Check for updates

Received: 5 March 2026 | Accepted: 3 July 2026

DOI: 10.1111/1365-2664.70561

RESEARCH ARTICLE

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

Tracking and modelling native woodland expansion in the Scottish Highlands using LiDAR

Sudina Thapa[1] | Aland H. Y. Chan[1] | Philippa R. Gullett[2,3] | William R. M. Flynn[1] | Liam Wakefield[1] | Mark H. Hancock[3] | Anil Madhavapeddy[4] | David A. Coomes[1]

[1]Conservation Research Institute and Department of Plant Sciences, University of Cambridge, Cambridge, UK

[2]Cairngorms Connect, Aviemore, UK

[3]RSPB Scotland, Centre for Conservation Science, Edinburgh, UK

[4]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

Correspondence
David A. Coomes
Email: dac18@cam.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/W00495X/1

Handling Editor: Javier Lopatin

Abstract

1. Sustained culling of deer in parts of the Scottish Highlands is resulting in natural recolonisation of open landscapes by woodlands. Tracking these changes and furthering our understanding of the drivers of expansion is challenging due to the limited scale of field-collected data.
2. We developed a new pipeline to automatically segment and map small regenerating trees with airborne LiDAR data (pulse density = 22 m^{-2}) and applied it to a 569 km^2 study site over Cairngorms Connect—the largest contiguous restoration partnership in the United Kingdom. The results were validated by young trees accurately geolocated and measured in the field. We then used the map of regenerating trees to build a Hurdle Model and assessed the influence of topographical and biophysical factors on regeneration success.
3. Results show that clusters of regenerating trees are detectable once they outgrow neighbouring shrubs (up to 1.1 m in height). Detection rates reached 83% and >95% for tree clusters 1.4–1.6 m and >1.8 m in height, respectively. The algorithm also gave good estimates of the heights (R^2 = 0.94) and canopy cover (R^2 = 0.90) of young trees, whilst keeping commission errors low (0.019%).
4. Deer control has led to large-scale natural forest colonisation in Cairngorms Connect. The new pipeline delineated 2.65 million young trees (<5 m in height) in the study area. It is estimated that 5325 ha of previously non-forested land now records sapling density >100 ha^{-1}. Geospatially, proximity to seed sources, elevation, habitat type and deer density all strongly affected tree establishment and sapling density.
5. *Synthesis and applications*. Forest restoration projects should utilise the new LiDAR-based pipelines to effectively map young trees, assess whether targets are met, and plan further action. Notably, the Scottish Government is now commissioning a national airborne LiDAR scan, providing a unique opportunity to upscale our algorithms to create national woodland expansion databases.

Sudina Thapa and Aland H. Y. Chan Both authors contributed equally.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

KEYWORDS
airborne LiDAR, crown segmentation, deer management, forest restoration, *Pinus sylvestris*, Scotland, Scots pine, tree regeneration

1 | INTRODUCTION

In the decade in which the United Nations has pledged to restore 30% of the planet's land to nature, Scotland holds substantial forest restoration potential that could deliver major gains for carbon sequestration and biodiversity. Forests covered 50%–80% of the total land area of Scotland around 3000 BCE (Smout, 2003). However, centuries of land conversion for agriculture and the subsequent proliferation of hunting estates means that only 4% of native forests remain, with the landscape now extensively covered by grazed heathland and grasslands with low structural diversity (Patterson et al., 2014). A recent study showed that an estimated 3.9 Mha of native woodland could be established in the Scottish uplands (Fletcher et al., 2021). This translates to a potential carbon sequestration of 696 million metric tonnes of CO_2 in the next century, equivalent to 35%–45% of the UK carbon removal targets through woodland creation suggested by the UK Climate Change Committee (Fletcher et al., 2021). This sentiment is echoed by Scotland's Forest Strategy, which set a target to raise forest cover from 5% to 21% by 2032, including creating 3000–5000 ha of native woodlands annually (Scottish Government, 2019). The restoration of lost woodlands at scale would also greatly benefit threatened species that are currently restricted to remaining forest fragments. These include iconic and charismatic fauna, such as Capercaillie (*Tetrao urogallus*) and red squirrel (*Sciurus vulgaris*), as well as rare flora, such as the one-flowered Wintergreen (*Moneses uniflora*) and twinflower (*Linnaea borealis*). Recognising these benefits, there is currently strong public support for large-scale forest expansion, with 82% of people of Scotland being in favour (Forest Research, 2025).

High establishment costs and deer browsing pose significant barriers to woodland expansion in Scotland. Traditionally, woodland creation has been driven mainly by active restoration, which incurs high financial and environmental costs. Unlike in forestry operations where costs are recouped by timber production, tree planting for restoration requires substantial financial investment (Watts, 2025). This includes physically planting saplings and setting up guards to protect planted trees. A shortfall in funding has led to several annual woodland creation targets in Scotland being missed in recent years (Confor, 2024). Active restoration also carries some upfront environmental impacts such as soil disturbance during tree planting, which often negates part of the benefits in carbon sequestration (Friggens et al., 2020). More recently, natural regeneration has been promoted as an alternative to tree planting (Scottish Government, 2019). Pressure from deer browsing, however, poses significant barriers to regenerating saplings. At high deer numbers, browsing heavily suppresses trees <3 m in height (Palmer & Truscott, 2003; Scottish Forestry, 2024; Tanentzap et al., 2013; Watts, 2025). Studies have shown that for natural regeneration to take hold, deer numbers need to be reduced to approximately 3–5 deer/km^2 (Rao S, 2017), lower than what the Scottish Government's Deer Working group suggested for open rangeland habitats (10 deer/km^2) (Pepper et al., 2019). As a result, surveys have found that 33% of native Scottish forests had no tree regeneration at all, whilst another 54% only had one or two regenerating saplings of less palatable species (Forestry Commission, 2014).

In recent years, several large projects have been established in Scotland to restore forests through natural colonisation (e.g. natural regeneration in open habitats from seeds produced in neighbouring woodlands), which created a need for effective monitoring. Notably, the Cairngorms Connect partnership (CC) brings together four land managers in the Cairngorms National Park under a shared 200-year vision to restore forests to the natural tree line through deer control (Gullett et al., 2023). This is mainly achieved through culling as deer fencing was identified as one of the major causes of death for capercaillie and other grouse species (Baines & Summers, 1997). Projects like the CC gave rise to the need for large-scale monitoring of regenerating trees. Unlike tree planting, the density of regenerating trees across an area depends heavily on biotic (herbivores, seed dispersers and seed sources) and abiotic (elevation, wind and soil conditions) factors that vary across space and time (Broughton et al., 2021). It is therefore important to track the distribution of regenerating trees to allow managers to (1) assess whether woodland creation targets are met, (2) estimate the proportion of saplings that manage to survive past the browse line, (3) predict future restoration trajectories and (4) evaluate whether interventions, such as active tree planting or reprofiling shrub structure by fire or cattle, might be needed (Broughton et al., 2021; P. R. Gullett et al., 2023; Palmer & Truscott, 2003; Thomas et al., 2015).

Remote sensing offers an effective approach to monitoring and modelling naturally regenerating trees, but existing pilot studies are limited to small study areas. LiDAR technology, such as airborne laser scanning (ALS), terrestrial laser scanning (TLS) and drone (UAV) LiDAR, can now accurately quantify 3D structures of vegetation. In particular, ALS datasets can now be used to survey millions of trees over thousands of hectares, providing an effective alternative to field surveying (Hill et al., 2017; Luscombe et al., 2023; Míguez & Fernández, 2023). Several pilot studies have demonstrated the ability for ALS to detect young trees. For instance, Du and Pang (2024) used ALS to detect regenerating trees under forest canopies over a 0.05 km^2 plot in Hebei, China. The study reported a 121% detection rate (over-detection) and 82% success rate when compared with field measurements. Commission errors mainly arose from misclassifying shrubs as saplings. Another study in Norway used ALS to identify regenerating trees near forest-tundra boundaries across a 0.12 km^2 plot. The study reported a success rate of 46% in detecting regeneration from 0.04 to 6.3 m in height (Hauglin & Næsset, 2016).

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

There is, however, a paucity of studies that leverage ALS to map and model regenerating trees across wider landscapes of conservation interest. To build trust in ALS-based workflows, we need more validation across sites with different underlying vegetation structures and background topographies such as mountainous landscapes with rocks and patches of tall vegetation (e.g. bracken *Pteridium aquilinum*). Results should also be translatable: land managers, such as partners in CC, often have monitoring programmes for tree regeneration based on established field survey protocols; if ALS-based methods were to be used more widely, we need to understand how they compare with these existing field-based methods, especially across wider landscapes. Lastly, ALS unlocks the exciting possibility of generating wall-to-wall maps of natural colonisation, which offers the potential for further analysis of the factors influencing successful tree establishment.

In this study, we used an ALS survey to map regenerating trees across a large study area (569 km^2). New methods were used to process LiDAR point clouds to map and measure heights of regenerating trees. The results were validated by field measurements and compared with existing regeneration surveys. The wall-to-wall forest regeneration maps were then used to study how management history, topography, and other biophysical gradients affected regeneration success. By doing so, we demonstrated how ALS can be used to monitor and study forest regeneration patterns across tens of thousands of hectares. With an upcoming national LiDAR scan expected to cover the whole of Scotland (Williams, 2026), this study provides a timely proof-of-concept and benchmark for monitoring woodland expansion in the region.

2 | MATERIALS AND METHODS

2.1 | Study area

The study area consists of a 569 km^2 stretch of land within the Cairngorms National Park, Scotland (red rectangle, Figure 1). The elevation spans from under 200 m in the Spey Valley to over 1100 m on the tops of the Cairngorms mountain range. The area

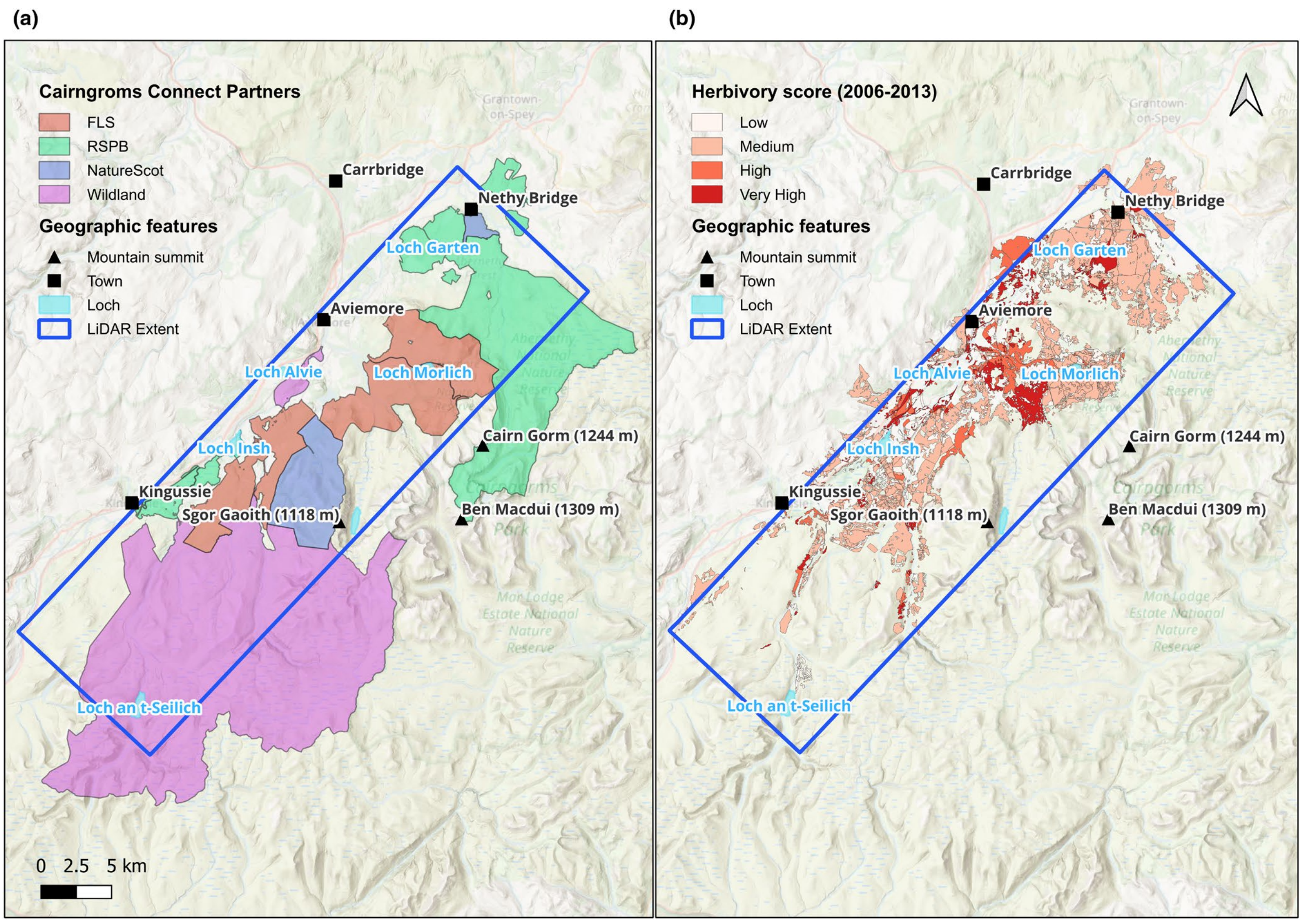


FIGURE 1 Maps of the study site. (a) land managed by the four partners of Cairngorms Connect. The estates include Glenmore and Rothiemurcus (FLS), Inshriach (FLS), Abernethy (RSPB), Insh Marshes (RSPB), Invereshie and Inshriach (NatureScot), Dell Woods (NatureScot), Glenfeshie and Gaick (Wildland Ltd) and Kinrara (Wildland Ltd); (b) the herbivory impact of land parcels covered by the LiDAR survey assessed by the Native Woodland Survey of Scotland in 2006–2013 (Forestry Commission Scotland, 2023). The deer numbers have since been reduced through culling. The blue rectangle indicates the extent of the airborne LiDAR dataset (569 km^2).

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

has a temperate climate with strong oceanic characteristics in the lowlands, which transcends to tundra climate in the mountains. The area is vegetatively diverse. In 2022, NatureScot and Space Intelligence (2023b) produced a EUNIS Level-2 Scotland Habitat and Land Cover Map based on satellite imagery and machine learning. The study estimates that the study area is 35% heathland, 32% forest, 15% bogs, 3.3% mesic grassland, 3% wet grassland, 3% alpine grassland, and 2% alpine scrub. Among the forested area are some of the best-preserved examples of Old Caledonian Forest (Scottish Forestry, 2016), which is dominated by Scots pine (*Pinus sylvestris*) but also contains some broadleaved trees such as downy birch (*Betula pubescens*) and rowan (*Sorbus aucuparia*) (Cairngorms National Park Authority, 2018; Summers et al., 1999). The area also encompasses some forestry operations, with plantations accounting for approximately 11% of the study area. Most of the study area is within Cairngorms Connect (CC), a partnership between RSPB Scotland, Wildland Ltd., Forestry and Land Scotland and NatureScot, supported by the Cairngorms National Park Authority (https://cairngormsconnect.org.uk/). Their landholdings are depicted in Figure 1a. CC partners have committed to a 200-year vision to restore biodiversity including to restore native forests, primarily through landscape-scale deer management to enable natural colonisation, supported by small-scale targeted planting (Cairngorms Connect, 2025; Gullett et al., 2023). The herbivory scores from surveys connected between 2006 and 2013 are shown in Figure 1b (Forestry Commission Scotland, 2023). Deer numbers and browsing pressure have dropped since the surveys, with current densities averaging below 4 deer/km^2 (Gullett et al., 2023).

2.2 | Aerial survey—ALS And aerial imagery

An Airborne LiDAR Survey of 569 km^2 that encompassed most of the CC partnership area was commissioned on the 17th July 2023 (Figure 1a). ALS LiDAR data were collected using a Riegl VQ780i airborne waveform laser scanner mounted on a manned Pilatus Porter PC-6 aircraft. The flight altitude was approximately 600 m above ground, and the dataset had a mean point density of 22 m^2. In addition to the LiDAR dataset, a set of four-band (RGBI) aerial images was also collected during the flight (0.1 m ground resolution).

2.3 | Environmental layers

We produced and collated environmental layers to support the detection and analysis of regenerating trees. A digital terrain model (DTM) at 1 m resolution was produced by ground classifying and rasterising the point cloud using the mcc() and tin() functions in the *lidR* package, with default parameters (Roussel et al., 2020). From the DTM, we further derived a slope raster of 1 m resolution. For the distribution of habitat types, we used the Scotland Habitat and Land Cover Map, which are derived from Sentinel 2 data and have ground resolutions of 10 m (NatureScot & Space Intelligence, 2023b). Recognising that the importance of seed sources in forest regeneration, we additionally calculated the distance from the nearest mature forest patch, also at 10 m resolution. The nearest forest patch is defined as any pixels being classed as forest in the Habitat and Land Cover Map. We included plantations into this calculation as they are also mainly comprised of Scots pine and are effective seed sources. Lastly, we obtained a mean wind speed raster for the study area from the Global Wind Atlas, which has a ground resolution of 250 m (DTU Wind Energy & World Bank Group, 2023). All raster-based operations and analyses was carried out using the *terra* package in *R 4.2.2* (Hijmans, 2025; R Core Team, 2025).

2.4 | Mapping regenerating trees—Alternative CHM pipeline

To map regenerating trees, the LiDAR point cloud was first processed into a 1-m canopy height model (CHM) using a new processing pipeline designed to retain small features (A H Y Chan et al., 2025) (Figure 2b). Traditional ground classification algorithms tend to misclassify small objects as ground due to the lack of vertical separation, which leads to poor performance on shrubs and regenerating trees (Chan et al., 2025; Gould et al., 2013; Greaves et al., 2016). Chan et al. (2025) found that this can be alleviated by (1) stretching the point cloud such that objects in the shrub layer appear as 'pseudo-forests' before ground classification and (2) normalising the ground-classified point cloud with interpolated digital terrain model (DTM-normalisation). We implemented this approach using the *lidarSHM* package (A H Y Chan et al., 2025). We mostly used the default parameters in the package, though a few changes were made to fit our use case. In particular, the original package focused on the shrub layer and removed trees. To retain trees, we changed *max_shrub_ht* parameter to 40 m. All processing was carried out using *lidR* (Roussel et al., 2020) and *lidarSHM* (A H Y Chan et al., 2025) in R 4.2.2 (R Core Team, 2025).

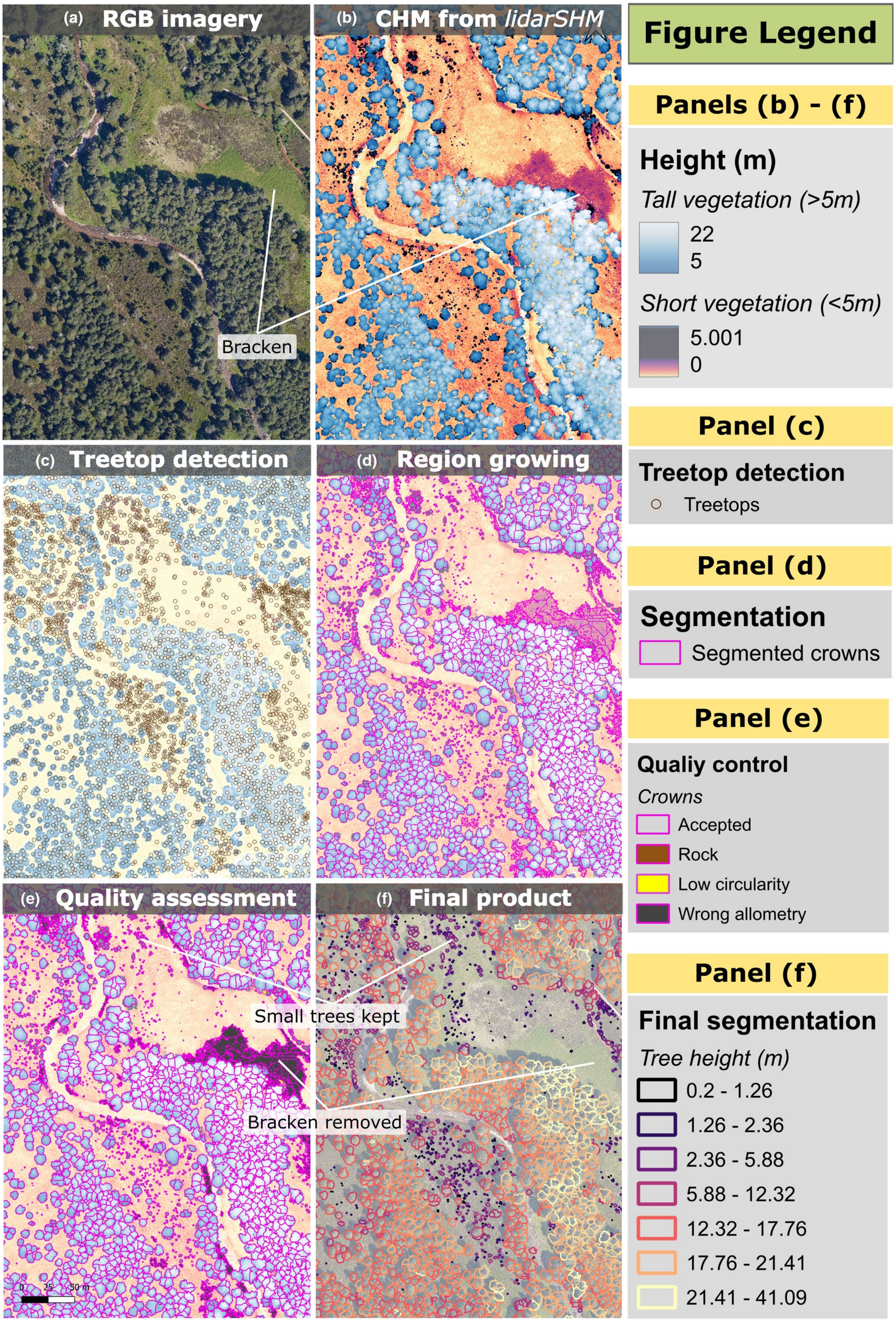
2.5 | Mapping regenerating trees—Tree top detection

After creating the CHM, we located the tops of regenerating trees using a dynamic window search approach. One issue we faced is that

FIGURE 2 Workflow for detecting tree colonisation from LiDAR data. (a) The RGB imagery of an example site. (b) the canopy height model (CHM) produced by the lidarSHM package, note that the algorithm reconstructs small features (bracken and young trees) well. (c) Detected treetops on the shrub-removed CHM. (d) The crowns created by region growing the treetops on the smoothed CHM produced by the lidarSHM package. (e) Allometry, circularity and rock filters to remove undesirable features. (f) The final delineations of both young and mature trees.

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology



13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the CHM created by the *lidarSHM* package was so sensitive to small features that it captured many bracken (*Pteridium aquilinum*, a tall fern) stands. As these bracken stands protrude above the ericaceous shrub layer, they could be mistaken as trees. To solve this issue, we specifically created a bracken-free CHM for treetop detection by subtracting the mean height of neighbourhood pixels from the CHM (window size = 3.8 × 3.8 m). From this bracken-free CHM, we located treetops using the *locate_trees* function (Figure 2c) (Roussel et al., 2020). The algorithm locates treetops by identifying the highest point across a variable search window. We defined the window size as:

$$w(x) = -3 \cdot e^{-0.15(x-1)} + 5$$

The function creates larger search windows for taller trees to prevent over-segmentation (Coomes et al., 2017). To prevent shrubs from being misclassified as trees, we removed treetops that were protruding <0.5 m above the shrub level.

2.6 | Mapping regenerating trees—Region growing and height estimation

Detected treetops were then passed on to a region growing algorithm to produce crown polygons that delineated trees (Figure 2d). To make the polygons less jagged, we first used the *focal* function in the *terra* package to smooth the canopy height model. We then applied the *its_dalponte2016* function through the *lidR* package to grow the detected treetops (Dalponte & Coomes, 2016). Pixels are only added to the crown if they are taller than (1) *th_seed* multiplied by the height of the tree and (3) *th_cr* multiplied by the mean height of pixels already in the crown. In this study, we set *th_seed* to 0.4 and *th_cr* to 0.5. This prevents low vegetation and ground pixels from being included in most cases. We estimated the heights of all delineated trees by extracting the treetop height from the CHM created by the *lidarSHM* package.

2.7 | Mapping regenerating trees—CHM Quality control

We implemented several quality control steps to ensure delineated crowns were analysis-ready (Figure 2e). Potential bracken artefacts were first screened using a crown diameter–height (D_e/H) threshold, followed by a circularity filter for borderline cases. We further removed likely rock misclassifications using an image-based classifier and excluded crowns associated with problematic topography or LiDAR noise. Finally, the analyses below were restricted to trees between 0.5 and 5 m in height to focus on regeneration. Full details of these procedures are provided in the Supplemental Information (Notes S1).

2.8 | Validation

To evaluate the performance of ALS-based forest regeneration mapping, we collected three sets of field data:

1. Geolocated and height-measured trees in the field (887 trees across 32 sites, Figure S2). This allowed us to calculate detection rate and validate LiDAR estimations of tree height. LiDAR-delineated crowns were matched with field-measured trees if they were within 1 m of each other. If a crown overlapped with multiple trees, we considered overlapping features as a 'tree cluster' when estimating detection rates. We paired the crown polygon with the tallest field-measured tree in the 'cluster' when validating height measurements. Lastly, we also used this dataset to derive a correction factor to estimate the total number of regenerating trees across the CC area from the number of LiDAR-detected tree clusters.
2. Canopy cover estimated across a network of circular plots (20 m diameter, $n = 305$) in the study area. This allowed us to validate the area of segmented crowns at a plot-level and estimate commission errors.
3. Surveys conducted by the four partner organisations of CC. To report eligibility for government grants and plan restoration action, surveys were conducted to estimate the area with regenerating densities >100 ha^{-1} for selected land parcels, with each partner adopting slightly different methodologies. These numbers were compared to LiDAR-derived estimates to evaluate whether LiDAR can replace existing monitoring at a landscape scale. In the Supplemental Information, we additionally did a more in-depth site-level comparison for the dataset from RSPB Abernethy, which includes 2318 circular plots (40 m diameter) surveyed for five classes of stem abundance (0, 1–12, 13–50, 51–140, >140) and average height (<50 cm, 50–99 cm, 100–149 cm, 150–199 cm and >200 cm).

More detailed description of the structure of the three datasets and methods used to validate the segmented crowns can be found in Figure S2 and Notes S2. The fieldwork plan was approved by the relevant land managers and consent has been granted by NatureScot to carry out the work in SSSI areas within the study site.

2.9 | Evaluating drivers of regeneration using Hurdle Models

After producing the LiDAR-based regeneration map, we built Hurdle Models to investigate the topographical or biophysical factors that facilitated or inhibited tree regeneration (Mabire-Yon, 2025; Mullahy, 1986). We decided to use a Hurdle Model as the dataset is heavily zero-inflated, with many tiles having no regenerating trees at all (Notes S4; Figure S4). Hurdle Models tackles this by having two

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

components—one for the zeros (i.e. whether there are regenerating trees) and another for counts (i.e. once the hurdle is cleared, how many trees are there).

We started by extracting data from a 50×50 m grid overlain on the study area. We chose this grid size as it captures enough variability to represent the spatial patterns in the environment without being too large to obscure local variation. We focused on trees colonising open areas, so we removed grid cells that were within or <25 m away from forests in the Scotland Habitat and Land Cover Map (NatureScot & Space Intelligence, 2023a, 2023b). For each non-forested grid cell, we extracted six variables: (1) the number of young trees (<5 m in height); (2) the habitat type; (3) the distance to forest edge; (4) the mean elevation; (5) the average wind speed; and (6) the mean slope. We focused on the six most prominent habitats, namely bog, mesic grassland, wet grassland, alpine grassland, alpine scrub and heathland. If the grid cell contained more than one class, the most abundant habitat type was adopted. Data extraction and processing were performed using the *terra* and *sf* packages in R (Hijmans, 2025).

We then built the Hurdle Model to relate the number of young trees with the other extracted environmental variables. We used a negative binomial distribution with a log-link to model the counts as it provided a better fit than Poisson (likelihood ratio test, $p < 0.0001$). An area offset term was added to the model to account for tiles that were partially masked out due to challenging topography (see Section 2.7). A set of diagnostic plots were produced by the *DHARMa* package (Hartig, 2026) to evaluate model fit (Figures S5–S7). To visualise the results, we constructed a forest plot with scaled coefficients of the Hurdle Model to compare the relative importance of different variables in affecting forest regeneration. We then used the Hurdle Model to predict how different variables affected the expected number of regenerating trees. In the main analysis, we allowed other variables to covary with the variable being scrutinised. This gives more realistic model predictions (e.g. alpine grasslands would have higher elevation) and allowed us to overlay model predictions on raw data. We do, however, recognise that the relationships visualised this way can be driven by other covariates. Therefore, we also produced an alternative set of results where we held all covariates at their mean values (Notes S5; Figures S8 and S9).

2.10 | Impacts of herbivory

We analysed the impact of herbivory on regeneration success by comparing herbivory impact data from the Native Woodland Survey of Scotland (2025) with the density of LiDAR-detected young trees. The woodland survey was carried out between 2006 and 2013. Surveyors categorised herbivore impact in each surveyed area into four categories (low, medium, high and very high) by assessing features such as bark stripping, ground disturbance, epicormic shoots and presence of palatable plant species (Forestry Commission Scotland, 2023). We then overlaid a 50×50 m grid onto the woodland survey polygons and selected non-forested grid cells that had herbivory level surveyed ($n = 4802$). For these grid cells, we analysed how herbivory level affected the density of regeneration by building a Hurdle Model and conducting post hoc pairwise tests using the *emmeans* package (Lenth & Piaskowski, 2026). We are aware that covariates can confound the relationship between herbivory and regeneration (Aland H Y Chan et al., 2024; Markoulidakis et al., 2022). To address this, we reweighted the dataset before building the Hurdle model such that four key environmental variables (distance from forest, elevation, slope and wind speed) had comparable weighted distributions amongst the different herbivory classes. See Notes S6 for more details.

3 | RESULTS

3.1 | Segmentation overview

A total of 6.17 million trees were segmented within the 569 km^2 study area. Amongst these trees, 2.65 million (43.0%) were young regenerating trees <5 m in height. There were two peaks in tree height distribution, one at lower heights representing young trees and one at around 20 m, representing the typical canopy height of mature forest in the area (Figure 3).

Deer management had led to large-scale forest expansion in the Cairngorm Connect area (Figure 4). Amongst the 50×50 m grid cells outside forests, 47.2% had at least one regenerating tree detected by the LiDAR segmentation algorithm. Excluding existing forests and correcting for under-segmented trees, it is estimated that 5235 ha had densities of young trees exceeding 100 ha^{-1}, including 1984 ha with densities over 400 ha^{-1}, meeting the government density thresholds for woodland creation by natural regeneration (Forestry Commission, 2026; Scottish Forestry, 2025; Figure 5).

3.2 | Accuracy of segmentation—Individual tree validation

Validating LiDAR-segmented crowns with trees measured in the field demonstrated the ability of the algorithm to detect small trees. As trees were geolocated in the field with a differential GPS to a sub-metre level, they produced reliable matches with crowns segmented in the LiDAR dataset (Figure 5). We noticed that some trees did not protrude above the heather layer (white points, Figure 5) and other trees formed inseparable clusters with larger trees (black points, Figure 5), so we went on to test how tall trees needed to get for the cluster to be detected (Figure 6). Previous studies in the study area have reported a maximum (99th percentile) heather height of 1.14 m in heathlands and forests (A H Y Chan et al., 2025). We found that the algorithm almost immediately detected trees once they grew beyond this heather layer. Short trees that fail to protrude above the heather stands (<1 m) were ignored (detection rate <1%), but trees only had to grow slightly taller (1.4–1.6 m) for the algorithm to detect

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

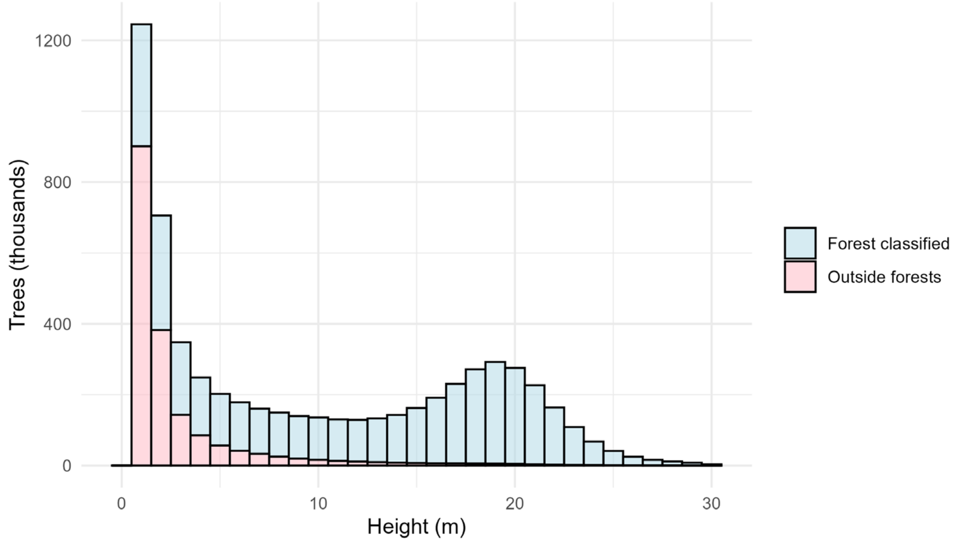


FIGURE 3 Height distribution of all delineated crowns under 30 m tall. The classification of forests was based on the Level 2 Scotland habitat and Land Cover Map 2022 (NatureScot & Space Intelligence, 2023b).

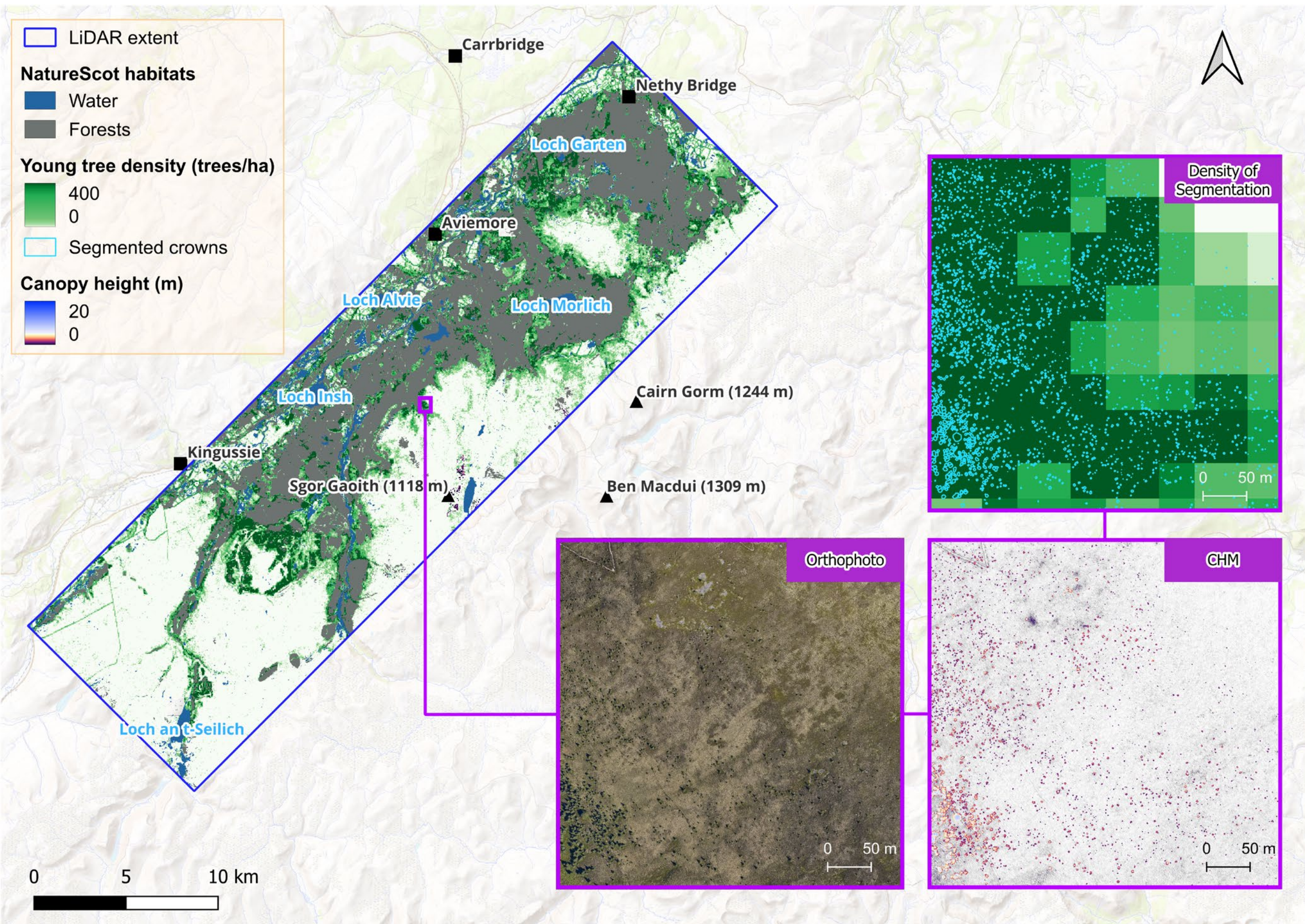


FIGURE 4 The forest is expanding at scale under management to reduce deer numbers. Main figure shows the density of young regeneration trees (<5 m) in open areas. The grey and blue regions represent existing forest and water bodies from the 2022 NatureScot Level 2 Scotland Habitat and Land Cover Map (NatureScot & Space Intelligence, 2023b). The background map is provided by OpenTopoMap. The insets demonstrate tree colonisation on a slope of Creag Dhubh, with panels showing the RGB orthophoto, canopy height model produced by the pseudo-forest pipeline and the segmented crowns overlaid on the density map.

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

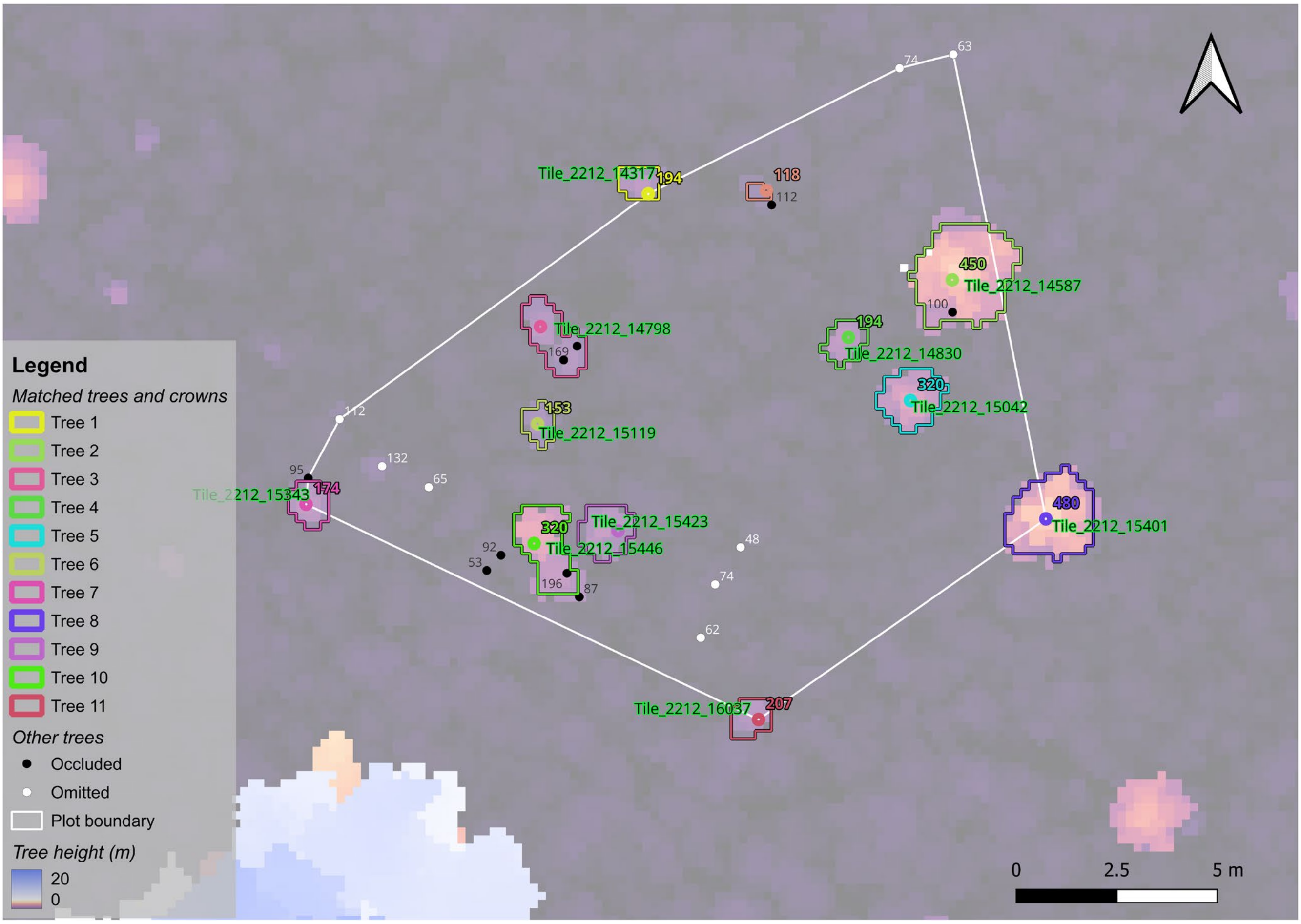


FIGURE 5 Matching of field-recorded trees (points) with LiDAR segmented crowns (polygons). The numbers are field-measured tree height in centimetres. Coloured points and polygons represent matched features. White points indicate undetected trees. Black points indicate occluded trees that were <1 m from a detected crown but were not the tallest trees in the clusters.

them with high accuracy (83%). Beyond 1.8 m, nearly all clusters were detected (>95%) (Figure 6).

Our algorithm also measured the heights of regenerating trees well (Figure 7). The LiDAR-estimated heights tightly correlated with field-measured heights of the tallest tree in the cluster (R^2 = 0.94). Note that the LiDAR data slightly underestimated sapling heights across all size classes (Figure 7). This is partly attributable to growth in the year between the LiDAR scan and the field surveys. For reference, by measuring distances between whorls of branches, we estimated that a batch of 136 *Pinus sylvestris* saplings (mean height = 2.1 m) grew by 17.6 cm in the year prior to the field surveys.

3.3 | Accuracy of segmentation—Canopy cover

The area of crowns segmented from the LiDAR dataset scaled well with canopy cover estimated in the field (Figure 8). Zero-inflated beta regression through the dataset gave a Bayesian R^2 of 0.90. Overall, the canopy covers estimated from the segmented crowns were higher than that recorded in the field, likely because of differences in the definition of canopy cover (Figure 8). Amongst the plots (10 m radius), 78 had no trees recorded in the field. Tallying the area of wrongly detected/segmented crowns in these tree-free plots, we estimated the commission error of the segmentation pipeline to be 0.019% of area processed.

3.4 | Accuracy of segmentation—Cairngorms Connect Partner surveys

Comparing the LiDAR tree density estimates with field surveys by Cairngorms Connect partners demonstrated its ability to make landscape-scale assessments of forest expansion. When estimating the area with >100 stems per hectare, the LiDAR-based estimates were of the same order as interpolations based on field surveys after applying a correction factor to convert the number of tree clusters to number of stems (Table 1). Notably, the LiDAR-based stem counts had a more stringent definition of regenerating trees (>3 years old), so the area estimates were slightly (5%–23%) lower.

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

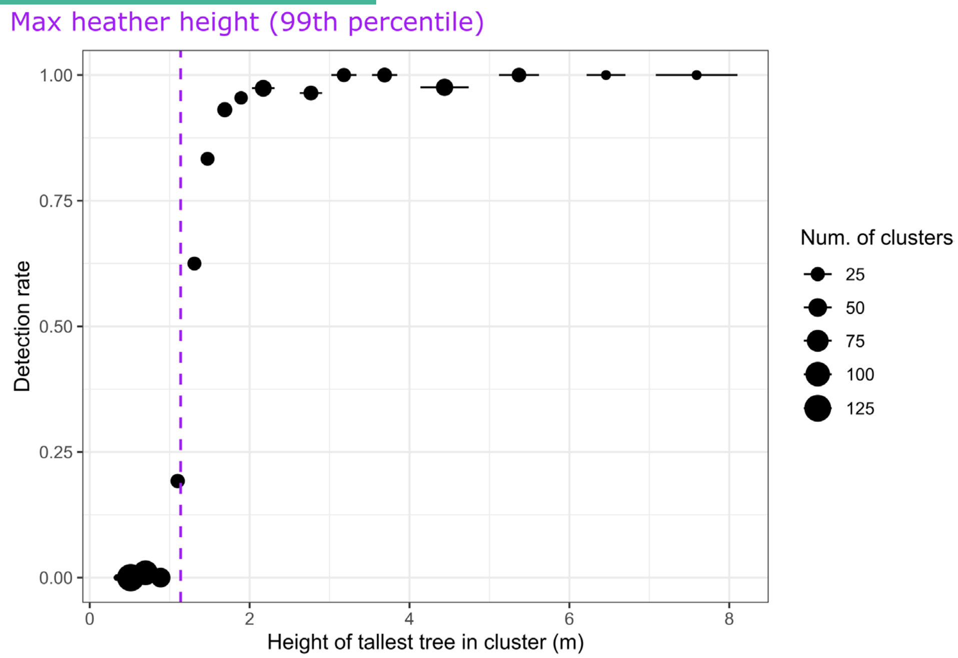


FIGURE 6 Detection rate of young trees across different height classes. The *x*-axis represents the height of a single tree or the tallest tree in a cluster if more than one tree were within a 1 m of the crown. The *y*-axis represents the proportion of clusters being detected by the algorithm. The data were binned by height class. The error bars show the standard deviation of heights in the bin. The size of the points indicating the number of clusters in the bin. The dashed purple line represents the 99th percentile of heather heights in heathlands and woodlands (Chan et al., 2025).

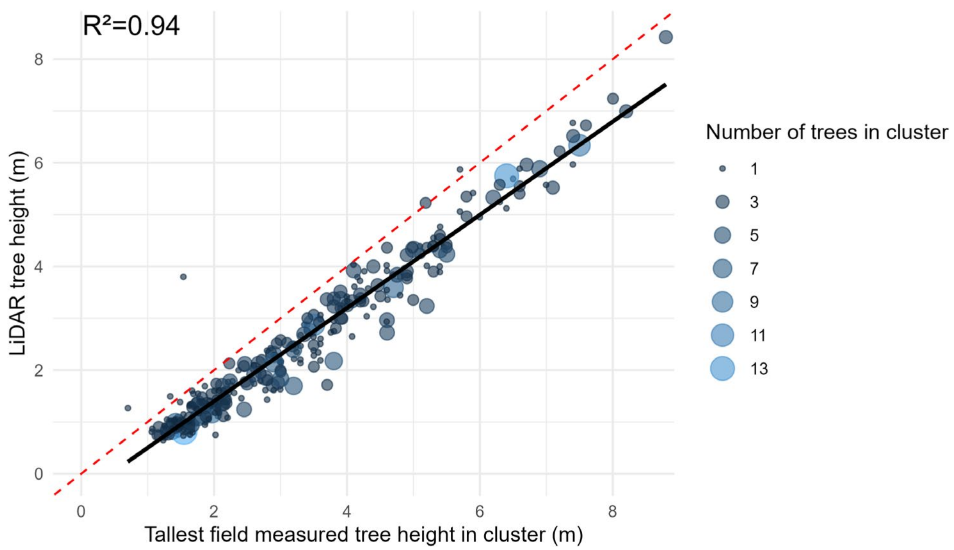


FIGURE 7 Measuring heights of saplings with LiDAR. The *x*-axis shows the height of the tallest tree in a cluster measured in the field. The *y*-axis shows the heights of saplings detected and measured in the LiDAR dataset. The red dashed line is the 1:1 line. The black solid line is the linear regression line through the data.

3.5 | Evaluating drivers of regeneration using a Hurdle Model

The Hurdle Model built on the density of young trees uncovered how various biophysical and topographical variables correlated with regeneration success. In Figure 9, we compared the effect sizes and directions by calculating the incident rate ratios (zero model) or odds ratios (count model) for all model predictors. In Figures 10 and 11, we overlaid Hurdle Model predictions onto actual data to investigate how the density of regenerating trees scaled with different environmental variables. Overall, the Hurdle Model that assumes a negative binomial distribution of tree counts provided a good fit to the data and gave realistic predictions.

Distance to forest showed the strongest effect on young tree density amongst all continuous variables. As expected, sites further away from forest edges had fewer regenerating trees, showing a clear

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

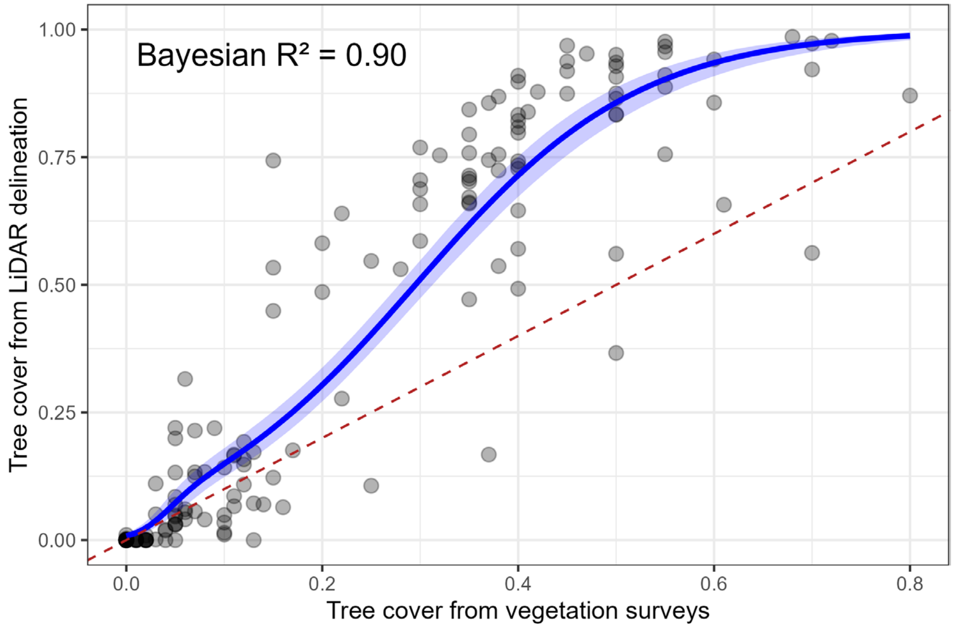


FIGURE 8 Canopy cover estimated from the LiDAR-delineated crowns against that estimated by eye in vegetation surveys. Blue line refers to estimates from a Bayesian Zero-inflated Beta Regression model with shading indicating the 95th confidence interval. Red dashed line is the 1:1 line indicating perfect agreement.

TABLE 1 Estimated area (ha) with high density of regenerating trees (>100 ha^{-1})—a comparison between field surveys carried out by partners of Cairngorms Connect and corrected numbers from the LiDAR pipeline.

Partnership	Estimated area from field surveys (ha)	Estimated area from LiDAR surveys (ha)
RSPB Abernethy	1278	1089
Wildland	619	523
Forestry and Land Scotland	515.8	492
NatureScot	277	212

Note: The comparison was made for selected sites that were field-surveyed by the Cairngorms Connect Partners, not the entire landscape. Also note that the LiDAR dataset had a more stringent definition of regenerating trees (>3 years old).

negative exponential relationship (Figures 9 and 10). Elevation was also negatively associated with both success of colonisation and tree counts (Figure 9), but showed a more nuanced relationship with its effects only manifesting >400 m as sites edge closer to the tree line (Figure 10). Slope and wind had weaker and mixed effects on tree colonisation (Figure 9).

The habitat class also strongly influenced regeneration success and tree density. Notably, as one might expect, bogs consistently had fewer regenerating trees compared to heathlands (Figures 9 and 11). The two grassland habitats were disproportionately located near lower altitude forests and were therefore associated with high densities of regenerating trees (Figure 11). However, the habitats themselves were not particularly favourable to forest regeneration as many of these regions were used as pastures for livestock (Figure 9). In contrast, alpine grasslands and shrublands had very low predicted and actual densities of regenerating trees (Figure 11), but this was mainly driven by their long distance from forests and harsh conditions associated with high elevations, not the habitat structure (Figure 9).

3.6 | Impacts of herbivory

The number of segmented tree crowns varied significantly across herbivore impact categories in the native woodland surveys (Hurdle Model, $p < 0.001$) (Figure 12). Conducting pairwise post hoc tests on the model revealed a significant difference in young tree density between the two lower herbivory classes (low and medium) and the two higher herbivory classes (high and very high) (Tukey's $p < 0.0001$). Surprisingly, the land parcels reported to have 'very high' herbivory scores (108 ha^{-1}) had significantly ($p = 0.0001$) more regenerating trees than those in the 'high' category (78 ha^{-1}). The dataset had been reweighted to create a balanced dataset for forest distance, elevation, slope and wind speed, so the patterns observed here were not due to these covariates.

4 | DISCUSSION

4.1 | Towards large-scale mapping of regenerating trees with airborne LiDAR

This paper demonstrates that airborne LiDAR can be used to map regenerating trees at scale. Our segmentation algorithm worked well for trees that protrude above the ericaceous shrub layer, detecting

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

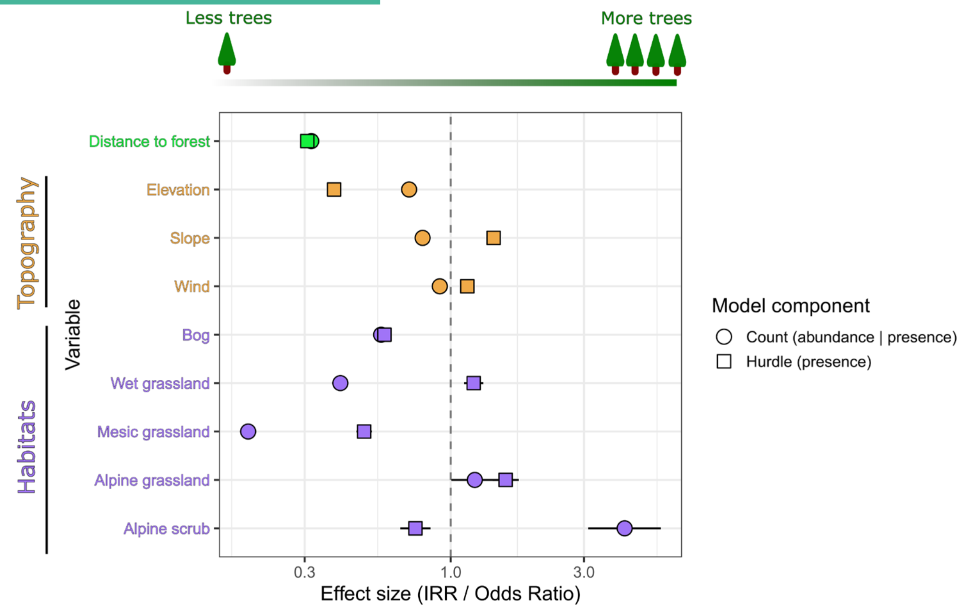


FIGURE 9 Forest plot showing the importance of variables in predicting the presence/absence of regenerating trees (zero model) and density of regenerating trees (count model). The *x*-axis shows either the incident rate ratio (IRR) for the hurdle or the odds ratio for the count model. Predictor variables were scaled for comparability. For the habitat classes, 'heathland' is set as the reference class, so the IRR or odds ratio for the other classes reflects the abundance of regeneration compared to heathland.

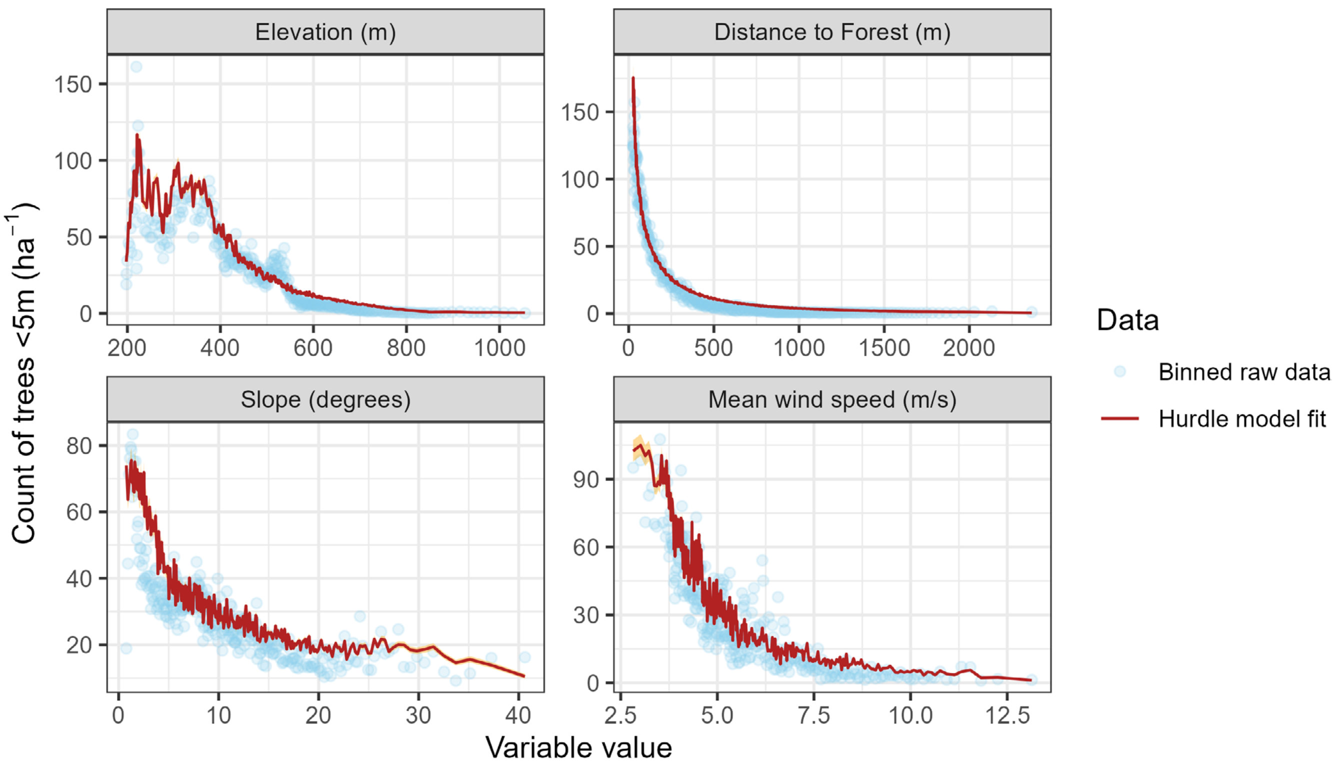


FIGURE 10 Number of saplings (<5 m in height) against four environmental variables. The red line shows predictions by the Hurdle Model. Other predictors were allowed to covary with the variable being interrogated. Orange shading represents model uncertainty (95% confidence interval, barely visible due to large dataset). Actual data is shown as pale blue points in the background. For each variable, the 50 × 50 m tiles were grouped into 400 bins according to variable value. Each point represents the mean variable value and density of regenerating trees in that bin.

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

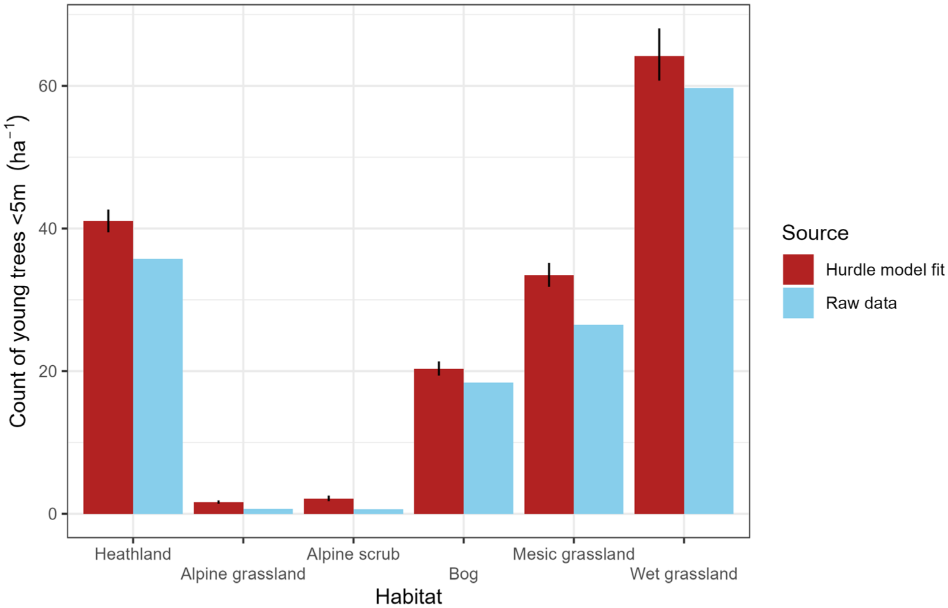


FIGURE 11 Number of saplings (<5 m in height) in different habitat classes. Habitat classes were derived from the NatureScot Scotland Habitat and Land Cover map. Red bars represent Hurdle Model predictions, with error bars showing the uncertainties (95% confidence intervals) of the model. Blue bars represent actual density of regenerating trees as detected by the LiDAR segmentation algorithm.

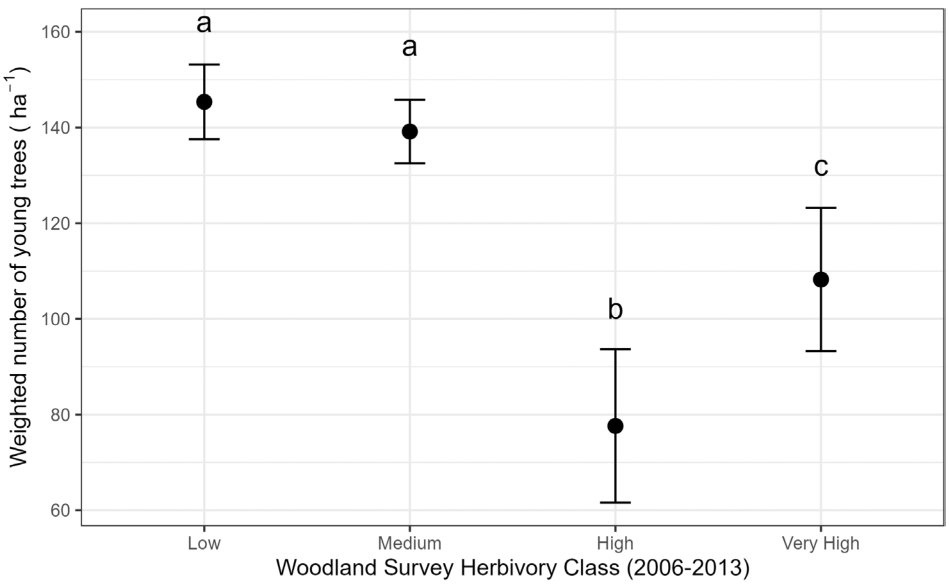


FIGURE 12 Density of young regenerating trees across sites with different herbivory classes in the Native Woodland Survey of Scotland. Error bars show standard errors. A Hurdle Model and post hoc test were built and carried out to gauge statistical significance. Points demarcated with the same alphabet were not significantly different.

83% of tree clusters 1.4–1.6 m tall and > 95% of tree clusters above 1.8 m in height. Beyond detection, the algorithm also demonstrated strong performance in estimating tree height (R^2 = 0.94) and canopy cover (R^2 = 0.90), two important metrics commonly used to gauge forest restoration success (Gullett et al., 2023; Potapov et al., 2021). At a landscape level, LiDAR-based regeneration maps were able to provide comparable estimates of woodland creation area (stem density >100 ha^{-1}) as field surveys. To our knowledge, this represents the first study to use LiDAR to map regenerating trees at this scale (>500 km^2). Our results highlight the potential for airborne LiDAR to complement or partially replace field monitoring (Table 1). Traditionally, one of the barriers of adopting an ALS-based approach is the high cost associated with collecting the data. However, as LiDAR is increasingly incorporated into a range of forestry, civil infrastructure and hydrological modelling workflows, the cost is also increasingly being spread across multiple projects and agencies,

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and some governments are now conducting regular national LiDAR scans (Chan et al., 2024; Shi et al., 2025). We believe that field validation would still be needed to account for variations between LiDAR scans, but the use of LiDAR for regular, large-scale mapping of regenerating trees is now more realistic than ever.

4.2 | Limitations of LiDAR-based young tree detection and height measurement

We faced some difficulties in detecting and segmenting very small regenerating trees due to the limitations of airborne LiDAR data. In previous work, Hauglin and Næsset (2016) used ALS data and found detection rates increasing with tree size, ranging from 15.8% to 80.4% for regeneration from 0.04 to 6.3 m tall (Hauglin & Næsset, 2016). In our study, the detection threshold sat at approximately 1.4 m, below which trees were generally omitted (Figure 6). The main challenge is the difficulty in separating out short regenerating trees from shrubs of similar heights. Chan et al. (2025) reported that heather can reach over 1.1 m in these Caledonian pinewoods and heathlands, with bracken further rising above the ericaceous shrub layer (Figure 2). Our approach partially solves this issue by setting a moving window to account for neighbouring shrub height when searching for treetops, which allows us to specifically focus on objects protruding above the shrub layer. However, treetops still get omitted when their heights were similar to or shorter than neighbouring shrubs or bracken. By detecting clusters of regenerating trees, our algorithm is sufficient for assessing areas of successful regeneration and ecological modelling on its own. However, a correction factor is still needed to convert the number of tree clusters to number of stems, accounting for occluded or undersegmented trees. Another factor that can affect detection rates is the density, spatial uniformity and footprint size of the LiDAR pulses. As regenerating trees are small, they are susceptible to laser pulses not hitting the top of trees, leading to trees being omitted. The dataset we used had a mean pulse density of 22 points/m^2, and it allowed us to detect clusters of trees at high accuracies once trees were 30–40 cm above the maximum shrub height. Further investigation is needed to ascertain how detection accuracies might change for lower point densities, different flight altitudes and different footprint sizes. Finally, it is worth noting that validating the LiDAR segmentations is far from straightforward due to geolocation errors. Here we adopted a criteria where misalignments <1 m were allowed when matching LiDAR segmentations with field-recorded trees. Statistically, relaxing this criteria can artificially inflate 'detection' rates (Figure 6) in the expense of R^2 in the height–height regression line (Figure 7). In this study, the high precision Global Navigation Satellite System (GNSS) performed sufficiently well that the spatial patterns of LiDAR- and field-recorded trees always recognisably matched up (Figure 5), but care should be taken not to relax the matching criteria too much in the case of substantial GNSS geolocation errors.

4.3 | Commission errors in detecting regenerating trees

The commission errors of the tree segmentation algorithm were low, estimated to account for approximately 0.019% of the tree-free validation area. Commission errors mainly occur due to shrubs and rocks being mistakenly classified as trees. In this study, we filtered out rocks by identifying rock coloured pixels in the RGB imagery using a random forest model trained by pixels extracted from a dozen hand drawn polygons. Whilst this is a simple approach that works for brightly coloured boulders in full sunlight, closer inspection reveals weaknesses with smaller rocks in shadows. Refining rock classification is outside the scope of this study, but it may be worth exploring more elaborate rock classification methods, such as approaches based on deep learning or LiDAR intensity, if this young tree segmentation pipeline is to be used more widely (Watts, 2025). As for shrub misclassifications, two different improvements could potentially be explored in the future. Firstly, one might consider a data fusion approach that combines spectral data to remove crowns purely consisting of common shrub species. Secondly, it may be worth assigning detection or segmentation confidence scores based on size and allometry of crowns during the quality assessment step, which allows for more stringent or liberal definitions of trees depending on application.

4.4 | Uncovering the ecology of tree colonisation in the Scottish Highlands

The mapping of regenerating trees across landscapes allowed us to carry out robust analysis on the factors associated with restoration success at unprecedented scales. As expected, shorter distances from existing forests was the strongest facilitators of regeneration (Aland H.Y. Chan & Coomes, 2024) (Figure 10). Past studies have reported typical effective seed dispersal distances for Scots pine to be up to 175 m, with some seeds travelling several kilometres by zoochory (faunal dispersal via hooves, fur or gut) or snowstorms (Picard & Baltzinger, 2012). The availability of seed sources therefore poses a significant barrier to rapid forest expansion, even in areas where browsing pressure is sufficiently low. This is particularly relevant in the Cairngorms due to the dense layer of ericaceous shrubs and mosses that develop under the oceanic climate (Summers, 2018). Only a small proportion of seeds manages to germinate and establish under the low light conditions at ground level coupled with high levels of consumption by voles or slugs (Häggström et al., 2023; Hancock et al., 2025). Our analysis also captured the negative relationship between elevation and tree colonisation (Figure 10). Previous work reported a natural tree line at 650 m in Creag Fhiaclach, which has thought to be stable in the past 1000 years (Nagy et al., 2013). In this study, we report tapering densities of regenerating trees beyond 400 m to the tree line (Figure 10). This reflects the effects of shorter growing seasons and the difficulty for trees to escape the browsing zone through

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

rapid growth (Garbarino et al., 2023; Jobbágy & Jackson, 2000). Interestingly, Simonsen (2013) studied Scots pine in Sweden and found natural colonisation to be more effective than tree planting in the harsh conditions, especially if seeds are abundant. As seed sources increase in the Cairngorms under elevated deer management, it would be worth monitoring how pine colonisation near the tree line changes over time (Garbarino et al., 2023). We also found significant differences in the density of young trees under different habitat and herbivory classes (Figures 11 and 12). The high water tables in bogs have long been known to suppress tree growth (Mchaffiea et al., 2009; Summers, 2018; Summers et al., 2008), so it is not surprising to see it also limiting tree establishment and sapling density (Figure 11). Similarly, it is no surprise that higher historical browsing pressure leads to lower sapling density (Figure 12). Intriguingly, sites with 'very high' herbivory pressure actually had significantly more young trees than sites with a 'high' herbivory score (Figure 12). This may be due to browsers opening up the vegetation and exposing the mineral soil when the deer browsing was surveyed (2006–2013), which favours tree colonisation once deer numbers are later lowered. The RSPB is currently trialling the use of cattle and artificial cutting to simulate these effects (Gullett et al., 2025; Hancock et al., 2023; Hancock et al., 2025).

4.5 | LiDAR-monitored natural regeneration as a woodland creation strategy for Scotland

Looking forward, the ability to map regenerating trees at scale represents a significant breakthrough for forest management and restoration. In Scotland, land managers are often interested in knowing the number of trees that grow past the browse line. The maximum browsing height of deer is typically between 1.1 m (roe deer) and 1.8 m (sika or red deer) (Palmer & Truscott, 2003; Scottish Forestry, 2024). Given that our pipeline can accurately locate tree clusters >1.4 m in height, we now have an effective tool to identify regions where natural colonisation translates to woodland creation (Figure 6). Compared to field surveys, a major advantage of LiDAR is its ability to create continuous maps. This is important for identifying young 'pioneers' in remote areas, which are particularly important for long-term forest expansion but are currently poorly recorded due to prohibitive survey costs (Hampe, 2011; Jacquemyn et al., 2003). Encouragingly, our regeneration maps provided clear evidence of young trees establishing extensively across Cairngorms Connect. A decade since the partnership began, large areas (5235 ha) of previously open habitats now support tree densities >100 ha^{-1}. This demonstrates that recruitment by natural regeneration can contribute significantly to Scotland's target of creating 3000–5000 ha of native woodlands annually (Scottish Government, 2019). The Scottish Government is now planning a new LiDAR scan with full coverage across the country (Williams, 2026). Our automated pipeline provides a timely, scalable tool to monitor landscape-scale changes as Scotland strives to meet its forest restoration targets.

5 | CONCLUSION

Airborne LiDAR provides a practical and scalable method for mapping regenerating trees across large restoration landscapes. Our workflow reliably detected and measured young trees that are taller than the shrub layer. Results aligned with existing field surveys and enabled robust analysis of the environmental correlates of tree establishment and colonisation. Although very small saplings remain difficult to detect, the resulting wall-to-wall maps offer a powerful tool for monitoring natural colonisation and evaluating restoration progress. As landscape-scale restoration efforts expand, automated detection of regenerating trees using approaches like this study will play an increasingly important role in supporting management.

AUTHOR CONTRIBUTIONS

Sudina Thapa, Aland H.Y. Chan and David A. Coomes conceived the ideas and designed the methodology. Aland H.Y. Chan collected the field data with support from William R.M. Flynn, Liam Wakefield and Philippa R. Gullett. Sudina Thapa analysed the data. Sudina Thapa and Aland H.Y. Chan led the writing of the manuscript. Philippa R. Gullett, Mark H. Hancock, Anil Madhavapeddy and David A. Coomes revised the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank the Natural Environment Research Council (NERC) for funding the work as part of the Changing the Environment Programme (grant number NE/W00495X/1). We are grateful to the wonderful people in all five partner organisations of Cairngorms Connect (RSPB, Wildland Ltd., NatureScot, Forestry and Land Scotland, and Cairngorms National Park Authority) for their help with this work. We are also incredibly thankful to those who helped us in the field—Issy Stuart and Catriona Forrest for measuring heights of tree saplings and Michael Durance for field support. Finally, we acknowledge support from Fugro for collecting the airborne LiDAR dataset.

CONFLICT OF INTEREST STATEMENT

We declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The up-to-date R code for generating canopy height models can be found in https://github.com/alandhyc/lidarSHM. The functions used to delineate crowns and carry out quality assessments for delineated crown polygons can be found in https://github.com/alandhyc/minitree. The original code used, identified treetops and delineated crowns are available in https://doi.org/10.17863/CAM.132432 (Chan et al., 2026).

STATEMENT OF INCLUSION

Our study brings together authors from different cultural and professional backgrounds. This includes two co-authors based in Scotland. Amongst the co-authors are those with decades of experience planning and carrying out restoration in the study area. This helped the

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

team navigate through fieldwork, be made aware of local sensitivities and prepare a manuscript that accurately represents the state of the study area with local references cited where appropriate.

ORCID

Aland H. Y. Chan https://orcid.org/0000-0001-6740-0680
Philippa R. Gullett https://orcid.org/0000-0001-7867-2315
David A. Coomes https://orcid.org/0000-0002-8261-2582

REFERENCES

Baines, D., & Summers, R. W. (1997). Assessment of bird collisions with deer fences in Scottish forests. *The Journal of Applied Ecology*, *34*(4), 941. https://doi.org/10.2307/2405284

Broughton, R. K., Bullock, J. M., George, C., Hill, R. A., Hinsley, S. A., Maziarz, M., Melin, M., Mountford, J. O., Sparks, T. H., & Pywell, R. F. (2021). Long-term woodland restoration on lowland farmland through passive rewilding. *PLoS One*, *16*(6), e0252466. https://doi.org/10.1371/journal.pone.0252466

Cairngorms Connect. (2025). *Forest restoration*. https://cairngormsconnect.org.uk/forest-restoration

Cairngorms National Park Authority. (2018). *Cairngorms National Park Forest Strategy 2018*. https://cairngorms.co.uk/wp-content/uploads/2019/03/CairngormsNationalParkForestStrategy2019Final.pdf

Chan, A. H. Y., & Coomes, D. A. (2024). Fire traps in the wet subtropics: New perspectives from Hong Kong. *Journal of Applied Ecology*, *61*(4), 884–898. https://doi.org/10.1111/1365-2664.14575

Chan, A. H. Y., Gullett, P. R., Flynn, W., Wakefield, L., Forrest, C., Stuart, I., Cowie, N. R., Hancock, M. H., & Coomes, D. (2025). *Shrub Height Model (SHM): Uncovering Shrub Structure with Airborne LiDAR*. Social Science Research Network. https://doi.org/10.2139/ssrn.5527379

Chan, A. H. Y., Jackson, T. D., Ki Law, Y., Rau, E.-P., Coomes, D. A., Fu Lam, P., & Kong, H. (2024). Forest dynamics where typhoon winds blow. *The New Phytologist*, *245*(6), 2496–2511. https://doi.org/10.1111/NPH.20350

Chan, A. H. Y., Thapa, S., Gullett, P. R., Flynn, W. R. M., Wakefield, L., Hancock, M. H., Madhavapeddy, A., & Coomes, D. A. (2026). LiDAR-detected treetops and delineated crowns in the Cairngorms Connect area. *Apollo-University of Cambridge Repository*. https://doi.org/10.17863/CAM.132432

Confor (2024). *What the Scottish Government Budget 2024/25 means for woodland creation*. https://www.confor.org.uk/media/3733213/scottish-government-budget-2024-25-confor-briefing-on-woodland-creation-jan-2024.pdf

Coomes, D. A., Dalponte, M., Jucker, T., Asner, G. P., Banin, L. F., Burslem, D. F. R. P., Lewis, S. L., Nilus, R., Phillips, O. L., Phua, M.-H., & Qie, L. (2017). Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. *Remote Sensing of Environment*, *194*, 77–88. https://doi.org/10.1016/J.RSE.2017.03.017

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*(11), 1236–1245. https://doi.org/10.1111/2041-210X.12575

DTU Wind Energy, & World Bank Group. (2023). *Global Wind Atlas*. https://globalwindatlas.info/en/

Du, L., & Pang, Y. (2024). Identifying regenerated saplings by stratifying Forest overstory using airborne {LiDAR} data. *Plant Phenomics*, *6*, 145. https://doi.org/10.34133/plantphenomics.0145

Fletcher, T. I., Scott, C. E., Hall, J., & Spracklen, D. V. (2021). The carbon sequestration potential of Scottish native woodland. *Environmental Research Communications*, *3*(4), 41003. https://doi.org/10.1088/2515-7620/abf467

Forest Research. (2025). *Public Opinion of Forestry Survey 2025: Scotland*. https://www.forestresearch.gov.uk/tools-and-resources/statistics/publications/public-opinion-of-forestry/

Forestry Commission. (2014). *Results from the Native Woodland Survey of Scotland*. https://cdn.forestresearch.gov.uk/2014/01/fcms126.pdf

Forestry Commission. (2026). *Guidance on completing the natural colonisation site assessment for your EWCO application*. https://www.gov.uk/government/publications/natural-colonisation-site-assessment-england-woodland-creation-offer/guidance-on-completing-the-natural-colonisation-site-assessment-for-your-ewco-application#fn:2

Forestry Commission Scotland. (2023). *The Woodland Herbivore Impact Assessment Method: User Guide*. https://www.forestry.gov.scot/publications/woodland-herbivore-impact-assessment-method-user-guide

Friggens, N. L., Hester, A. J., Mitchell, R. J., Parker, T. C., Subke, J. A., & Wookey, P. A. (2020). Tree planting in organic soils does not result in net carbon sequestration on decadal timescales. *Global Change Biology*, *26*(9), 5178–5188. https://doi.org/10.1111/gcb.15229

Garbarino, M., Morresi, D., Anselmetto, N., & Weisberg, P. J. (2023). Treeline remote sensing: From tracking treeline shifts to multi-dimensional monitoring of ecotonal change. *Remote Sensing in Ecology and Conservation*, *9*(6), 729–742. https://doi.org/10.1002/rse2.351

Gould, S. B., Glenn, N. F., Sankey, T. T., & McNamara, J. P. (2013). Influence of a dense, low-height shrub species on the accuracy of a lidar-derived DEM. *Photogrammetric Engineering & Remote Sensing*, *79*(5), 421–431. https://doi.org/10.14358/PERS.79.5.421

Greaves, H. E., Vierling, L. A., Eitel, J. U. H., Boelman, N. T., Magney, T. S., Prager, C. M., & Griffin, K. L. (2016). High-resolution mapping of aboveground shrub biomass in Arctic tundra using airborne lidar and imagery. *Remote Sensing of Environment*, *184*, 361–373. https://doi.org/10.1016/J.RSE.2016.07.026

Gullett, P., Hancock, M., Mason, L., & Weatherall, A. (2025). United Kingdom/Scotland: Assisted regeneration to restore lost forests. In K. Lapin, J. Oettel, M. Braun, & H. Konrad (Eds.), *Ecological connectivity of forest ecosystems*. Springer. https://doi.org/10.1007/978-3-031-82206-3_37

Gullett, P. R., Leslie, C., Mason, R., Ratcliffe, P., Sargent, I., Beck, A., Cameron, T., Cowie, N. R., Hetherington, D., MacDonell, T., Moat, T., Moore, P., Teuten, E., & Hancock, M. H. (2023). Woodland expansion in the presence of deer: 30 years of evidence from the Cairngorms Connect landscape restoration partnership. *Journal of Applied Ecology*, *60*(11), 2298–2308. https://doi.org/10.1111/1365-2664.14501;PAGEGROUP:STRING:PUBLICATION

Häggström, B., Gundale, M. J., & Nordin, A. (2023). Environmental controls on seedling establishment in a boreal forest: Implications for Scots pine regeneration in continuous cover forestry. *European Journal of Forest Research*, *143*(1), 95–106. https://doi.org/10.1007/S10342-023-01609-1

Hampe, A. (2011). Plants on the move: The role of seed dispersal and initial population establishment for climate-driven range expansions. *Acta Oecologica*, *37*(6), 666–673. https://doi.org/10.1016/j.actao.2011.05.001

Hancock, M. H., Gullett, P. R., Clift, M., Golder, P., Marshall, G., Rao, S. J., Willi, J., & Cowie, N. R. (2025). Sapling establishment by Scots pine (*Pinus sylvestris*) following experimental prescribed burning and deer exclusion on heathland beside native forest. *European Journal of Forest Research*, *144*, 1323–1336. https://doi.org/10.1007/s10342-025-01809-x

Hancock, M. H., Gullett, P. R., Golder, P., Marshall, G., & Cowie, N. R. (2023). Emulating natural disturbances with trial management in Scotland: Effects of burning, mowing and cattle on habitat measures important for capercaillie *Tetrao urogallus*. *European Journal of Forest Research*, *142*(3), 593–609. https://doi.org/10.1007/s10342-023-01544-1

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hartig, F. (2026). *DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models*. https://github.com/florianhartig/dharma

Hauglin, M., & Næsset, E. (2016). Detection and segmentation of small trees in the forest-tundra ecotone using airborne laser scanning. *Remote Sensing*, *8*(5), 407. https://doi.org/10.3390/rs8050407

Hijmans, R. J. (2025). *terra: Spatial data analysis*. https://rspatial.org/

Hill, S., Latifi, H., Heurich, M., & Müller, J. (2017). Individual-tree- and stand-based development following natural disturbance in a heterogeneously structured forest: A LiDAR-based approach. *Ecological Informatics*, *38*, 12–25. https://doi.org/10.1016/j.ecoinf.2016.12.004

Jacquemyn, H., Butaye, J., & Hermy, M. (2003). Impacts of restored patch density and distance from natural forests on colonization success. *Restoration Ecology*, *11*(4), 417–423. https://doi.org/10.1046/j.1526-100X.2003.rec0237.x

Jobbágy, E. G., & Jackson, R. B. (2000). Global controls of forest line elevation in the northern and southern hemispheres. *Journal of Biogeography*, *27*(3), 715–730. https://doi.org/10.1046/j.1365-2699.2000.00162.x

Lenth, R. V., & Piaskowski, J. (2026). *emmeans: Estimated marginal means, aka least-squares means*. https://rvlenth.github.io/emmeans/

Luscombe, D. J., Gatis, N., Anderson, K., Carless, D., & Brazier, R. E. (2023). Rapid, repeatable landscape-scale mapping of tree, hedgerow, and woodland habitats (THaW), using airborne LiDAR and spaceborne SAR data. *Ecology and Evolution*, *13*(5), e10103. https://doi.org/10.1002/ece3.10103

Mabire-Yon, R. (2025). Hurdle models in psychology—A practical guide for inflated data. *International Journal of Psychology*, *60*(3), e70042. https://doi.org/10.1002/ijop.70042

Markoulidakis, A., Taiyari, K., Holmans, P., Pallmann, P., Busse, M., Godley, M. D., & Griffin, B. A. (2022). A tutorial comparing different covariate balancing methods with an application evaluating the causal effects of substance use treatment programs for adolescents. *Health Services and Outcomes Research Methodology*, *23*, 1–34. https://doi.org/10.1007/S10742-022-00280-0/FIGURES/3

Mchaffiea, H. S., Leggb, C. J., Proctor, M. C. F., & Amphlettd, A. (2009). Vegetation and ombrotrophy on the mires of abernethy forest, Scotland. *Plant Ecology and Diversity*, *2*(1), 95–103. https://doi.org/10.1080/17550870902722317

Míguez, C., & Fernández, C. (2023). Evaluating the combined use of the NDVI and high-density lidar data to assess the natural regeneration of *P. pinaster* after a high-severity fire in NW Spain. *Remote Sensing*, *15*(6), 1634. https://doi.org/10.3390/rs15061634

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, *33*(3), 341–365. https://doi.org/10.1016/0304-4076(86)90002-3

Nagy, J., Nagy, L., Legg, C. J., & Grace, J. (2013). The stability of the *Pinus sylvestris* treeline in the Cairngorms, Scotland over the last millennium. *Plant Ecology and Diversity*, *6*(1), 7–19. https://doi.org/10.1080/17550874.2013.770933

Native Woodland Survey of Scotland. (2025).

NatureScot, & Space Intelligence. (2023a). *Scotland Habitat and Land Cover Map-2022: EUNIS Level 1 Classification*. https://gis-downloads.nature.scot/HLCM-2022-L1_SCOTLAND_TIFF_27700.zip

NatureScot, & Space Intelligence. (2023b). *Scotland Habitat and Land Cover Map - 2022: EUNIS Level 2 Classification*. https://gis-downloads.nature.scot/HLCM-2022-L2_SCOTLAND_TIFF_27700.zip

Palmer, S. C. F., & Truscott, A. M. (2003). Browsing by deer on naturally regenerating Scots pine (*Pinus sylvestris* L.) and its effects on sapling growth. *Forest Ecology and Management*, *182*(1–3), 31–47. https://doi.org/10.1016/S0378-1127(03)00026-4

Patterson, G., Nelson, D., Robertson, P., & Tullis, J. (2014). *Scotland's native woodlands: Results from the native woodland survey of Scotland*. Forestry Commission Scotland.

Pepper, S., Barbour, A., & Glass, J. (2019). *The Management of Wild Deer in Scotland*. The Scottish Government.

Picard, M., & Baltzinger, C. (2012). Hitch-hiking in the wild: Should seeds rely on ungulates? *Plant Ecology and Evolution*, *145*(1), 24–30. https://doi.org/10.5091/plecevo.2012.689

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C. E., Armston, J., Dubayah, R., Blair, J. B., & Hofton, M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, *253*(4), 112165. https://doi.org/10.1016/j.rse.2020.112165

R Core Team. (2025). *R: A language and environment for statistical computing*. https://www.r-project.org/

Rao S, J. (2017). Effect of reducing red deer Cervus elaphus density on browsing impact and growth of Scots pine *Pinus sylvestris* seedlings in semi-natural woodland in the Cairngorms, UK. *Conservation Evidence*, *14*, 22.

Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J. F., de Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of airborne laser scanning (ALS) data. *Remote Sensing of Environment*, *251*, 112061. https://doi.org/10.1016/J.RSE.2020.112061

Scottish Forestry. (2016). *Caledonian pinewood inventory*. https://spatialdata.gov.scot/geonetwork/srv/eng/catalog.search#/metadata/FC0E9D2A-798B-4B6A-8251-9750DCCCC382

Scottish Forestry. (2024). *Guidance on distinguishing between browsing by different mammal species*. https://www.forestry.gov.scot/publications/guidance-distinguishing-between-browsing-different-mammal-species

Scottish Forestry. (2025). Natural Regeneration. https://www.forestry.gov.scot/natural-regeneration

Scottish Government. (2019). *Scotland's Forestry Strategy 2019–2029*. https://www.gov.scot/publications/scotlands-forestry-strategy-20192029/documents/

Shi, Y., Wang, J., & Kissling, W. D. (2025). Multi-temporal high-resolution data products of ecosystem structure derived from country-wide airborne laser scanning surveys of The Netherlands. *Earth System Science Data*, *17*(7), 3641–3677. https://doi.org/10.5194/essd-17-3641-2025

Simonsen, R. (2013). Optimal regeneration method—planting vs. natural regeneration of Scots pine in northern Sweden. *Silva Fennica*, *47*(2), 23. https://doi.org/10.14214/sf.928

Smout, T. C. (2003). *People and woods in Scotland: A history*. Edinburgh University Press.

Summers, R. W., Mavor, R. A., MacLennan, A. M., & Rebecca, G. W. (1999). The structure of ancient native pinewoods and other woodlands in the highlands of Scotland. *Forest Ecology and Management*, *119*(1--3), 231–245. https://doi.org/10.1016/S0378-1127(98)00515-5

Summers, R. W. (2018). *Abernethy forest: The history and ecology of an old Scottish pinewood*. Royal Society for the Protection of Birds.

Summers, R. W., Wilkinson, N. I., & Wilson, E. R. (2008). Age structure and history of stand types of *Pinus sylvestris* in Abernethy Forest, Scotland. *Scandinavian Journal of Forest Research*, *23*(1), 28–37. https://doi.org/10.1080/02827580701646513

Tanentzap, A. J., Zou, J., & Coomes, D. A. (2013). Getting the biggest birch for the bang: Restoring and expanding upland birchwoods in the Scottish highlands by managing red deer. *Ecology and Evolution*, *3*(7), 1890–1901. https://doi.org/10.1002/ece3.548

Thomas, H. J. D., Paterson, J. S., Metzger, M. J., & Sing, L. (2015). Towards a research agenda for woodland expansion in Scotland. *Forest Ecology and Management*, *349*(4), 149–161. https://doi.org/10.1016/j.foreco.2015.04.003

Watts, S. H. (2025). Montane willow scrub restoration in Scotland: Reviewing 30 years of progress to reestablish the altitudinal treeline. *Restoration Ecology*, *33*(1), e14290. https://doi.org/10.1111/rec.14290

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Williams, M. (2026). *Scotland's LiDAR revolution: First data release to reveal Scotland's landscape in unprecedented detail*. Scottish Government. https://blogs.gov.scot/digital/2026/01/29/scotlands-lidar-revolution-first-data-release-to-reveal-scotlands-landscape-in-unprecedented-detail/

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1: The effective diameter of segmented tree plotted against the height of the tree. Effective diameters are calculated by converting the areas of polygons to diameters assuming the crown is circular. Note the points that are far away from the general trend line, which corresponds to bracken stands being misclassified as 'crowns'. The blue line represents the threshold we used to remove bracken.

Table S1. Confusion matrix of the rock classifier. A random forest model was built on manually delineated pixels on the RGBI orthophoto of the Cairngorms Connect area.

Figure S2. Distribution of the validation data. Only survey data within the LiDAR extent were used. The basemap is provided by OpenTopoMap.

Figure S3. (a) Boxplot and Tukey HSD post-hoc test showing the segmented regeneration compared to the count data collected by the RSPB field surveys. The categories correspond to the following counts: 0, 1–12, 13–50, 51–140 and >140. (b) Boxplot showing the average heights of the segmented regeneration (according to the LiDAR data) compared to the height data collected by the RSPB. The categories correspond to the following ranges: 50–99 cm, 100–149 cm, 150–199 cm and >200 cm. The class <50 cm was not shown as it falls below the detection threshold of LiDAR data. Isolated outliers have been removed for visualisation. Letters show statistical significance, with groups sharing the same letter not being significantly different from each other.

Figure S4. Zero-inflation shown by the counts of 50 × 50 m grid cells having different numbers of detected young trees (<5 m height). The x-axis is truncated at 15 for clarity.

Figure S5. The main diagnostic plot for the Hurdle Model produced by the DHARMa package (Hartig, 2026). The left panel shows the QQ plot. The right panel shows the DHARMa residuals plotted against the predictions, with the shading indicating the distribution of residuals and the dashed red line showing the 50th percentile.

Figure S6. The zero-inflation diagnostic plot produced by the DHARMa package (Hartig, 2026) for the regeneration count Hurdle Model.

Figure S7. The diagnostic plot for nonparametric dispersion produced by the DHARMa package (Hartig, 2026) for the regeneration count Hurdle Model.

Figure S8. The predicted number of detected young trees <5 m in height (ha^{-1}) against four environmental variables. Predictions were made by a Hurdle Model and conditioned upon holding other covariates constant at their respective mean values. The shading shows the confidence intervals (97.5th percentile) of model predictions.

Figure S9. The predicted number of detected young trees <5 m in height (ha^{-1}) across six different habitat types. Predictions were made by a Hurdle Model and conditioned upon holding other covariates constant at their respective mean values. The error bars show the confidence intervals (97.5th percentile) of model predictions.

How to cite this article: Thapa, S., Chan, A. H. Y., Gullett, P. R., Flynn, W. R. M., Wakefield, L., Hancock, M. H., Madhavapeddy, A., & Coomes, D. A. (2026). Tracking and modelling native woodland expansion in the Scottish Highlands using LiDAR. *Journal of Applied Ecology*, *63*, e70561. https://doi.org/10.1111/1365-2664.70561

13652664, 2026, 9, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70561 by Anil Madhavapeddy Two - Cambridge University Library , Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License