDD 2002)*, 29(4), 293–313. [https://doi.org/10.1016/S0306-4379\(03\)00072-3](https://doi.org/10.1016/S0306-4379(03)00072-3)
- Vauhkonen, J., Ene, L., Gupta, S., Heinzl, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K. M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., & Maltamo, M. (2012). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry: An International Journal of Forest Research*, 85(1), 27–40. <https://doi.org/10.1093/forestry/cpr051>

- Vierling, K., Vierling, L., Gould, W., Martinuzzi, S., & Clawges, R. (2008). LiDAR: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and The Environment - FRONT ECOL ENVIRON*, 6, 90–98. <https://doi.org/10.1890/070001>
- Vosselman, G., & Maas, H.-G. (Eds.). (2010). *Airborne and terrestrial laser scanning* (Repr). Whittles Publ.
- Wallace, L. O., Lucieer, A., & Watson, C. S. (2012). Assessing the feasibility of UAV-based LiDAR for high resolution forest change detection. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39(B7), 499–504. <https://doi.org/10.5194/isprsarchives-XXXIX-B7-499-2012>
- Wehr, A., & Lohr, U. (1999). Airborne laser scanning—An introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2), 68–82. [https://doi.org/10.1016/S0924-2716\(99\)00011-8](https://doi.org/10.1016/S0924-2716(99)00011-8)
- Weiss, A. D. (2001). *Topographic Position and Landforms Analysis* [Poster presentation]. The Nature Conservancy. https://www.jennessent.com/downloads/TPI-poster-TNC_18x22.pdf
- Witkin, A. (1984). Scale-space filtering: A new approach to multi-scale description. *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 9, 150–153. <https://doi.org/10.1109/ICASSP.1984.1172729>
- Yu, X., Hyypä, J., Kukko, A., Maltamo, M., & Kaartinen, H. (2006). Change Detection Techniques for Canopy Height Growth Measurements Using Airborne Laser Scanner Data. *Photogrammetric Engineering & Remote Sensing*, 72(12), 1339–1348. <https://doi.org/10.14358/PERS.72.12.1339>
- Zhang, W., Cai, S., Liang, X., Shao, J., Hu, R., Yu, S., & Yan, G. (2020). Cloth simulation-based construction of pit-free canopy height models from airborne LiDAR data. *Forest Ecosystems*, 7(1), 1. <https://doi.org/10.1186/s40663-019-0212-0>
- Zhao, H., Morgenroth, J., Pearse, G., & Schindler, J. (2023). A Systematic Review of Individual Tree Crown Detection and Delineation with Convolutional Neural Networks (CNN). *Current Forestry Reports*, 9(3), 149–170. <https://doi.org/10.1007/s40725-023-00184-3>

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8–36.
<https://doi.org/10.1109/MGRS.2017.2762307>

Appendix A – Summary statistics for tested height thresholds

This appendix presents summary statistics for the tested height thresholds applied to both CHM resolutions (0.5 meter and 1.0 meter). For each threshold, the area and proportion of the study area above the threshold are reported together with the mean and standard deviation of canopy height.

Table A1. Summary statistics for the high-resolution CHM (0.5 m) at different tested height thresholds.

Threshold (m)	Area $\geq$ threshold (ha)	Proportion $\geq$ threshold	Mean height (m)	Standard deviation (m)
1	2117.6	0.46	8.79	4.51
2	1971.4	0.43	9.34	4.19
3	1882.6	0.41	9.66	4.01
4	1762.6	0.38	10.08	3.80
5	1623.5	0.35	10.56	3.57

Table A2. Summary statistics for the low-resolution CHM (1.0 m) at different tested height thresholds.

Threshold (m)	Area $\geq$ threshold (ha)	Proportion $\geq$ threshold	Mean height (m)	Standard deviation (m)
1	2392.7	0.51	8.90	4.44
2	2281.5	0.49	9.26	4.23
3	2182.0	0.47	9.57	4.06
4	2027.3	0.44	10.04	3.83
5	1861.1	0.40	10.54	3.60

Appendix B – Sensitivity analysis of thinning distance

This appendix presents the sensitivity analysis of the thinning distance parameter for both the high- and low-resolution CHM. The number of detected treetops and the relative decrease compared to the 0.5 m thinning distance are presented.

Table B1. Sensitivity analysis of thinning distance, high-resolution CHM.

Thinning distance (m)	Treetops	Decrease (%) vs 0.5 m
0.5	1,650,154	0.0
0.8	1,603,812	2.81
1.1	1,433,147	13.15

Table B2. Sensitivity analysis of thinning distance, low-resolution CHM.

Thinning distance (m)	Treetops	Decrease (%) vs 0.5 m
0.5	854,102	0
0.8	853,568	0.06
1.1	846,840	0.85

Appendix C – Segmented crowns per tile

This appendix presents the number of segmented tree crowns per 1000 x 1000 m tile for the high-resolution CHM. Tiles were processed with a 10 m overlap to reduce edge effects but only crowns whose centroid fell within the core tile extent are included in the counts. Tiles without segmented crowns are excluded from the table.

Table C1. Segmented crowns per tile (high-resolution CHM).

Tile	X-min	X-max	Y-min	Y-max	Crowns
1	390000	391000	6851109	6852109	807
2	390000	391000	6852109	6853109	21,099
3	390000	391000	6853109	6854109	37,173
4	391000	392000	6849109	6850109	189
5	391000	392000	6850109	6851109	13,489
6	391000	392000	6851109	6852109	33,807
7	391000	392000	6852109	6853109	43,121
8	391000	392000	6853109	6854109	16,910
9	392000	393000	6848109	6849109	5,655
10	392000	393000	6849109	6850109	35,419
11	392000	393000	6850109	6851109	40,844
12	392000	393000	6851109	6852109	35,313
13	392000	393000	6852109	6853109	28,327
14	392000	393000	6853109	6854109	488
15	393000	394000	6846109	6847109	4,300
16	393000	394000	6847109	6848109	16,610
17	393000	394000	6848109	6849109	36,217
18	393000	394000	6849109	6850109	33,093
19	393000	394000	6850109	6851109	28,317
20	393000	394000	6851109	6852109	34,775
21	393000	394000	6852109	6853109	46,696
22	393000	394000	6853109	6854109	28,544
23	393000	394000	6854109	6855109	27,786
24	393000	394000	6855109	6855927	3,254
25	394000	395000	6844109	6845109	96
26	394000	395000	6845109	6846109	26,318
27	394000	395000	6846109	6847109	43,799
28	394000	395000	6847109	6848109	49,373
29	394000	395000	6848109	6849109	30,749
30	394000	395000	6849109	6850109	24,784
31	394000	395000	6850109	6851109	29,962
32	394000	395000	6851109	6852109	50,569

33	394000	395000	6852109	6853109	33,900
34	394000	395000	6853109	6854109	48,497
35	394000	395000	6854109	6855109	39,429
36	394000	395000	6855109	6855927	10,822
37	395000	396000	6844109	6845109	11,932
38	395000	396000	6845109	6846109	46,297
39	395000	396000	6846109	6847109	804
40	395000	396000	6849109	6850109	10,470
41	395000	396000	6850109	6851109	29,037
42	395000	396000	6851109	6852109	36,235
43	395000	396000	6852109	6853109	31,155
44	395000	396000	6853109	6854109	39,223
45	395000	396000	6854109	6855109	36,329
46	395000	396000	6855109	6855927	13,071
47	396000	397000	6845109	6846109	180
48	396000	397000	6849109	6850109	837
49	396000	397000	6850109	6851109	2,124
50	396000	397000	6851109	6852109	19,950
51	396000	397000	6852109	6853109	38,657
52	396000	397000	6853109	6854109	39,815
53	396000	397000	6854109	6855109	36,753
54	396000	397000	6855109	6855927	21,974
55	397000	398000	6851109	6852109	18,532
56	397000	398000	6852109	6853109	33,445
57	397000	398000	6853109	6854109	36,453
58	397000	398000	6854109	6855109	54,399
59	397000	398000	6855109	6855927	24,905
60	398000	399000	6851109	6852109	16,053
61	398000	399000	6852109	6853109	29,748
62	398000	399000	6853109	6854109	43,748
63	398000	399000	6854109	6855109	35,771
64	398000	399000	6855109	6855927	30,366
65	399000	399333	6851109	6852109	3,669
66	399000	399333	6852109	6853109	9,018
67	399000	399333	6853109	6854109	4,071
68	399000	399333	6854109	6855109	343

Appendix D – Percentile thresholds per strata for the high-resolution dataset

This appendix presents the group-wise percentile values for tree height, crown diameter and annual growth in the high-resolution dataset. The values were used for selecting percentile-based thresholds.

Table D1. Percentile values used during threshold calibration for the high-resolution dataset.

Strata	variabel	n	p15	p20	p25	p30	p70	p75	p80
EH_dry-mesic_TH	cr_dia	96,819	2.19	2.33	2.46	2.59	3.91	4.15	4.41
EH_dry-mesic_TL	cr_dia	47,337	2.19	2.33	2.46	2.59	3.87	4.15	4.41
EH_dry-mesic_TM	cr_dia	92,486	2.26	2.39	2.52	2.65	3.83	4.03	4.3
EH_mesic-moist_TH	cr_dia	14,165	2.46	2.71	2.88	3.04	4.33	4.55	4.79
EH_mesic-moist_TL	cr_dia	59,425	2.46	2.65	2.82	2.99	4.26	4.48	4.72
EH_mesic-moist_TM	cr_dia	145,174	2.52	2.76	2.93	3.09	4.33	4.55	4.75
EH_moist-wet_TH	cr_dia	1,132	2.19	2.33	2.52	2.65	3.95	4.15	4.37
EH_moist-wet_TL	cr_dia	7,274	2.19	2.39	2.59	2.71	4.15	4.37	4.62
EH_moist-wet_TM	cr_dia	36,301	2.19	2.39	2.52	2.71	3.95	4.15	4.37
EL_dry-mesic_TH	cr_dia	186,951	2.65	2.82	3.04	3.19	4.3	4.48	4.69
EL_dry-mesic_TL	cr_dia	142,124	2.71	2.93	3.09	3.24	4.51	4.69	4.92
EL_dry-mesic_TM	cr_dia	64,405	2.71	2.88	3.04	3.19	4.41	4.58	4.79
EL_mesic-moist_TH	cr_dia	4,292	2.71	2.93	3.09	3.29	4.58	4.75	4.95
EL_mesic-moist_TL	cr_dia	63,773	2.71	2.93	3.14	3.29	4.58	4.79	5.01
EL_mesic-moist_TM	cr_dia	22,964	2.76	2.93	3.14	3.29	4.51	4.72	4.92
EL_moist-wet_TH	cr_dia	325	2.26	2.52	2.65	2.76	3.94	4.07	4.27
EL_moist-wet_TL	cr_dia	8,152	2.26	2.52	2.65	2.82	4.11	4.3	4.51
EL_moist-wet_TM	cr_dia	7,688	2.19	2.33	2.52	2.71	3.91	4.11	4.33
EM_dry-mesic_TH	cr_dia	192,153	2.52	2.71	2.82	2.99	4.15	4.33	4.55
EM_dry-mesic_TL	cr_dia	100,844	2.59	2.76	2.88	3.04	4.33	4.51	4.75
EM_dry-mesic_TM	cr_dia	89,248	2.52	2.65	2.76	2.93	4.07	4.3	4.51
EM_mesic-moist_TH	cr_dia	4,626	2.65	2.76	2.93	3.09	4.51	4.72	4.98
EM_mesic-moist_TL	cr_dia	62,225	2.52	2.71	2.88	3.04	4.37	4.58	4.82
EM_mesic-moist_TM	cr_dia	33,503	2.46	2.65	2.76	2.93	4.26	4.44	4.69
EM_moist-wet_TH	cr_dia	334	2.19	2.39	2.52	2.76	3.96	4.18	4.41
EM_moist-wet_TL	cr_dia	9,478	2.19	2.39	2.59	2.71	4.03	4.22	4.44
EM_moist-wet_TM	cr_dia	8,736	2.19	2.33	2.52	2.65	3.83	4.03	4.22
EH_dry-mesic_TH	g_ann	96,819	0.12	0.15	0.19	0.22	0.38	0.39	0.41
EH_dry-mesic_TL	g_ann	47,337	0.11	0.14	0.18	0.22	0.39	0.41	0.43
EH_dry-mesic_TM	g_ann	92,486	0.11	0.15	0.19	0.22	0.37	0.39	0.41
EH_mesic-moist_TH	g_ann	14,165	0.05	0.07	0.08	0.1	0.3	0.32	0.35
EH_mesic-moist_TL	g_ann	59,425	0.06	0.07	0.09	0.11	0.32	0.34	0.37
EH_mesic-moist_TM	g_ann	145,174	0.05	0.06	0.07	0.09	0.3	0.32	0.35
EH_moist-wet_TH	g_ann	1,132	0.02	0.03	0.03	0.04	0.09	0.1	0.12
EH_moist-wet_TL	g_ann	7,274	0.03	0.03	0.04	0.05	0.12	0.14	0.16
EH_moist-wet_TM	g_ann	36,301	0.02	0.03	0.03	0.04	0.09	0.1	0.12

EL_dry-mesic_TH	g_ann	186,951	0.13	0.15	0.17	0.18	0.35	0.38	0.4
EL_dry-mesic_TL	g_ann	142,124	0.12	0.14	0.16	0.18	0.37	0.39	0.42
EL_dry-mesic_TM	g_ann	64,405	0.13	0.15	0.17	0.19	0.37	0.39	0.42
EL_mesic-moist_TH	g_ann	4,292	0.06	0.07	0.09	0.11	0.26	0.29	0.32
EL_mesic-moist_TL	g_ann	63,773	0.08	0.1	0.11	0.13	0.32	0.35	0.39
EL_mesic-moist_TM	g_ann	22,964	0.06	0.07	0.08	0.09	0.23	0.26	0.3
EL_moist-wet_TH	g_ann	325	0.03	0.03	0.04	0.05	0.12	0.12	0.15
EL_moist-wet_TL	g_ann	8,152	0.05	0.06	0.07	0.08	0.2	0.22	0.26
EL_moist-wet_TM	g_ann	7,688	0.03	0.04	0.05	0.06	0.14	0.16	0.18
EM_dry-mesic_TH	g_ann	192,153	0.13	0.15	0.17	0.19	0.35	0.37	0.39
EM_dry-mesic_TL	g_ann	100,844	0.12	0.14	0.17	0.19	0.36	0.38	0.4
EM_dry-mesic_TM	g_ann	89,248	0.13	0.16	0.19	0.22	0.37	0.39	0.41
EM_mesic-moist_TH	g_ann	4,626	0.06	0.08	0.1	0.12	0.31	0.34	0.37
EM_mesic-moist_TL	g_ann	62,225	0.07	0.09	0.11	0.13	0.33	0.35	0.38
EM_mesic-moist_TM	g_ann	33,503	0.06	0.08	0.1	0.12	0.32	0.35	0.38
EM_moist-wet_TH	g_ann	334	0.02	0.03	0.04	0.05	0.1	0.12	0.14
EM_moist-wet_TL	g_ann	9,478	0.04	0.05	0.06	0.07	0.17	0.2	0.22
EM_moist-wet_TM	g_ann	8,736	0.03	0.04	0.04	0.05	0.14	0.16	0.19
EH_dry-mesic_TH	h2024	96,819	5.62	5.97	6.33	6.71	12.23	12.95	13.75
EH_dry-mesic_TL	h2024	47,337	5.71	6.03	6.33	6.64	11.86	12.64	13.5
EH_dry-mesic_TM	h2024	92,486	5.85	6.24	6.61	7	11.53	12.14	12.78
EH_mesic-moist_TH	h2024	14,165	6.27	6.77	7.28	7.78	12.14	12.73	13.37
EH_mesic-moist_TL	h2024	59,425	6.31	6.81	7.34	7.87	12.35	12.87	13.53
EH_mesic-moist_TM	h2024	145,174	6.51	7.09	7.64	8.19	12.26	12.77	13.38
EH_moist-wet_TH	h2024	1,132	5.11	5.4	5.65	5.93	8.73	9.1	9.55
EH_moist-wet_TL	h2024	7,274	5.5	5.87	6.26	6.65	10.3	10.84	11.52
EH_moist-wet_TM	h2024	36,301	5.19	5.53	5.85	6.17	9.08	9.56	10.11
EL_dry-mesic_TH	h2024	186,951	7.6	8.53	9.35	10.06	14.49	15.07	15.68
EL_dry-mesic_TL	h2024	142,124	7.75	8.61	9.46	10.22	15.27	15.98	16.72
EL_dry-mesic_TM	h2024	64,405	7.84	8.81	9.64	10.32	14.81	15.44	16.1
EL_mesic-moist_TH	h2024	4,292	6.64	7.33	8.02	8.6	13.32	14.03	14.86
EL_mesic-moist_TL	h2024	63,773	7.26	7.91	8.58	9.23	13.98	14.76	15.62
EL_mesic-moist_TM	h2024	22,964	6.76	7.3	7.86	8.41	12.78	13.46	14.24
EL_moist-wet_TH	h2024	325	5.08	5.29	5.57	5.8	7.94	8.34	8.83
EL_moist-wet_TL	h2024	8,152	5.77	6.24	6.68	7.13	10.76	11.36	12.01
EL_moist-wet_TM	h2024	7,688	5.27	5.64	6	6.35	9.37	9.87	10.42
EM_dry-mesic_TH	h2024	192,153	6.9	7.39	7.88	8.4	13.82	14.46	15.08
EM_dry-mesic_TL	h2024	100,844	7.04	7.52	7.99	8.48	14.15	14.92	15.67
EM_dry-mesic_TM	h2024	89,248	6.86	7.29	7.69	8.09	13.11	13.98	14.84
EM_mesic-moist_TH	h2024	4,626	6.75	7.29	7.79	8.29	13.22	14.03	14.81
EM_mesic-moist_TL	h2024	62,225	6.64	7.13	7.57	8.01	12.53	13.29	14.16
EM_mesic-moist_TM	h2024	33,503	6.37	6.82	7.23	7.62	11.71	12.52	13.47
EM_moist-wet_TH	h2024	334	4.88	5.18	5.41	5.7	8.91	9.56	10.1
EM_moist-wet_TL	h2024	9,478	5.42	5.79	6.14	6.51	9.9	10.44	11.14
EM_moist-wet_TM	h2024	8,736	5.24	5.55	5.9	6.21	8.89	9.39	9.92

Appendix E – Percentile thresholds per strata for the low-resolution dataset

This appendix presents the group-wise percentile values for tree height, crown diameter and annual height growth in the low-resolution dataset. The values were used for selecting percentile-based thresholds.

Table E1. Percentile values used during threshold calibration for the low-resolution dataset.

Strata	variabel	n	p15	p20	p25	p30	p70	p75	p80
EH_dry-mesic_TH	cr_dia	51,406	3.57	3.91	4.07	4.37	5.86	6.18	6.38
EH_dry-mesic_TL	cr_dia	24,479	3.57	3.74	4.07	4.37	5.97	6.18	6.48
EH_dry-mesic_TM	cr_dia	46,194	3.57	3.91	4.07	4.37	5.86	6.08	6.38
EH_mesic-moist_TH	cr_dia	8,885	3.74	4.07	4.37	4.65	6.18	6.38	6.68
EH_mesic-moist_TL	cr_dia	31,915	3.91	4.22	4.37	4.65	6.28	6.48	6.77
EH_mesic-moist_TM	cr_dia	80,830	3.91	4.22	4.51	4.65	6.28	6.48	6.77
EH_moist-wet_TH	cr_dia	854	2.76	2.99	3.24	3.57	5.29	5.41	5.64
EH_moist-wet_TL	cr_dia	4,484	2.99	3.39	3.57	3.91	5.75	6.08	6.38
EH_moist-wet_TM	cr_dia	24,142	2.99	3.19	3.39	3.57	5.29	5.53	5.86
EL_dry-mesic_TH	cr_dia	102,65	4.22	4.51	4.65	4.92	6.18	6.38	6.58
EL_dry-mesic_TL	cr_dia	77,194	4.37	4.65	4.79	5.05	6.38	6.58	6.86
EL_dry-mesic_TM	cr_dia	34,317	4.22	4.51	4.79	4.92	6.28	6.48	6.68
EL_mesic-moist_TH	cr_dia	2,633	4.07	4.37	4.65	4.92	6.48	6.77	7.05
EL_mesic-moist_TL	cr_dia	36,208	4.22	4.51	4.79	5.05	6.48	6.77	6.96
EL_mesic-moist_TM	cr_dia	13,313	3.91	4.22	4.51	4.79	6.38	6.58	6.86
EL_moist-wet_TH	cr_dia	219	2.76	3.19	3.57	3.57	5.17	5.47	5.64
EL_moist-wet_TL	cr_dia	4,827	3.19	3.57	3.74	4.07	5.86	6.13	6.38
EL_moist-wet_TM	cr_dia	4,745	2.99	3.19	3.39	3.57	5.41	5.64	5.97
EM_dry-mesic_TH	cr_dia	105,97	4.07	4.37	4.51	4.79	6.08	6.28	6.48
EM_dry-mesic_TL	cr_dia	55,557	4.07	4.37	4.65	4.79	6.28	6.48	6.68
EM_dry-mesic_TM	cr_dia	45,828	4.07	4.22	4.51	4.65	6.08	6.28	6.48
EM_mesic-moist_TH	cr_dia	2,828	3.91	4.22	4.51	4.79	6.38	6.58	6.86
EM_mesic-moist_TL	cr_dia	34,959	3.91	4.22	4.51	4.65	6.28	6.58	6.77
EM_mesic-moist_TM	cr_dia	18,924	3.74	4.07	4.22	4.51	6.18	6.38	6.68
EM_moist-wet_TH	cr_dia	227	2.99	3.19	3.39	3.57	5.41	5.75	6.18
EM_moist-wet_TL	cr_dia	5,831	2.99	3.39	3.57	3.91	5.64	5.97	6.28
EM_moist-wet_TM	cr_dia	5,440	2.99	3.19	3.39	3.57	5.29	5.53	5.86
EH_dry-mesic_TH	g_ann	51,406	0.09	0.12	0.16	0.19	0.38	0.39	0.41
EH_dry-mesic_TL	g_ann	24,479	0.09	0.12	0.15	0.19	0.39	0.41	0.44
EH_dry-mesic_TM	g_ann	46,194	0.09	0.12	0.16	0.19	0.36	0.39	0.41
EH_mesic-moist_TH	g_ann	8,885	0.04	0.05	0.06	0.07	0.29	0.32	0.35
EH_mesic-moist_TL	g_ann	31,915	0.04	0.06	0.08	0.1	0.31	0.34	0.36
EH_mesic-moist_TM	g_ann	80,830	0.04	0.05	0.06	0.08	0.29	0.32	0.35
EH_moist-wet_TH	g_ann	854	0.01	0.01	0.01	0.03	0.07	0.09	0.1
EH_moist-wet_TL	g_ann	4,484	0.01	0.02	0.03	0.03	0.11	0.13	0.15
EH_moist-wet_TM	g_ann	24,142	0.01	0.01	0.01	0.03	0.07	0.09	0.11
EL_dry-mesic_TH	g_ann	102,65	0.1	0.12	0.15	0.17	0.35	0.38	0.41
EL_dry-mesic_TL	g_ann	77,194	0.1	0.12	0.14	0.16	0.36	0.39	0.43
EL_dry-mesic_TM	g_ann	34,317	0.1	0.13	0.15	0.17	0.36	0.4	0.42

EL_mesic-moist_TH	g_ann	2,633	0.05	0.06	0.07	0.09	0.27	0.3	0.34
EL_mesic-moist_TL	g_ann	36,208	0.06	0.09	0.1	0.12	0.32	0.36	0.4
EL_mesic-moist_TM	g_ann	13,313	0.04	0.05	0.06	0.07	0.24	0.27	0.3
EL_moist-wet_TH	g_ann	219	0.01	0.01	0.03	0.03	0.09	0.1	0.12
EL_moist-wet_TL	g_ann	4,827	0.03	0.04	0.05	0.06	0.19	0.21	0.26
EL_moist-wet_TM	g_ann	4,745	0.01	0.03	0.03	0.04	0.12	0.14	0.16
EM_dry-mesic_TH	g_ann	105,97	0.1	0.12	0.15	0.17	0.34	0.36	0.39
EM_dry-mesic_TL	g_ann	55,557	0.1	0.12	0.15	0.17	0.36	0.38	0.41
EM_dry-mesic_TM	g_ann	45,828	0.11	0.14	0.16	0.19	0.36	0.39	0.41
EM_mesic-moist_TH	g_ann	2,828	0.05	0.07	0.09	0.11	0.3	0.34	0.36
EM_mesic-moist_TL	g_ann	34,959	0.06	0.08	0.1	0.12	0.32	0.35	0.38
EM_mesic-moist_TM	g_ann	18,924	0.05	0.06	0.08	0.1	0.31	0.34	0.38
EM_moist-wet_TH	g_ann	227	0.01	0.01	0.02	0.03	0.09	0.11	0.13
EM_moist-wet_TL	g_ann	5,831	0.03	0.03	0.04	0.05	0.16	0.19	0.22
EM_moist-wet_TM	g_ann	5,440	0.01	0.03	0.03	0.04	0.12	0.15	0.17
EH_dry-mesic_TH	h2022	51,406	5.5	5.9	6.4	6.9	13.17	13.8	14.4
EH_dry-mesic_TL	h2022	24,479	5.5	5.9	6.2	6.6	13	13.7	14.5
EH_dry-mesic_TM	h2022	46,194	5.8	6.2	6.6	7.1	12	12.7	13.4
EH_mesic-moist_TH	h2022	8,885	6.4	7	7.5	8	12.5	13.08	13.8
EH_mesic-moist_TL	h2022	31,915	6.5	7.1	7.8	8.4	12.84	13.4	14.11
EH_mesic-moist_TM	h2022	80,830	6.7	7.3	7.99	8.5	12.52	13.1	13.7
EH_moist-wet_TH	h2022	854	5.2	5.5	5.8	6.2	8.8	9.2	9.6
EH_moist-wet_TL	h2022	4,484	5.7	6.12	6.54	6.96	10.5	11	11.7
EH_moist-wet_TM	h2022	24,142	5.4	5.7	6	6.35	9.1	9.6	10.1
EL_dry-mesic_TH	h2022	102,65	8.5	9.5	10.3	11	15.12	15.7	16.2
EL_dry-mesic_TL	h2022	77,194	8.7	9.8	10.6	11.38	16.1	16.7	17.4
EL_dry-mesic_TM	h2022	34,317	8.7	9.74	10.5	11.1	15.5	16	16.6
EL_mesic-moist_TH	h2022	2,633	7	7.74	8.4	9.07	14.2	14.9	15.6
EL_mesic-moist_TL	h2022	36,208	7.8	8.6	9.4	10.1	14.9	15.6	16.34
EL_mesic-moist_TM	h2022	13,313	6.9	7.5	8.1	8.7	13.31	14	14.8
EL_moist-wet_TH	h2022	219	5.1	5.3	5.45	5.8	8.1	8.43	8.93
EL_moist-wet_TL	h2022	4,827	6	6.5	7	7.4	11.18	11.7	12.4
EL_moist-wet_TM	h2022	4,745	5.6	5.9	6.3	6.6	9.4	9.9	10.46
EM_dry-mesic_TH	h2022	105,97	6.9	7.5	8.2	8.9	14.48	15	15.5
EM_dry-mesic_TL	h2022	55,557	7.1	7.7	8.3	9	15	15.6	16.3
EM_dry-mesic_TM	h2022	45,828	6.8	7.3	7.8	8.3	14.2	14.81	15.5
EM_mesic-moist_TH	h2022	2,828	6.7	7.3	8	8.62	13.9	14.66	15.3
EM_mesic-moist_TL	h2022	34,959	6.8	7.3	7.9	8.4	13.5	14.2	14.99
EM_mesic-moist_TM	h2022	18,924	6.3	6.8	7.3	7.7	12.6	13.4	14.28
EM_moist-wet_TH	h2022	227	4.99	5.22	5.5	5.85	8.64	9.15	9.84
EM_moist-wet_TL	h2022	5,831	5.62	6	6.34	6.7	10.2	10.8	11.5
EM_moist-wet_TM	h2022	5,440	5.4	5.8	6.1	6.4	9	9.5	10.01

Appendix F – Distribution of candidate trees and strata area statistics (high-resolution)

This appendix summarizes the distribution of candidate trees across the 27 strata derived from combinations of elevation soil moisture, and TPI classes in the high-resolution dataset. For each stratum, the number of candidates, total area (ha), and candidate density (candidates per ha) are presented.

Table F1. Candidate tree distribution and area statistics for the high-resolution dataset.

Strata	Candidates	Area (ha)	Candidates per ha
EH_mesic-moist_TM	6,573	180.55	36.41
EM_dry-mesic_TH	4,268	219.14	19.48
EL_dry-mesic_TL	3,514	189.7	18.52
EL_dry-mesic_TH	3,365	228.67	14.72
EH_dry-mesic_TH	3,314	98.72	33.57
EM_dry-mesic_TL	3,031	124.36	24.37
EH_dry-mesic_TM	2,942	92.7	31.74
EH_mesic-moist_TL	2,691	71.71	37.53
EM_mesic-moist_TL	2,166	77.78	27.85
EM_dry-mesic_TM	2,153	99.89	21.55
EL_mesic-moist_TL	1,903	88.47	21.51
EH_dry-mesic_TL	1,642	48.35	33.96
EM_mesic-moist_TM	1,217	39.87	30.53
EL_dry-mesic_TM	1,181	82.07	14.39
EH_moist-wet_TM	791	37.09	21.33
EL_mesic-moist_TM	676	30.92	21.86
EH_mesic-moist_TH	592	17.47	33.89
EH_moist-wet_TL	262	8.05	32.55
EM_moist-wet_TL	231	9.99	23.13
EL_moist-wet_TL	184	8.96	20.53
EM_mesic-moist_TH	180	6.15	29.28
EL_mesic-moist_TH	122	5.9	20.68
EM_moist-wet_TM	116	8.46	13.71
EL_moist-wet_TM	92	7.67	11.99
EH_moist-wet_TH	20	1.15	17.38
EL_moist-wet_TH	2	0.32	6.18
EM_moist-wet_TH	2	0.35	5.72

Appendix G – Distribution of candidate trees and strata area statistics (low-resolution)

This appendix summarizes the distribution of candidate trees across the 27 strata derived from combinations of elevation, soil moisture, and TPI classes in the low-resolution dataset. For each stratum, the number of candidates, total area (ha), and candidate density (candidates per ha) are presented.

Table G1. Candidate tree distribution and area statistics for the low-resolution dataset.

Strata	Candidates	Area (ha)	Candidates per ha
EM_dry-mesic_TH	4,524	261.27	17.32
EH_mesic-moist_TM	4,363	208.18	20.96
EL_dry-mesic_TH	3,391	265.04	12.79
EL_dry-mesic_TL	2,868	213.04	13.46
EH_dry-mesic_TH	2,849	117.12	24.33
EM_dry-mesic_TL	2,622	144.91	18.09
EH_dry-mesic_TM	2,372	104.09	22.79
EM_dry-mesic_TM	2,004	112.75	17.77
EM_mesic-moist_TL	1,778	91.07	19.52
EH_mesic-moist_TL	1,768	82.12	21.53
EL_mesic-moist_TL	1,477	101.61	14.54
EH_dry-mesic_TL	1,368	56.4	24.25
EL_dry-mesic_TM	1,100	90.43	12.16
EM_mesic-moist_TM	911	46.36	19.65
EH_moist-wet_TM	542	44.2	12.26
EL_mesic-moist_TM	499	35.14	14.2
EH_mesic-moist_TH	474	22.29	21.26
EH_moist-wet_TL	184	9.55	19.27
EM_moist-wet_TL	173	12.12	14.27
EL_moist-wet_TL	160	10.66	15.01
EM_mesic-moist_TH	138	7.56	18.24
EL_mesic-moist_TH	102	7.41	13.77
EM_moist-wet_TM	86	10.06	8.55
EL_moist-wet_TM	75	9.03	8.31
EH_moist-wet_TH	13	1.48	8.77
EM_moist-wet_TH	2	0.44	4.6
EL_moist-wet_TH	0	0.39	0

Appendix H – Field site photographs

Tree with cavities



Tree marked with conservation value



Tree with dead branches



Site overview



Appendix I – Code availability

The scripts used in this study are available in a public GitHub repository:

Bäremo, A., & Spik, C. (2026). *LiDAR-CHM-Forest-Analysis* [Computer software].
GitHub. <https://github.com/abaremo/LiDAR-CHM-Forest-Analysis>