remote sensing

MDPI

Article

Towards Forest Condition Assessment: Evaluating Small-Footprint Full-Waveform Airborne Laser Scanning Data for Deriving Forest Structural and Compositional Metrics

Matthew J. Sumnall [1], Ross A. Hill [2,*] and Shelley A. Hinsley [3]

[1] Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[2] Department of Life and Environmental Sciences, Bournemouth University, Poole, Dorset BH12 5BB, UK
[3] UK Centre for Ecology and Hydrology, Wallingford, Oxfordshire OX10 8BB, UK
* Correspondence: rhill@bournemouth.ac.uk

Abstract: Spatial data on forest structure, composition, regeneration and deadwood are required for informed assessment of forest condition and subsequent management decisions. Here, we estimate 27 forest metrics from small-footprint full-waveform airborne laser scanning (ALS) data using a random forest (RF) and automated variable selection (Boruta) approach. Modelling was conducted using leaf-off (April) and leaf-on (July) ALS data, both separately and combined. Field data from semi-natural deciduous and managed conifer plantation forests were used to generate the RF models. Based on NRMSE and NBias, overall model accuracies were good, with only two of the best 27 models having an NRMSE > 30% and/or NBias > 15% (Standing deadwood decay class and Number of sapling species). With the exception of the Simpson index of diversity for native trees, both NRMSE and NBias varied by less than ±4.5% points between leaf-on only, leaf-off only and combined leaf-on/leaf-off models per forest metric. However, whilst model performance was similar between ALS datasets, model composition was often very dissimilar in terms of input variables. RF models using leaf-on data showed a dominance of height variables, whilst leaf-off models had a dominance of width variables, reiterating that leaf-on and leaf-off ALS datasets capture different aspects of the forest and that structure and composition across the full vertical profile are highly inter-connected and therefore can be predicted equally well in different ways. A subset of 17 forest metrics was subsequently used to assess favourable conservation status (FCS), as a measure of forest condition. The most accurate RF models relevant to the 17 FCS indicator metrics were used to predict each forest metric across the field site and thresholds defining favourable conditions were applied. Binomial logistic regression was implemented to evaluate predicative accuracy probability relative to the thresholds, which varied from 0.73–0.98 area under the curve (AUC), where 11 of 17 metrics were >0.8. This enabled an index of forest condition (FCS) based on structure, composition, regeneration and deadwood to be mapped across the field site with reasonable certainty. The FCS map closely and consistently corresponded to forest types and stand boundaries, indicating that ALS data offer a feasible approach for forest condition mapping and monitoring to advance forest ecological understanding and improve conservation efforts.

Keywords: remote sensing; lidar; ALS; forest management; conservation; condition; random forest

Citation: Sumnall, M.J.; Hill, R.A.; Hinsley, S.A. Towards Forest Condition Assessment: Evaluating Small-Footprint Full-Waveform Airborne Laser Scanning Data for Deriving Forest Structural and Compositional Metrics. *Remote Sens.* **2022**, *14*, 5081. https://doi.org/10.3390/rs14205081

Academic Editors: Iñigo Molina, Jan Komarek and Marlena Kycko

Received: 9 September 2022
Accepted: 8 October 2022
Published: 11 October 2022

1. Introduction

The spatial data and analysis requirements for modern environmental monitoring and management are increasingly high [1,2]. Forest locations, in particular, are subject to a number of management pressures and external threats [3–5], being exploited as a resource for a range of cultural and economic activities, as well as being important habitats for a variety of organisms and storing substantial amounts of above-ground carbon [6]. All of

these aspects can influence forest condition, the evaluation of which requires a diverse array of indicator metrics to be measured and interpreted.

The assessment of a given forest site cannot include measurements of every facet of the environment due to temporal and logistical constraints, thus the choice of what to measure is critical. These measurements will be indicative of some ecological function or resource [7,8]. At the scale of a stand, indicators are usually placed into one of two categories: those based on the identification of key species or of key structures [9]. The rationale for key structures is that ecosystems containing a variety of structural components are considered likely to have a variety of resources and species that utilise these resources [9].

Numerous approaches have been suggested for the assessment of the conservation conditional status of forested areas at national and international levels (e.g., [10]). Forest condition could be defined in terms of biodiversity, species richness, structural complexity, plant health, ecosystem function and/or productivity, dependent upon the management objective. Furthermore, little consensus exists in the definition of what conditions constitute as 'favourable' within each habitat, and definitions can be regionally specific [11]. Regardless of which, a large number of indicator metrics of favourable conservation status (FCS) have been proposed for this task, for example, the approach defined in [11] required 17 metrics, including tree size distributions, tree species diversity and understorey composition, amongst others.

The amount of data collected and the area sampled on the ground through fieldwork often tend to be quite limited. Indeed, the need to restrict the cost of fieldwork and gather data over large areas within a short time span often results in relatively small areas being sampled using field plots of typically 1 ha or less [12]. This emphasises that field data are a subsampling of the environment, with the potential to miss important changes even over short distances, whilst the feasibility of measuring everything is impractical. Large-scale monitoring requirements are really only feasible though the use of remote sensing technologies (e.g., [13]). With a need to supplement conventional field-based assessment and for standardisation between approaches, an increasing uptake of remote sensing methods is foreseen [14–18].

There have been several reviews which have evaluated the potential role of remote sensing (in particular, hyper-spectral, hyper-spatial and hyper-temporal optical data and active airborne laser scanning or radar data) for vegetation condition assessment (e.g., [19–22]). These reviews demonstrate a broad range of remote sensing approaches used to estimate various indicators and then infer vegetation or habitat condition; for example, measures of forest extent or canopy cover [23,24], composition or diversity [25,26], structure, physiology or function [24,27–30], health, stress or disease [31–35], seasonal dynamics [36], disturbance, degradation, vulnerability or recovery [37–39].

Recent studies have found that airborne laser scanning, hereafter referred to as ALS, can be a powerful predictor of different vegetation attributes, such as height, basal area, stem density and other vegetation structure parameters at the plot or larger scales [40–44]. ALS uses laser pulses to directly measure ground and vegetation height, as well as the vertical distribution of intercepted surfaces, making it an ideal tool for mapping vegetation structure with no saturation at high biomass values [45]. Thus, ALS measurements have been shown to produce more accurate estimates of vegetation structure parameters than other remotely sensed data, because ALS has the ability to penetrate forest canopies and to detect three-dimensional forest structures [46,47]. This is typically accomplished through deriving some statistical relationship between field training data and an ALS metric, although such correlations are often site or ALS acquisition specific (e.g., [48,49]). To date, only one study has used ALS to assess individual tree condition, using structural metrics related to crown density and tree height [50].

Typically, the use of ALS to predict compositional metrics, such as species type, has proven difficult, as ALS metrics are associated with the structural arrangement of features within and below the canopy. Some studies have explored the potential of ALS to model the diversity of insects, spiders and birds [51–53]. However, few studies have explicitly anal-

ysed the relationships between plant species diversity and ALS measurements [25,54,55]. Radiometric information recorded by ALS, commonly referred to as intensity and recorded from an infrared laser, has proven sufficient to classify species in some cases [56], however, differences in acquisition specifications and sensor design seriously impact performance. At the very least, classification generally between deciduous and coniferous vegetation is possible (e.g., [56,57]). The use of intensity has, however, remained contentious.

ALS return intensity (also referred to as amplitude) presents a number of challenges for its use in any analysis, given the nature of typical sensor design. According to [42], it is not possible to compare any two discrete-return intensity values. The research presented in [58] states that intensity data can vary in performance between sensors when used for species classification. In addition, sensors using different wavelengths can also preclude comparison [59]. Recent work on normalising intensity has been proposed, and the results seem promising (e.g., [56]). Other studies have combined ALS datasets with multi- or hyper-spectral data which have allowed the classification of overstorey species [60,61]. Alternatively, general species characteristics have been correlated with ALS derived metrics, such as mean vegetation return height and the diversity of forest species or land cover (e.g., [62–64]).

In general, the area-based statistical models developed for a given dataset are not transferable between acquisitions. For example, in [65] metrics estimated with models calibrated using leaf-on field and ALS data produced erroneous results when applied to leaf-off ALS data, implying there is a difference in the 3D distribution of returns when considering multi-temporal analysis. The combination of datasets from different time-points (e.g., summer and winter) has the potential to yield additional information. Studies have used multiple acquisitions of airborne ALS for the same site at different times in order to exploit the seasonality of different forest species, to estimate understorey presence [66] or to compare the accuracy of terrain mapping [67]. For the latter study, a greater proportion of the returns were from lower portions of the forested area when assessed with leaf-off data than with leaf-on data, due to the absence or presence of foliage. The research outlined in [68] combined leaf-on and leaf-off datasets and was able to exploit the difference in vegetation structure and return intensity between the acquisitions to classify tree species. The authors of [69] stated that above-ground biomass estimates were similar between leaf-on and leaf-off ALS data, but that stratification by species type improved estimates. In [70], a combination of leaf-on and leaf-off data was shown to model diameter at breast height (DBH) and diversity in crown dimensions more accurately, whilst models derived from leaf-on only data performed the poorest in terms of accuracy. Conversely, [71] stated that differences in canopy conditions manifested in leaf-on or leaf-off datasets have an insignificant impact on estimates of above-ground biomass. The model strength was instead dependent on environmental conditions and the modelling method implemented. The authors of [55] estimated 23 forest structural or compositional forest metrics using a combination of leaf-on and/or leaf-off metrics, where only ten of these were best estimated using a combination of both datasets. This implies that there are potentially significant changes in vegetation structure captured by acquisitions from different dates, however, some metrics can be estimated from either.

A more recent development of ALS is full-waveform (FW). This sensor type, as opposed to more conventional discrete-return ALS, provides connected profiles of the three-dimensional scene per pulse, which potentially contain more detailed information about the structure of the illuminated surfaces [72]. Additional processing is required to provide a conventional 3D point cloud. Such a sensor design can potentially return more information from below the canopy [73] and higher return densities than conventional discrete-return ALS, as noted in [74], which is of interest in forest studies [75–77]. From the research presented in [55], the analysis of full-waveform (FW) airborne ALS data can yield estimates of both compositional and structural metrics of substantial potential for assessing forests for a variety of possible applications.

Thus, utilising FW ALS data and a combination of leaf-on and leaf-off acquisitions could conceivably yield the highest potential for estimating forest metrics for forest condition assessment. The goal of the current research is to evaluate the capabilities of FW ALS to provide a range of forest structural and compositional metrics across the full vertical profile and to assess the accuracy of these estimates for forest condition monitoring. This research will assess the potential utility that ALS acquisition can provide in future forest monitoring applications, specifically for forest contexts in southern Britain. The field site comprises both semi-natural and managed plantation forests in close proximity. Twenty-seven field metrics for the semi-natural forest have been defined for this location from previous research, which is outlined in [11,55]. The objectives of the current research were therefore to: (i) estimate the accuracy of these structural and compositional metrics across a range of forest types; (ii) evaluate if there are any benefits to the use of leaf-on, leaf-off or combined ALS acquisitions in the estimation of forest metrics; (iii) compare the relative merit of height, amplitude and width ALS metrics in derived models; and (iv) assess the use of a subset of these metrics for forest favourable conservation status (FCS) monitoring.

2. Materials and Methods

2.1. Study Site

The study site is located within the New Forest National Park in the south of the UK (50°50′N; 1°30′W). Many land cover types exist in close proximity and are managed for multiple land uses. Much of the forested areas are managed plantations, although semi-natural forest also exists [78]. The analysis was focused on a ca. 22 km^2 area in the proximity of the towns of Brockenhurst, Beaulieu and Lyndhurst. The underlying terrain within this site varied only gently, with typical elevations occurring between 5 m and 45 m above sea level. A number of the plantation forest locations are managed enclosures in order to reduce ungulate browsing. Unenclosed areas are not subject to felling operations and are permanently open to grazing by large ungulates (e.g., ponies and deer). Enclosed areas are typically located in plantation coniferous forest.

The area contains several types of semi-natural and plantation coniferous and deciduous forests (as stated in [79]), which present a wide range of available structural and compositional variation, such as canopy gaps and the presence of deadwood or understorey. Deciduous species include: oaks (*Quercus robur* and *Quercus petraea*), common beech (*Fagus sylvatica*), common alder (*Alnus glutinosa*), silver birch (*Betula pendula*), sweet chestnut (*Castanea sativa*) and holly (*Ilex aquifolium*). Coniferous species include: Corsican pine (*Pinus nigra* var. maritime), Scots pine (*Pinus sylvestris*), Douglas fir (*Pseudotsuga menziesii*) and Norway spruce (*Picea abies*). The distribution of the broad habitat types as classified by the Forestry Commission are summarised in Figure 1.

2.2. Definition of Field Metrics

A number of field-based indicators of FCS were defined in [11] for the New Forest (Table 1). They identified 17 ecological indicators that related to structure and composition, deadwood, tree regeneration and ground vegetation (from: [80–83]). These indicators and an additional 10 metrics were recorded as part of a later study [55] by a plot-based method (Table 1). A summary of field measured ranges for each of the 27 indicator metrics is presented in Table S1.

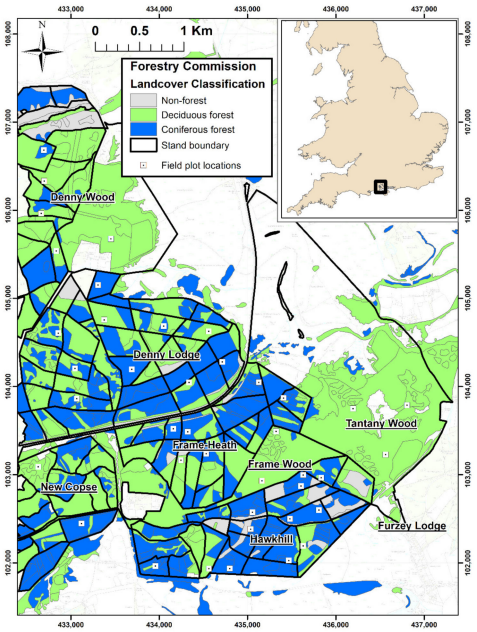

Figure 1. The complete study area extent in the New Forest, UK. Forestry Commission primary habitat codes are provided for reference illustrating the locations of coniferous and deciduous forest. Black lines represent stand boundaries and grey lines represent subcompartment boundaries. Field plot locations are also indicated.

Table 1. List of key ecological factors recorded for each of the field plots installed within the New Forest field site. The acronym DBH stands for diameter at breast height. The indicators marked with an * were used as determinants of 'favourable conservation status' following from [11].

Key Ecological Factor	Indicator
Forest canopy structure	* Number of trees (ha^{-1}) * Basal area (m^2ha^{-1}) * Mean DBH (cm) * Standard deviation of tree diameters (cm) * Percentage of big trees (DBH > 40 cm) (%) * Mean tree height (m) Mean tree stem spacing (m) Mean height to the living crown (m) Mean crown horizontal area (m^2) Sum of crown horizontal area (m^2)
Forest canopy composition	Number of tree species Number of native tree stems * Shannon–Wiener index for native trees Simpson index of diversity
Forest regeneration: saplings	* Number of saplings (ha^{-1}) * Number of native saplings (ha^{-1}) Number of sapling species
Forest regeneration: seedlings	* Number of seedlings (ha^{-1}) * Number of native seedlings (ha^{-1}) Number of seedling species * Shannon–Wiener index for native seedlings
Forest deadwood	* Volume of downed deadwood (m^3ha^{-1}) * Downed deadwood decay class (0–1) * Volume of standing deadwood (m^3ha^{-1}) * Standing deadwood decay class (0–1)
Ground layer	* Number of ground vegetation species (ha^{-1}) Percentage bare ground cover (%)

2.3. Field Data Collection

Using pre-existing forest inventory data (made available from the Forestry Commission and described in [55]), the woodland areas of the study site were split into coniferous, deciduous and mixed woodland compartments and stratified according to forest inventory information. A total of 41 field plots were then located randomly across this stratification to enumerate a range of forest types and canopy conditions. An initial 21 plots were visited in the summer (June to September) of 2010, with the remaining 20 plots visited in a further field campaign in the summer of 2012 (June to October). Each of the field plots was established at a minimum of 10 m away from a stand boundary or any non-forested areas in order to reduce any potential edge effects. Plots in coniferous stands totaled 20, whilst 16 were located in deciduous stands, and 5 plots were located in mixed stands.

We installed 41 north-oriented square plots sized 30 m × 30 m, with a square 10 m × 10 m subplot in the south-west corner. Field plot corners were located using a combination of Leica GPS 500 (Leica Geosystems Ltd, Milton Keynes, UK) and Sokkia 6F total station (Sokkia Topcon Company Ltd, Kangawa, Japan). Leica Geo-office software (version 8.2) was used for post-processing tasks.

Total horizontal positional error was calculated as ≤0.08 m. A total of 27 metrics (as in Table 1) were recorded in each of the 41 field plots. Plot-level totals and averages were calculated for each field recorded metric. Within each plot, tree stems with a diameter at breast height (DBH) greater than 10 cm were considered as a tree. Where appropriate, DBH was measured at a height of 1.4 m above ground. Measurements per tree included: stem DBH, canopy top height (m) and species (with Scots pine, common alder, oak, beech, silver birch, holly and sweet chestnut considered as native or naturalised species). Vertical height measurements were calculated using a clinometer. Plot-level basal area was calculated by summing the area of a circle calculation applied to each DBH measurement. The number of saplings and their species types were recorded within each field plot. Saplings were defined as tree stems > 1.3 m in height with DBH < 10 cm. The total number of seedlings and their species types were also recorded within the 10 m × 10 m subplot. Seedlings were defined as tree stems < 1.3 m in height. The number of ground vegetation (i.e., vascular plant) species was also recorded within the subplot. In the field, height to the live crown (HTLC) was defined as the distance between the ground and the lowest live branch, and measured using a clinometer. Crown diameter was recorded for the horizontal projection north to south and then east to west. Measurements were made at ground level with a tape held at the edge of the crown being measured. The edges of the crown were defined as the perimeter of the crown that was visible and identifiable from the ground directly below. Crown horizontal area was calculated as the area of an ellipse.

The Shannon–Wiener index (SH) [84] for all native trees and saplings was calculated as:

$$\mathrm{SH} = \sum_{i=1}^{n} p_i \, log_e \, p_i, \quad (1)$$

where p_i = the proportion of individuals (plot stem number) in $i = 1, \ldots, n$, where n is the number of species. The Simpson index [85] is expressed as:

$$\mathrm{SI} = 1 - \sum_{i=1}^{n} x_i^2, \quad (2)$$

where x_i is the relative abundance of the i-th species.

Each of the standing deadwood items, or snags, within a field plot was recorded. Snags were defined as standing deadwood > 10 cm DBH [86]. Snag volume was calculated using the formula for determining cylindrical volume using height and DBH measurements. Downed deadwood (DDW) was defined as deadwood logs or branches of at least 10 cm diameter lying on the ground [86]. Measurements for DDW were made in the 10 m × 10 m subplot only. Length and girth around the two ends of the log were recorded. Estimates of DDW volume were determined using the equation for a frustum of a cone. Deadwood decay states for snags and DDW were divided into three decay classes according to the following criteria, as defined in [11]: (i) logs with a low decay state, no surface breakdown,

bark still intact, wood structure firm; (ii) logs with a moderate decay state, with some surface breakdown, wood structure weaker but bole mostly sound; and (iii) logs with high decay state, extensive surface breakdown, bark mostly absent, bole with no sound wood present and colonised with vegetation. A size-weighted average decay class score was then calculated at the plot level, and rescaled to a value between 0 and 1.

2.4. Airborne Laser Scanner Datasets

Small-footprint FW ALS data were acquired for the study area in 2010 under leaf-off (8 April) and leaf-on (6 July) conditions. The ALS instrument used was the Leica ALS50-II airborne laser scanner. On both dates, the ALS data were acquired at a flying altitude of ~1600 m, with a pulse repetition frequency (PRF) of 147 kHz, a beam divergence of 0.22 mr and a scan half angle of 10°. Flight lines were established with approximately 30% overlap. The geometric accuracy for the scanner is stated by the manufacturer (Leica Geosytems Ltd., St. Gallen, Switzerland) as a nominal vertical accuracy of 0.05 m to 0.10 m, and horizontal accuracy of 0.13 m to 0.61 m for flying altitudes of 200 m to 6000 m. With the chosen flight and sensor configuration, the average pulse density for both the leaf-on and leaf-off data was approximately 2.5 pulses per m^2 on the ground at nadir.

2.5. Airborne Laser Scanner Processing

The FW ALS data (supplied in LAS version 1.3 format) required pre-processing steps in order to derive a conventional point cloud. The Sorted Pulse Data library (SPDlib) (version 1.0.0) [87,88] was used to perform Gaussian decomposition upon each of the returned waveforms to derive a point cloud. Between 1 and 10 returns per pulse were derived using this approach. In addition to elevation, each return had an associated value relating to waveform amplitude (or intensity) and peak-width. Ground elevation returns were classified through a progressive morphological filter within SPDlib as outlined in [89]. Above-ground heights were then calculated by subtracting ALS elevation returns from an interpolated terrain surface created from ground returns (using the nearest neighbour interpolation method). Intensity/amplitude range normalisation was implemented using the method documented in [90].

In order to derive ALS variables from the field plot areas, all ALS returns which lay within field plot boundaries were extracted from the datasets. Nine key variables were generated from the complete vertical distribution of the return height, amplitude and width values: mean, median, minimum, maximum, standard deviation (StDev), variance, absolute deviation (AbsDev), skewness and kurtosis. These metrics were selected for comparison to previous work [55]. Given the potential importance of amplitude and width metrics from returns near ground (<0.2 m; as in [55]), the nine variables for both were also calculated for all returns below 0.2 m height above ground. Percentiles were calculated using all returns from the complete vertical distribution of return heights, plus amplitude and width data, every 5th percentile (excluding the 50th and 100th percentiles as these are the median and maximum, respectively). In addition, dominant height was calculated as the arithmetic mean of heights of returns ≥80th percentile (as in [91]), and canopy cover (CC) was calculated as follows:

$$\mathrm{CC} = \left(\frac{\sum R_{>T}}{\sum R_{all}} \right) \times 100, \tag{3}$$

where $R_{>T}$ denotes those returns above a threshold height T, and R_{all} is all returns. Here, T was set to 1 m height above ground. These variables were generated for both leaf-on and leaf-off datasets, resulting in 99 metrics for each (198 in total).

2.6. Statistical Analysis

A random forest (RF) modelling approach was selected in order to limit the potential issues caused by high correlations between many of the ALS metrics [92,93]. In addition,

linear regression methods can have difficulty in structurally complex forest when compared to machine learning approaches [94].

All statistical analyses were performed using R software (version 4.1.1) (http://www.r-project.org/) (accessed on 10 October 2021). Statistical analysis was conducted using the Boruta package (version 7.0.0, [95]) for automated variable selection and the caret (version 6.0-91, [96]) and randomForest (version 4.7-1, [97]) packages for model fitting. Logistic regression analysis was performed using the pscl [98] and pROC [99] packages.

The RF approach [100] is a widely used machine learning technique using an ensemble method to perform classification or regression tasks (e.g., [101–103]). Multiple RF models were developed for each of the 27 indicator metrics. This was carried out for each ALS dataset: (i) leaf-on only; (ii) leaf-off only; and (iii) a combination of leaf-on and leaf-off (i.e., 27 models per dataset). Analyses were performed for all 41 field plots for each of the three datasets irrespective of dominant tree species type (i.e., deciduous, coniferous or mixed). No dedicated validation dataset was used here. A lower number of variables relative to the number of observations will reduce the likelihood of overfitting, thus the Bortua R package was used as a means to automatically select important variables. This package is a wrapper built around the RF algorithm [95]. Briefly, the approach duplicates and randomly reorders the training dataset and implements an RF classifier. This approach was repeated 500 times, and the predictor accuracy/impurity was assessed through the calculation of a Z-score. Comparisons of every iteration select a variable as important if it outperforms its duplicate. At the end of this process, the seven most important predictor variables (if available) were used as input into the RF model.

Tuning and construction of the RF model was implemented using the caret R package. A 10-fold cross-validation control was used in all permutations here. Tuning consisted of three elements evaluated in sequence: (i) MTRY: the number of variables randomly sampled as candidates at each branch split; (ii) maxnodes: the maximum number of terminal nodes (limiting the potential RF tree size); and (iii) ntrees: the number of trees to grow. The ranges of values tested as part of the tuning process were: 1–20; 3–20; and 250–2000, respectively. The optimal value for each of the three values was taken to be the one which produced the lowest root mean square error (RMSE) and mean average error (MAE), in addition to the highest adjusted coefficient of determination (R^2). Once the three tuning parameters had been determined, the RF model was then run using the predictor variable(s) identified through the variable selection procedure.

The predicted values from the statistical models were quantified by calculating the adjusted coefficient of determination (adj. R^2), absolute and normalised root mean square error (RMSE and NRMSE, respectively) and absolute and normalised bias (Bias and NBias). These functions are expressed as:

$$\text{adj. } R^2 = 1 - \frac{(\sum_i (x_i - y_i)^2)/(n-k)}{(\sum_i (x_i - \overline{x})^2)/(n-1)}, \tag{4}$$

$$\text{RMSE } \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}, \tag{5}$$

$$\text{NRMSE} = \frac{RMSE}{(x_{max} - x_{min})}, \tag{6}$$

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i), \tag{7}$$

$$\text{NBias} = \frac{Bias}{(x_{max} - x_{min})}, \tag{8}$$

where y is the predicted value, x is the observed field measured value for the individual plot i, n is the number of samples, $\overline{x}$ is the mean of field measured values and k is the number of predictor variables in the model. Here, we define acceptable model accuracy in terms of NRMSE, where: <10% indicates a high correspondence between predicted and observed,

10–20% denotes a small difference, 20–30% is a moderate difference and an NRMSE $\geq$ 30% represents a large and unacceptable difference [104,105]. Likewise, an NBias of $\geq$15% was considered unacceptable model accuracy [103].

2.7. Mapping Forest Favourable Conservation Status (FCS)

Favourable conservation status for each of the 41 field plots was calculated with regard to the threshold values given in [11], and accounting for scale (see Table S2). The most accurate RF models relevant to the 17 favourable status indicator metrics were then used to predict each field metric per plot, which was then assessed relative to the threshold values. The predictions were assigned a value of 0 or 1, depending on whether the estimate was under or over the threshold value for favourable status, giving a potential range of 0–17, where a maximum score means a high FCS status.

In order to evaluate the ability of the RF models to estimate the metrics to a suitable accuracy, specifically to an accuracy level which would allow us to estimate if an area has a metric above or below the threshold value defined in Table S2, a binomial logistic regression was implemented on the results. Both field plot data and predictions from the RF models were assigned a 0 or 1 depending on threshold value. These values were entered into a generalised linear model binomial logistic regression, in order to evaluate the ability of the RF models to predict a dichotomous dependent variable. The proportion of true positives (sensitivity) and true negatives (specificity) was calculated. As a more general metric to evaluate how well the logistic regression model does at classifying the data, we calculated the area under the curve (AUC) using the pROC package [99]. AUC values can vary from 0.5 (which suggests no relationship) to 1.0 (which suggests observed and predicted values are identical), where higher values are considered better. Here, we classify a value of 0.7 or above as acceptable [106].

Subsequently, each of the 17 indicator metrics was estimated for 30 $\times$ 30 m cells across the study site for the entire area of ALS data acquisition, with each assessed per cell against the relevant FCS threshold. Each cell per indicator metric was assigned a value of 0 or 1, depending on whether the estimate was under or over the threshold value. These values were summed giving an index between 0 and 17 per cell.

3. Results

3.1. Summary of Models Using Metrics from the Leaf-On Acquisition Only

Model tuning for all of the final RF models produced from metrics generated from the leaf-on ALS acquisition is summarised for reference in Table S3. The predictor variables used in each of the 27 RF models are summarised in Table S4.

The final RF models each used between one and seven input ALS variables. In total across the 27 models, height variables were included 55 times, width variables 43 times and amplitude variables 26 times. The most commonly occurring ALS variables were AbsDev of return height (six models), Variance of return height (seven models) and StDev of return height (eight models). Across the 27 RF models, height variables were used in 22 models, width variables in 16 models and amplitude variables in 13 models. Only four models exclusively used height predictors (Shannon–Weiner index for native trees, Downed deadwood decay class, Number of seedlings and Number of ground vegetation species), three models used only amplitude predictors (Simpson index of diversity, Number of saplings and Standing deadwood decay class) and one used width only (Number of sapling species). There were nine models that combined height and width predictors, four that combined height and amplitude predictors and only one model that combined amplitude with width predictors. There were five models which combined height, width and amplitude predictor variables.

The RF models created for each of the 27 forest metrics were assessed against field data, summarised in Table 2. Across all models, the adj. R^2 ranged from 0.59 to 0.82. In terms of NRMSE, eight models were below 15%, 18 were between 15 and 30% and one was above 30% (Standing deadwood decay class), here implying an unsuitable model. The

majority of models had an NBias of below 15%, while three did not (Standing deadwood decay class, Simpson index of diversity and Number of sapling species). Thus, in total three models of the 27 were considered of unacceptable accuracy.

Table 2. Summary statistics of each random forest model using metrics from the leaf-on acquisition, with absolute and normalised root mean square error (RMSE and NRMSE); absolute bias and normalised bias (NBias); and the adjusted coefficient of determination (Adj. R^2).

Variables	RMSE	NRMSE (%)	Bias	NBias (%)	Adj. R^2
Number of trees (plot^{-1})	12.04	14.50	0.10	0.12	0.78
Mean DBH (cm)	6.21	14.39	0.04	0.10	0.82
Standard deviation of tree diameters (cm)	7.57	15.53	−0.06	−0.13	0.70
Basal area (m^{2}plot^{-1})	0.65	20.32	0.00	0.06	0.77
Percentage of big trees (DBH > 40 cm)	15.49	18.07	−1.04	−1.33	0.80
Mean tree height (m)	2.41	11.89	−0.05	−0.26	0.80
Mean tree stem spacing (m)	0.98	14.85	0.11	1.70	0.70
Mean height to the living crown (m)	2.20	19.49	0.14	1.28	0.78
Mean crown horizontal area (m^2)	18.50	17.31	0.42	0.40	0.79
Total crown horizontal area (m^2)	263.43	16.38	0.18	0.01	0.79
Number of tree species	1.25	20.84	0.08	1.28	0.59
Number of native tree stems	8.95	17.21	0.15	0.30	0.79
Shannon–Wiener index for native trees	0.37	25.03	0.01	0.47	0.60
Simpson index of diversity	0.22	21.82	0.22	22.38	0.73
Number of saplings (plot^{-1})	14.58	13.50	−2.60	−2.43	0.61
Number of native saplings (plot^{-1})	4.22	12.42	−2.01	−6.09	0.79
Number of sapling species	0.97	24.31	−0.50	−16.67	0.64
Number of seedlings (plot^{-1})	230.12	26.63	−4.14	−0.48	0.66
Number of native seedlings (plot^{-1})	165.10	22.93	0.56	0.08	0.59
Shannon–Wiener index for native seedlings	0.40	24.84	−0.14	−9.77	0.73
Number of seedling species	1.90	14.59	−0.12	−0.99	0.81
Volume of downed deadwood (m^{3}plot^{-1})	4.93	12.86	0.17	0.44	0.60
Downed deadwood decay class	0.11	24.30	0.00	0.29	0.70
Volume of standing deadwood (m^{3}plot^{-1})	2.29	22.66	−0.50	−4.97	0.64
Standing deadwood decay class	0.33	33.27	−0.15	−22.34	0.61
Number of ground vegetation species	1.24	20.72	−0.01	−0.20	0.73
Percentage bare ground cover (%)	24.89	27.66	−4.15	−4.64	0.80

3.2. Summary of Models Using Metrics from the Leaf-Off Acquisition Only

Final RF model tuning parameters for the leaf-off ALS acquisition are provided in Table S5. Predictor variables used in each model are summarised in Table S6.

The final RF models created using leaf-off ALS data also had between one and seven input variables. In contrast with the leaf-on models, there was a greater prevalence of width variables. Thus, across the 27 models, width variables were included 54 times, and both height and amplitude variables 36 times. The most commonly occurring ALS variables were Kurtosis of return width (six models), Variance of return width and StDev of return width (seven models) and AbsDev of return width (eight models). Height, width and amplitude predictors were included in 14, 17 and 14 models, respectively, and used exclusively as input for one (Simpson index of diversity), six (Standard deviation of tree diameter, Number of native saplings, Standing deadwood decay class, Number of native seedlings, Mean height to the living crown and Mean crown horizontal area) and three (Total crown horizontal area, Number of sapling species, Percentage bare ground cover) of the RF models, respectively. A combination of height and width variables formed the inputs to three models, height and amplitude for six models and amplitude and width for four models. All three predictor variable types were included in four models.

Comparisons of RF model predictions against field data (Table 3) produced models with adj. R^2 which ranged from 0.59 to 0.89. In terms of NRMSE values, eight models were below 15%, 17 were between 15 and 30% and two were above 30% (Standing deadwood decay class and Percentage bare ground cover). The majority of models (25 of 27) had an NBias of less than 15%, the exceptions being: Standing deadwood decay class and Number

of sapling species. Thus, as with the leaf-on only dataset, three of the 27 RF models were of unacceptable accuracy, of which two were for the same forest metric.

Table 3. Summary statistics of each random forest model using metrics from the leaf-off acquisition, with absolute and normalised root mean square error (RMSE and NRMSE); absolute bias and normalised bias (NBias); and the adjusted coefficient of determination (Adj. R^2).

Variables	RMSE	NRMSE (%)	Bias	NBias (%)	Adj. R^2
Number of trees (plot−1)	11.22	13.52	0.04	0.05	0.83
Mean DBH (cm)	7.04	16.32	0.26	0.60	0.73
Standard deviation of tree diameters (cm)	7.00	14.36	0.03	0.05	0.79
Basal area (m^2plot−1)	0.74	23.05	0.01	0.29	0.62
Percentage of big trees (DBH > 40 cm)	14.27	16.65	−0.34	−0.43	0.76
Mean tree height (m)	2.55	12.58	0.09	0.44	0.85
Mean tree stem spacing (m)	0.97	14.66	0.00	0.01	0.73
Mean height to the living crown (m)	1.83	16.19	0.06	0.57	0.89
Mean crown horizontal area (m^2)	21.00	19.65	0.02	0.02	0.76
Total crown horizontal area (m^2)	298.18	18.65	−3.50	−0.22	0.81
Number of tree species	1.04	17.29	0.05	0.85	0.78
Number of native tree stems	9.29	17.86	0.28	0.54	0.73
Shannon–Wiener index for native trees	0.34	23.10	−0.05	−3.88	0.69
Simpson index of diversity	0.22	21.96	−0.04	−4.66	0.59
Number of saplings ($plot^{-1}$)	12.19	11.28	−3.72	−3.48	0.63
Number of native saplings ($plot^{-1}$)	4.08	12.01	−2.39	−7.25	0.59
Number of sapling species	1.11	27.68	−0.60	−19.90	0.61
Number of seedlings ($plot^{-1}$)	209.24	24.22	−3.71	−0.43	0.73
Number of native seedlings ($plot^{-1}$)	148.42	20.61	2.42	0.34	0.69
Shannon–Wiener index for native seedlings	0.37	22.86	−0.14	−10.27	0.75
Number of seedling species	1.89	14.57	−0.05	−0.41	0.73
Volume of downed deadwood (m^3plot^{-1})	4.87	12.69	0.15	0.39	0.69
Downed deadwood decay class	0.11	22.70	0.00	−0.41	0.70
Volume of standing deadwood (m^3plot^{-1})	2.10	20.75	−0.38	−3.79	0.73
Standing deadwood decay class	0.33	32.93	−0.14	−21.14	0.70
Number of ground vegetation species	1.24	20.74	−0.01	−0.09	0.72
Percentage bare ground cover (%)	28.52	31.68	−4.95	−5.53	0.69

3.3. Summary of Models Using Metrics from Both the Leaf-On and Leaf-Off Acquisition

Final RF model tuning parameters for the combination of leaf-on and leaf-off ALS acquisition are provided for reference in Table S7. Predictor variables used in each model are summarised in Table S8.

As with both the separate leaf-on and leaf-off models, the combined leaf-on/leaf-off RF models each had between one and seven input ALS variables. In total across the 27 models, width variables were included 55 times, height variables 45 times and amplitude variables 32 times, of which 50 were from leaf-on data and 82 from leaf-off data. The most common input variables (regardless of leaf-on or leaf-off) were: Maximum of return height, Variance of return height, AbsDev of return height and StDev of width (6 models), StDev of return height and Kurtosis of return width (7 models), Variance of return width (8 models), and AbsDev of return width (10 models). Height, width or amplitude ALS variables were present in 16, 18 and 14 models, respectively. Those models using exclusively height, width or amplitude variables constituted four (Mean tree height, Number of tree species, Shannon–Weiner index for native trees, Number of ground vegetation species), seven (Standard deviation of tree diameters, Mean height to living crown, Mean crown horizontal area, Number of seedlings, Number of native seedlings, Number of sapling species, Downed deadwood decay class) and two (Total crown horizontal area and Number of saplings) RF models, respectively. Models combining height and width, height and amplitude or amplitude and width variables accounted for two, three and two models, respectively. Models produced using all three predictor variable types accounted for seven models.

Predictor variables derived from leaf-on acquisition were present in 20 models, whilst leaf-off variables were present in 23 models. Both leaf-on and leaf-off variables were present together in 16 RF models, whilst four models were derived using leaf-on only variables

and seven using leaf-off only variables. For the four models containing only leaf-on ALS data, these were either exactly the same (Mean tree spacing, Number of saplings), a subset (Percentage bare ground) or a highly correlated alternative (Number of sapling species) to the input variables used in the leaf-on only models. By contrast, for the seven models containing only leaf-off ALS variables, four of the models had the same or a subset of the input variables used in the leaf-off only models (Standard deviation of tree diameter, Mean height to the living crown, Number of native seedlings and Number of seedling species), whilst the models for Number of tree species, Downed deadwood decay class and Shannon–Wiener index for native trees had very different input variables to their leaf-off only counterparts. As many of the input variables are highly correlated, in addition to the large number of predictor variables available, and the random subsetting of predictors in the Boruta package over a finite number of iterations (500), this could explain the different predictors being selected. It is conceivable that the predictors indicated in Table S6 would be repeated if there were enough Boruta iterations. Each of the predictions from these models is illustrated in Figure S1.

Comparisons of RF model predictions against field data (Table 4) produced models with adj. R^2 which ranged from 0.56 to 0.89. In terms of NRMSE values, nine models were below 15%, 17 were between 15 and 30% and only Standing deadwood decay class was above 30%, here implying an unsuitable model. The majority of models (25 of 27) had an NBias of less than 15%, with the exceptions being Standing deadwood decay class and Number of sapling species.

Table 4. Summary statistics of each random forest model using metrics from the leaf-on and leaf-off acquisition, with absolute and normalised root mean square error (RMSE and NRMSE); absolute bias and normalised bias (NBias); and the adjusted coefficient of determination (Adj. R^2).

Variables	RMSE	NRMSE (%)	Bias	NBias (%)	Adj. R^2
Number of trees (plot^{-1})	9.86	11.88	0.10	0.12	0.89
Mean DBH (cm)	6.32	14.67	0.14	0.32	0.79
Standard deviation of tree diameters (cm)	7.17	14.71	0.07	0.15	0.77
Basal area (m^{2}plot^{-1})	0.73	22.69	0.00	−0.14	0.75
Percentage of big trees (DBH > 40 cm)	13.88	16.19	−0.46	−0.59	0.85
Mean tree height (m)	2.84	13.99	−0.03	−0.12	0.79
Mean tree stem spacing (m)	0.98	14.85	0.11	1.70	0.70
Mean height to the living crown (m)	1.83	16.19	0.09	0.81	0.89
Mean crown horizontal area (m^2)	19.56	18.30	0.42	0.40	0.77
Total crown horizontal area (m^2)	275.02	17.10	−2.44	−0.15	0.82
Number of tree species	1.23	20.42	0.10	1.71	0.58
Number of native tree stems	8.10	15.58	0.15	0.30	0.76
Shannon–Wiener index for native trees	0.36	24.32	−0.02	−1.19	0.64
Simpson index of diversity	0.18	18.30	−0.01	−0.69	0.66
Number of saplings (plot^{-1})	14.58	13.50	−2.60	−2.43	0.61
Number of native saplings (plot^{-1})	4.58	13.48	−2.08	−6.31	0.60
Number of sapling species	1.09	27.31	−0.59	−19.70	0.69
Number of seedlings (plot^{-1})	195.51	22.63	−0.46	−0.05	0.70
Number of native seedlings (plot^{-1})	157.01	21.81	1.38	0.19	0.68
Shannon–Wiener index for native seedlings	0.36	22.65	−0.13	−9.49	0.77
Number of seedling species	1.89	14.57	−0.05	−0.41	0.73
Volume of downed deadwood (m^{3}plot^{-1})	5.15	13.42	0.18	0.46	0.56
Downed deadwood decay class	0.11	23.88	0.00	0.07	0.65
Volume of standing deadwood (m^{3}plot^{-1})	2.12	20.93	−0.36	−3.52	0.77
Standing deadwood decay class	0.31	30.95	−0.15	−22.68	0.68
Number of ground vegetation species	1.30	21.61	−0.01	−0.10	0.63
Percentage bare ground cover (%)	25.82	28.69	−4.28	−4.78	0.70

3.4. Best Overall Models

Comparing the leaf-on only and leaf-off only RF models, based on NRMSE and NBias, seven models were better using leaf-on data, 11 were better using leaf-off data and nine models were inconsistent between the two ALS datasets. However, in only two models (Number of sapling species and Volume of standing deadwood) were both the NRMSE and

NBias different by >1% point between the two datasets. Combining leaf-on and leaf-off data improved ten leaf-on only models and nine leaf-off only models, but only in three cases were both NRMSE and NBias increased by >1% point (Volume of standing deadwood for the leaf-on model and the Simpson index of diversity for both leaf-on and leaf-off models). Therefore, overall, with the exception of Simpson index of diversity which was modelled with a very high NBias using leaf-on only data (22.38%), there was very little or consistent improvement between the leaf-on only, leaf-off only and combined datasets in RF model output quality assessed by NRMSE and NBias.

The most accurate models overall for each of the 27 forest metrics, in terms of minimising NRMSE and NBias, and maximising adj. R^2 (summarised in Table 5 for the 27 field metrics), were derived from across all ALS datasets. Models derived from leaf-on only ALS data (Section 3.1) constitute 8 of the best models, those derived from the leaf-off data (Section 3.2) constitute 12 of the best models and those derived from a combination of leaf-on and leaf-off ALS data (Section 3.3) account for 7 models.

Table 5. Overall best models from leaf-on only, leaf-off only or combined leaf-on and leaf-off metrics (Combined). Model selection was determined by minimising RMSE/NRMSE and bias/NBias and maximising adj. R^2 values.

Variables	Best Model	NRMSE (%)	NBias (%)
Number of trees (plot^{-1})	Combined	11.88	0.12
Mean DBH (cm)	Leaf-on only	14.39	0.10
Standard deviation of tree diameters (cm)	Leaf-off only	14.36	0.05
Basal area (m^{2}plot^{-1})	Leaf-on only	20.32	0.06
Percentage of big trees (DBH > 40 cm)	Combined	16.19	−0.59
Mean tree height (m)	Leaf-on only	11.89	−0.26
Mean tree stem spacing (m)	Leaf-off only	14.66	0.01
Mean height to the living crown (m)	Leaf-off only	16.19	0.57
Mean crown horizontal area (m^2)	Leaf-on only	17.31	0.40
Total crown horizontal area (m^2)	Leaf-on only	16.38	0.01
Number of tree species	Leaf-off only	17.29	0.85
Number of native tree stems	Combined	15.58	0.30
Shannon–Wiener index for native trees	Leaf-off only	23.10	−3.88
Simpson index of diversity	Leaf-off only	21.96	−4.66
Number of saplings (plot^{-1})	Leaf-off only	11.28	−3.48
Number of native saplings (plot^{-1})	Leaf-on only	12.42	−6.09
Number of sapling species	Leaf-on only	24.31	−16.67
Number of seedlings (plot^{-1})	Combined	22.63	−0.05
Number of native seedlings (plot^{-1})	Leaf-off only	20.61	0.34
Shannon–Wiener index for native seedlings	Combined	22.65	−9.49
Number of seedling species	Leaf-off only	14.57	−0.41
Volume of downed deadwood (m^{3}plot^{-1})	Leaf-off only	12.69	0.39
Downed deadwood decay class	Leaf-off only	22.70	−0.41
Volume of standing deadwood (m^{3}plot^{-1})	Combined	20.93	−3.52
Standing deadwood decay class	Combined	30.95	−22.68
Number of ground vegetation species	Leaf-off only	20.74	−0.09
Percentage bare ground cover (%)	Leaf-on only	27.66	−4.64

Overall, slightly lower NRMSE and NBias values were achieved through this process of best model selection. For example, the difference in NRMSE between the leaf-on, leaf-off and combined leaf-on/leaf-off models per forest metric varies between only 0.02 and 4.02 percentage points. The average NRMSE across the 27 models was as follows: 19.57% for leaf-on, 19.28% for leaf-off, 19.06% for combined leaf-on/leaf-off and 18.36% for the best set of models from across all datasets. In terms of NRMSE for the best models, 26 of the models had a value below 30% (nine < 15%), with the exception being Standing deadwood decay class. For NBias, 25 of 27 models were <15%, with the exceptions being Standing deadwood decay class and Number of sapling species. These were therefore the only two RF models considered to be of unacceptable quality.

3.5. Assessment of Accuracy with Regard to Indicators of 'Favourable Conservation Status'

Favourable status values for the 41 field plots ranged from 3 to 11, where lower values on average were observed for plots located in coniferous stands (mean 6.8; standard error 0.4; n = 20), and higher values for sites in deciduous dominated stands (mean 8.6; standard error 0.5; n = 16) or mixed stands (mean 8.2; standard error 0.7; n = 5). The results of a t-test indicated that the FCS index means from coniferous and deciduous plots were significantly different ($p < 0.05$).

The most accurate RF models (as in Table 5) relevant to the 17 FCS indicator metrics were used to predict each field metric, which were then assessed relative to the threshold values. Binomial logistic regression was implemented, and the results are summarised in Table 6. Overall accuracy of the logistic regression per indicator metric varied between 51 to 100 percent correct (where 11 of 17 were >80% correct). The lowest accuracy (<65%) was observed for the Shannon–Wiener index for native trees and Volume of downed deadwood. All AUC values were above 0.7, implying acceptable discrimination for predicted values. AUC values ranged from 0.71 to 0.98. In addition, the AUC calculated for the Shannon–Wiener index for native seedlings was determined to be rank-deficient and potentially misleading, implying that one or more of the predictor variables were not linearly independent. All of the field measurements for Number of seedlings and Number of native seedlings did not exceed the threshold value and thus were always 0, preventing AUC calculation.

Table 6. Summary of logistic regression applied to the estimates from models created from leaf-on, leaf-off or combined models for the 17 indicator metrics identified in Table 1 (and Table S2). Predicted values were assessed against these indicator metric targets and assigned a 1 if they exceeded this, or a 0 if they did not. These values were then assessed against the equivalent field value. (# denotes the area under the curve (AUC) model was rank-deficient and therefore the fit may be misleading; - denotes no model could be run in these instances, all predicted values were below the FCS threshold and were therefore zero).

Variables	Overall Correct (%)	Sensitivity (%)	Specificity (%)	AUC (0–1)
Number of trees (plot^{-1})	82.93	90.48	75.00	0.87
Shannon–Wiener index for native trees	63.41	58.33	70.59	0.71
Basal area (m^{2}plot^{-1})	90.24	96.55	75.00	0.92
Mean DBH (cm)	70.73	86.36	52.63	0.86
Standard deviation of tree diameters (cm)	80.49	100.00	0.00	0.75
Percentage of big trees (DBH > 40 cm)	82.11	100.00	5.26	0.89
Mean tree height (m)	92.68	100.00	0.00	0.93
Number of saplings (plot^{-1})	80.49	40.00	86.11	0.86
Number of native saplings (plot^{-1})	95.12	100.00	95.00	0.98
Volume of downed deadwood (m^{3}plot^{-1})	51.22	100.00	50.00	0.89
Downed deadwood decay class	97.56	100.00	0.00	0.76
Volume of standing deadwood (m^{3}plot^{-1})	8049	77.78	81.25	0.91
Standing deadwood decay class	82.93	92.59	64.29	0.78
Number of seedlings (plot^{-1})	-	-	-	-
Number of native seedlings (plot^{-1})	-	-	-	-
Shannon–Wiener index for native seedlings	90.24	81.82	93.33	0.93 #
Number of ground vegetation species	80.49	96.15	53.33	0.82

The map presented in Figure 2 provides the index of favourable conservation status derived from ALS models for the New Forest study site. Generally, high index values are observed for areas containing native beech–oak deciduous forest, whereas low values are consistent with areas of coniferous plantation forest.

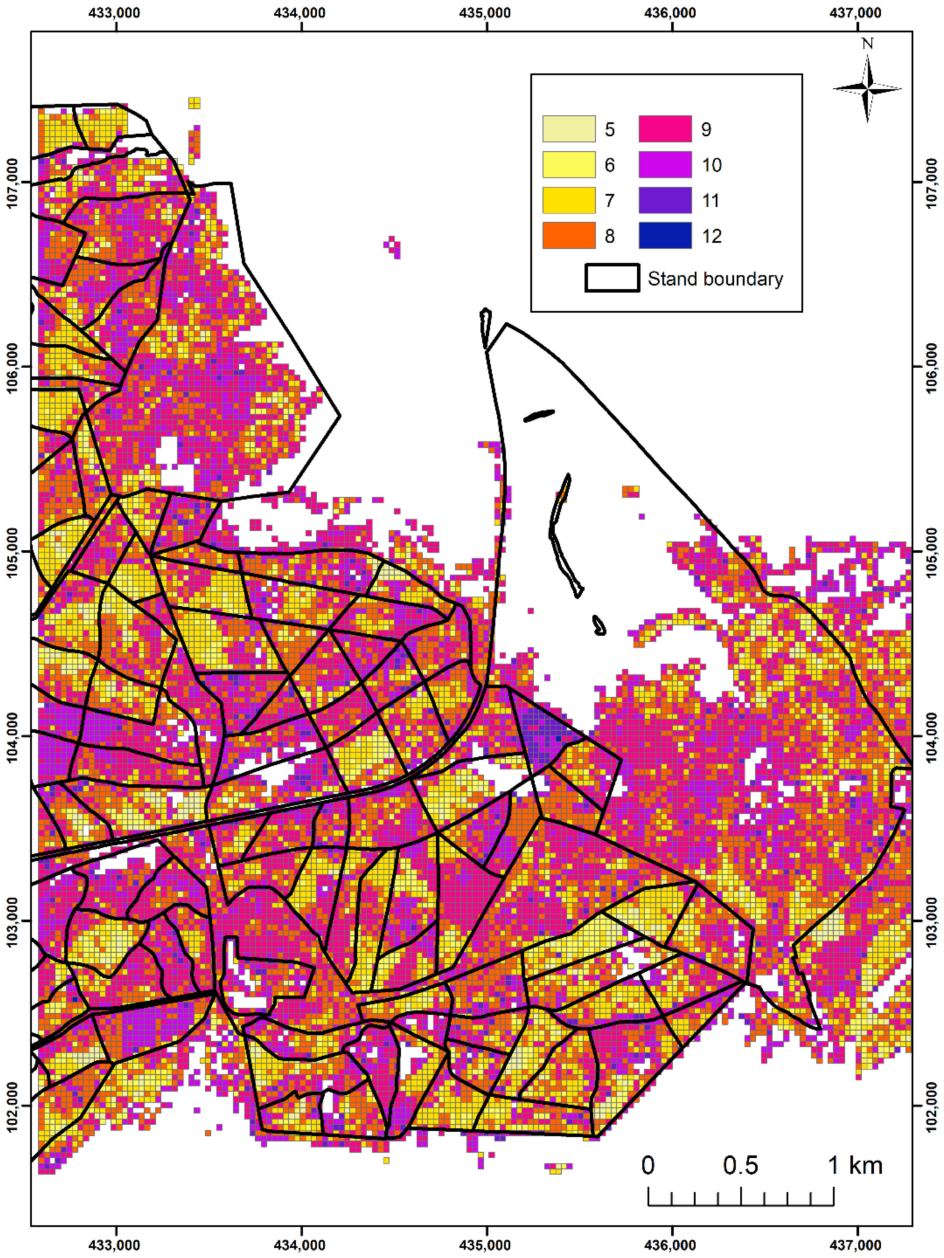

Figure 2. Predicted 'favourable conservation status' index map calculated for 30 × 30 m areas covering the entire study site extent. Index values ranged from 5 to 12, where lower values were observed for locations within managed coniferous stands, and higher values for cells located within the semi-ancient deciduous dominated stands.

4. Discussion

4.1. RF Model Differences between Leaf-On and Leaf-Off ALS Acquisitions

The timing of ALS acquisition in the current research, with leaf-off data (from April, prior to leaf flush) and leaf-on data (from July, in mid-summer) is generally expected to produce different ALS return distributions, specifically related to canopy penetration rates, (e.g., [67,107,108]). The analysis of both ALS acquisitions for field metric estimates by random forest modelling, when assessed individually, produced similar results in terms of NRMSE and NBias. With the exception of the Simpson index of diversity for native trees (which was modelled with a notably high NBias in leaf-on only data), both NRMSE and NBias varied by <±4.5% between leaf-on only and leaf-off only models (similarly to [109]). However, whilst these model differences were low, 11 of 27 metrics were more accurately

predicted from the leaf-off data, with most of these field metrics related to the number, spacing and variation of trees (in terms of species composition and size) or below-canopy features such as seedlings, deadwood and height to the live crown. Only seven leaf-on models were better than leaf-off models in terms of NRMSE and NBias, and these mostly related to tree size and crown area, saplings and bare ground.

When the two ALS acquisitions were combined into RF models, there was relatively little gain in accuracy per metric. In total, only in the case of Number of sapling species, Volume of standing deadwood and the Simpson index of diversity was there a difference of >1% point in both NRMSE and NBias when comparing leaf-on only, leaf-off only and/or combined leaf-on/leaf-off RF models. For the other 24 metrics, the difference between models in terms of NRMSE and NBias was negligible or inconsistent. However, whilst the models had similar accuracy, they differed considerably in their input ALS variables (see Section 4.2). When the best models were selected to minimise NRMSE and NBias, and maximise the adj. R^2, eight were leaf-on only models, 12 leaf-off only models and seven were combined leaf-on/leaf-off models. In the best models, metrics related to tree size, crown area, saplings or bare ground used leaf-on only data, whilst metrics related to tree number and spacing, tree, seedling and ground vegetation species composition and diversity and deadwood were dominated by leaf-off or combined leaf-on/leaf-off ALS variables.

Overall, our NRMSE results were similar to the accuracies reported in the surrounding literature (Table 7). The most direct comparison is from an earlier study using the same field and ALS data and an area-based linear multiple regression approach (using Akaike information criterion analysis) to estimate 23 metrics. With the exception of the estimates of the Volume of standing deadwood and Downed deadwood decay class, which had a lower RMSE in [55], all other RF model estimates have a lower RMSE in the current study, typically over 50% lower than their linear model derived counterpart. All metrics have a better or similar R^2 and/or RMSE than other published studies, with the exception of Mean tree height [110] and Shannon–Wiener index for native trees [61]. It should be noted, however, that Standing deadwood decay class was modelled in the current study with an NRMSE > 30%, which could be considered unreliably high.

Table 7. Statistics (root mean square error (RMSE), normalised RMSE (%), coefficient of determination (R^2) and standard error of the regressions (SER%)) from the literature for the same metrics used in the current study which were estimated using airborne laser scanning.

Variables	Overall Accuracy	References
Number of trees (plot^{-1})	R^2 0.67; RMSE 15.97 (22%) R^2 0.82; RMSE 33.25	[55] [111]
Mean DBH (cm)	R^2 0.80; RMSE 10.11 (35%) R^2 0.71–0.73; RMSE 3.5–3.6 R^2 0.68–0.95; RMSE 6.38–8.16 (14.75–23.18%)	[55] [112] [102]
Standard deviation of tree diameters (cm)	R^2 0.59; SER 30%	[46]
Basal area (m^{2}plot^{-1})	R^2 0.69; RMSE 0.91 (28%) R^2 0.65–0.81; RMSE 0.32–0.43 R^2 0.3–0.91; RMSE 0.09–0.18	[55] [112] [113]
Percentage of big trees (DBH > 40 cm)	-	-
Mean tree height (m)	R^2 0.98; RMSE 0.83 R^2 0.88; RMSE 0.09–1.78 RMSE 1.3 m R^2 0.93 to 0.96; RMSE 4.15 to 4.27	[110] [112] [114] [115]
Mean tree stem spacing (m)	R^2 0.91; RMSE 1.32 (33%)	[55]

Table 7. *Cont.*

Variables	Overall Accuracy	References
Mean height to the living crown (m)	R^2 0.88; RMSE 2.54 (24%) R^2 0.72–0.92; RMSE 1.54 to 1.62 R^2 0.88; RMSE 1.62 R^2 0.4; RMSE 1.25	[55] [116] [117] [118]
Mean crown horizontal area (m^2)	R^2 0.86; RMSE 24.39 (28%)	[55]
Total crown horizontal area (m^2)	R^2 0.69; RMSE 645.64 (45%)	[55]
Number of tree species	R^2 0.32; RMSE 1.66 (28%)	[55]
Number of native tree stems	R^2 0.61; RMSE 13.93 (27%)	[55]
Shannon–Wiener index for native trees	R^2 0.67; RMSE 0.58 (40%) R^2 0.83, RMSE 0.25	[55] [61]
Simpson index of diversity	R^2 0.55; RMSE 0.29 (22%)	[55]
Number of saplings ($plot^{-1}$)	R^2 0.97; RMSE 26.46 (25%) R^2 0.87; RMSE 22.31	[55] [119]
Number of native saplings ($plot^{-1}$)	R^2 0.97; RMSE 11.47 (30%)	[55]
Number of sapling species	R^2 0.70; RMSE 1.42 (48%)	[55]
Number of seedlings ($plot^{-1}$)	R^2 0.28; RMSE 331.96 (36%) R^2 0.32; RMSE 71.25	[55] [119]
Number of native seedlings ($plot^{-1}$)	R^2 0.68; RMSE 245.15 (34%)	[55]
Shannon–Wiener index for native seedlings	-	-
Number of seedling species	R^2 0.45; RMSE 3.10 (24%)	[55]
Volume of downed deadwood (m^3plot^{-1})	R^2 0.45; RMSE 2.49 (27%) R^2 0.61; RMSE 14.09 (51.6%) R^2 0.62; RMSE 0.22	[55] [120] [121]
Downed deadwood decay class	R^2 0.75; RMSE 0.60 (43%)	[55]
Volume of standing deadwood (m^3plot^{-1})	R^2 0.92; RMSE 1.36 (16%) R^2 0.02–0.62; RMSE 3.71–5.23 (56–90%) R^2 0.48; RMSE 14.74 (78.8%)	[55] [122] [120]
Standing deadwood decay class	R^2 0.59; RMSE 0.57 (29%)	[55]
Number of ground vegetation species	R^2 0.77; RMSE 1.70 (28%)	[55]
Percentage bare ground cover (%)	R^2 0.86; RMSE 39.82 (42%)	[55]

4.2. Exploring the Predictors Used in the Models

The RF model individual predictors were not consistent when estimating the same field metric using leaf-on or leaf-off data. The majority of RF models derived from leaf-on data (22 of 27) included predictor variables summarising the distribution of return heights, whereas a lower proportion of models derived from leaf-off data (14 of 27) incorporated height metrics. When considering the models produced from a combination of leaf-on and leaf-off acquisitions, the proportion of models which incorporated height predictors was more similar to leaf-off models (16 of 27). Width variables were present in 16–18 models and amplitude variables were present in 13–14 models across the leaf-on, leaf-off and combined leaf-on/leaf-off datasets. Given that the accuracy of the models was similar across the leaf-on, leaf-off and combined datasets (with the exception of the Simpson index of diversity), this implies that amplitude and width ALS variables from leaf-off data (often from a lower depth within the canopy) can provide equivalent or possibly better predictive ability than height variables from leaf-on data.

Using the best selected RF model (Table 5), of the 10 metrics defined as being related to forest canopy structure (see Table 1), five of these used leaf-on only predictors, three used leaf-off only predictors and two used a combination of leaf-on and leaf-off predictors. These RF models showed a dominance of return height and width predictors, both with 26 out of a total number of 65 ALS variables across all 10 models. The majority of height

predictor variables related to measures of variance (i.e., standard or absolute deviation, variance, skewness) or upper canopy height (i.e., ≥70th percentile, maximum or dominant height), with the exception of Basal area which had lower canopy height variables (i.e., ≤30th percentile and minimum non-ground height). The relationships between forest structure and height predictor ALS variables are widely documented (e.g., [115,123]. The RF models for both Standard deviation of tree diameters and Mean height to the live crown were composed of width metrics only (mostly relating to variance or mean). The authors of [124] defined echo-width as providing information on the range distribution of scattering surfaces within the laser footprint that contributes to a return echo and it is an indicator for surface roughness and ground slope. Authors such as [125] have demonstrated statistically significant relationships between roughness parameters and forest structure, which implies that echo-width predictors may be linked to the size and distribution of tree canopy components encountered in the current study. Many of the amplitude predictors were for the mid- to upper percentiles or measures relating to total or variance, which is similar to that observed in [126] whereby amplitude predictors combined with height predictors were strongly related to forest biomass. Other research has stated that the distribution of amplitude values in a forest is related to the presence and spatial arrangement of foliage [68,127]. The best RF model for the estimation of Total crown horizontal area had amplitude variables as the three most important predictors, highlighting a relationship between canopy biomass and ALS return amplitude.

Concerning the estimates of the four forest metrics related to forest canopy composition, the best RF models were dominated by height-related ALS variables (13 of a total of 22 input variables). Unlike the forest structure models, these tended to be more from the mid-profile (i.e., 40th–55th percentile) in addition to measures of variance, implying a link between structural diversity and overstorey composition. The model for the Simpson index of diversity only incorporated height metrics. Native species were typically deciduous species, characterised as being far more structurally complex (e.g., [128]), implying that the RF model is picking up on canopy height variation and understorey. The width and amplitude variables mostly related to ≥75th percentile or variance measures, indicating that species composition influences the spatial arrangement and roughness of the canopy.

The best RF models for regenerating saplings and seedlings were dominated by width input variables (17 out of a total of 32 variables used in the seven models), and these were spread across the full range (i.e., 5th–100th percentile) as well as measures relating to variance. The best RF models for the number of native saplings and seedlings were both created from only width variables. Overall, the number of regenerating saplings and seedlings encountered in the current study was low, and the majority of these were encountered in deciduous dominated plots. Leaf-off ALS data were thus the primary source of useful information (even with combined leaf-on/leaf-off models). The height variables used in the best RF models relate to mid-profile (i.e., 40th percentile) or to either maximum or variance. The best RF model for Number of seedlings contained only height predictors (related to top-canopy or variance). This highlights that seedling and sapling regeneration is linked to the surrounding forest structure. The Number of seedling species had amplitude as five of the seven input variables. Amplitude metrics are potentially related to living and dead biomass, bark and coniferous needles [122,129].

All RF models for the estimation of deadwood volume and decay class incorporated ALS metrics from both leaf-on and leaf-off acquisitions, and in particular amplitude and width variables (six and five, respectively, out of 15 total input ALS variables across the four models). These tend to relate to upper percentiles (≥70th percentile) and measures of variance. The authors of [130] suggested that forest ground-level vegetated elements are rougher whilst fallen stems have smooth surfaces that could be linked with echo-width. Standing deadwood volume was likely related to forest cover type and management, where larger proportions were located in deciduous dominated plots. As with downed deadwood, standing dead material likely has smoother surfaces, as represented in the width metrics used in the RF models. ALS amplitude variables used are possibly related to dominant

cover type and structure, being similar to the input variables used to predict Total crown horizontal area.

Finally, the best RF models for the ground layer metrics of Number of ground vegetation species and the Percentage of bare ground cover used very different ALS predictors. For the former model, the height predictor metrics were from the upper portion of the forest plot (≥80th percentile), implying that the upper canopy structure impacts ground vegetation species, whereas for the latter model the height predictors were from the lower portion of the profile (≤30th percentile), suggesting the importance of understorey in influencing ground vegetation coverage. In both cases, high percentile amplitude variables were also included, potentially related to the presence and spatial arrangement of foliage across the vertical profile.

4.3. Favourable Conservation Status

ALS datasets have been used to estimate forest structure and predict the habitat of various animal species, and potential diversity of said species (e.g., [63,131–133]). A common element is the identification of indicators of environmental conditions, and establishment of criteria for their assessment. For the current study, we assessed the accuracy of our RF model predictions with regard to 17 FCS indicator threshold values determined by [11] for semi-ancient deciduous forest in the UK. With the exception of assessments of seedling metrics (Number of seedlings, Number of native seedlings and Shannon–Wiener index for native seedlings), the 14 other metrics were predicted with acceptable accuracy with respect to their threshold values (i.e., with AUC values over 0.7, as specified in [106]). For all field plots, the numbers of seedlings encountered were consistently below the threshold value specified in [11].

The FCS index was calculated for the entire study site, and consistent values were observed within the extents of individual stand boundaries. Higher values were observed for locations containing semi-ancient deciduous forest, which we would expect given the nature of the index, and the natural structural and compositional diversity, and presence of native species, in comparison to the planted homogenous coniferous forest, which consistently scored low values. Conifer stands typically have little to no understorey due to management and shading. Conifer areas also have little or no deadwood, relatively low and invariant DBH values and low species diversity. This means that the grid-cells which intersect with these areas do not score very highly for anything except stem density and basal area. The number of regenerating trees (i.e., saplings and seedlings) is low across the entire study site, especially for coniferous areas, thus the highest value of 17 for the FCS index is not met for any of the New Forest study sites. This trend was recognised by [82] because of grazing by a high number of large ungulate species, such as deer and ponies, which are present throughout the unenclosed forest.

4.4. Concerns and Wider Implications

Sample size effects on model training and accuracy are the main issue in the current research project. None of our field plots achieved >11 in the FCS scoring, and so we must conclude that either more field data are needed to truly cover the complete range of conditions at the field site, or that this location simply does not have any sites with FCS index values > 11. The sample number was relatively low, although other studies have used similar sample numbers for RF analyses (e.g., [134]). The sample size precluded us from separating dedicated training and validation datasets. This therefore represents a source of uncertainty. Future work should ensure a large number of field plots are installed. RF models have been stated to be incapable of extrapolating beyond the range of values used in training the model [135].

In addition to sample size, plot location and FCS target thresholds should also be considered as possible causes of model outcome uncertainty. However, as in [136], minor imprecision in plot geo-reference has little impact on RF model performance for plot-level estimates. The FCS targets outlined in [11], or more specifically the target threshold values

(derived from: [80–83]), were collated for a specific semi-natural forest location. We have had to presume that these indicators are appropriate for all forest stands encountered in the current research and context. Data for semi-natural deciduous and plantation coniferous forest have been merged throughout this study.

There was a two-year time difference in this study between the two periods of field data collection, and therefore also between the 2012 field data and the 2010 ALS acquisitions. As our field plots were in mature, stable woodland, we must assume that any potential changes in our field site during this time were small. However, this is a potential source of uncertainty [137]. Future work would therefore need to ensure field data are acquired at an appropriate time relative to ALS acquisition to ensure any potential uncertainty is minimised.

Strong relationships were observed between ALS variables and fieldwork derived forest plot attributes. However, as with other similar studies, there are a number of potential issues which limit transferability of such models to other sites. These can relate to the range of forest types studied in order to derive the relationships, with many developed for a specific region and a limited number of species types. Tree species and canopy structure can influence the biases observed in ALS observations of height and can have an effect on the vertical penetration of ALS returns, which can be complicated when considering leaf-on and leaf-off data (e.g., [137,138]). In addition, considerations over differences in the acquisition parameters of the ALS data between studies pose problems [139,140], as do seasonal differences between field and remote sensing data capture, for example, models calibrated using leaf-on field and ALS data can produce erroneous results when applied to leaf-off ALS data (e.g., [64]). Different canopy penetration capabilities when comparing sensors have been demonstrated, even when applied at the same location [141,142]. Acquisition specific considerations in terms of sensor specifications, flying altitude, pulse repetition frequency, laser power, beam divergence and footprint size will influence target parameter characterisation and a system's ability to resolve a return [143–145]. The application of different processing methods (in particular, raster or point cloud approaches) can also yield differences in parameter retrieval [146].

Within the context of the current research, the ALS metrics selected were relatively basic with an aim of being more transferable [143,147]. The surrounding research literature contains additional metrics which could improve our models. The inclusion of vertically distributed canopy density or complexity indices may improve model accuracy as they have been demonstrated to correlate well with canopy conditions (e.g., [148,149]). Computationally similar light penetration indices appear to be transferable across acquisitions (e.g., [150]). Likewise, methods based on fitting functions to the frequency of returns by height, e.g., Weibull probability density functions [151], offer great potential. Future work will therefore need to assess the transferability or stability of metrics, in addition to the modelling methods, across different ALS acquisitions and environmental contexts.

One of the conclusions of the study presented in [55], when using linear regression analysis to estimate forest metrics, was that the leaf-on and leaf-off datasets captured different properties of the forest, which is supported here given the range of predictor metrics used across the models. Many of the best performing models in that study (10 of 23) included both leaf-on and leaf-off metrics. Overall, more accurate results were obtained in the current study, but relatively small differences were observed between leaf-on, leaf-off or combined model accuracies. This suggests that the RF approach may be better when applied to locally disparate forest types than linear models (as in [152]). There is also the potential implication that a combination of datasets of leaf-on and leaf-off acquisitions may not be necessary depending on the statistical modelling approach applied, at least within the context of the current study. The benefits that a combination of datasets provides in a regression approach may have been rendered redundant by the use of RF here. A comparison of different modelling methods would be a logical avenue for future work.

5. Conclusions

In this study undertaken in coniferous and deciduous forest in the New Forest, UK, we found that the differences between RF models derived from leaf-on, leaf-off and combined leaf-on/leaf-off datasets were slight, with the exception of estimating the Simpson index of diversity. Thus, whilst there were some detectable trends (e.g., higher accuracies were observed for forest overstorey structure when leaf-on ALS predictor variables were used, whilst forest canopy composition or understorey characteristics achieved a higher accuracy when either leaf-off or combined leaf-on and leaf-off ALS data were used), these were not significant enough to imply that consideration of ALS acquisition time will be required for optimal prediction of forest characteristics if using RF modelling. However, whilst model performance was similar between leaf-on, leaf-off and combined datasets, model composition was often very dissimilar, reiterating that these datasets capture different aspects of the forest and that structure and composition across the full vertical profile are highly inter-connected.

Estimated accuracy overall exceeded estimates produced using linear models in previous research. Whilst there is room for improvement, given the uncertainties associated with estimating a range of metrics and the small training sample size used, the ALS-based method was in good agreement with field-based assessments of FCS. In addition, continuously mapped estimates of FCS were created across the study site and corresponded closely to forest stands and compartments present, demonstrating the utility of such an approach for forest condition mapping and monitoring.

The value of ALS is its ability to estimate a variety of habitat variables related to forest three-dimensional structure. The current research demonstrates the feasibility of predicating spatially explicit forest metrics, derived from ALS covariates, for multiple forest stands. This approach potentially demonstrates a method to rapidly assess forest condition and FCS over large areas, reducing (but not eliminating) the need for costly field surveys. The main advantage of ALS acquisitions of this type is that they allow the sampling of forest structural characteristics at a high spatial resolution over large spatial extents. The ability to characterise these structures through ALS provides a proxy for biophysical processes in forests [126]. These findings are important for advancing the management of forest resources. Further investigation into such approaches could therefore yield additional data on many potential biological processes and distributions within a landscape. A similar type of approach to that documented in the current study would allow managers to investigate forest changes post-disturbance or treatment. Equipped with this knowledge, managers would have better information and have tools to address the impact of forest changes in the face of a potentially changing climate.

The results of this study show that full-waveform multi-temporal ALS holds a great deal of potential information which is useful for estimating forest characteristics, both within and below the main forest canopy, and for mapping continuously across large extents. The spatial resolution of the data allows for within stand assessment, whereas previously, areas defined as stands, regardless of size, were considered the smallest management unit. We expect that the use of remote sensing technologies and methods will be tested in other sites in the future in order to develop better estimates. The availability of high-resolution forest vegetation maps with acceptable agreement to field validation will significantly advance forest ecological understanding and improve conservation efforts.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs14205081/s1, Figure S1. Plots of predicted values from RF models against field measurements produced from the most accurate models; Table S1. A summary of field measured values for each of the 27 metrics across all 41 plot locations; Table S2. List of ecological indicators for the New Forest; Table S3. A summary of random forest model tuning parameters for models using metrics from the leaf-on acquisition; Table S4. A summary of variable inputs to each random forest model using metrics from the leaf-on acquisition, as determined by Boruta variable selection; Table S5. A summary of random forest model tuning parameters for models using metrics from the leaf-off acquisition; Table S6. A summary of variable inputs to each random forest model using

metrics from the leaf-off acquisition, as determined by Boruta variable selection; Table S7. A summary of random forest model tuning parameters for models using metrics from both the leaf-on and leaf-off acquisitions; Table S8. A summary of variable inputs to each random forest model using metrics from both leaf-on and leaf-off acquisitions, as determined by Boruta variable selection.

Author Contributions: Conceptualisation, R.A.H.; methodology, M.J.S.; software, M.J.S.; validation, M.J.S.; formal analysis, M.J.S.; investigation, M.J.S. and R.A.H.; resources, S.A.H. and R.A.H.; data curation, M.J.S. and R.A.H.; writing—original draft preparation, M.J.S. and R.A.H.; writing—review and editing, M.J.S., R.A.H. and S.A.H.; visualisation, M.J.S.; supervision, R.A.H. and S.A.H.; project administration, R.A.H. and M.J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: ALS data are available via the NERC Earth Observation Data Centre (NEODC) http://www.neodc.rl.ac.uk (accessed on 30 September 2022). Field plot data are available on request from the corresponding author.

Acknowledgments: We would like to acknowledge the Airborne Research and Survey Facility of the National Environment Research Council for their collaboration in the remote sensing survey and post-processing of the ALS data, in addition to the assistance of the Forestry Commission for providing site access and supplementary site and management information. Bournemouth University and the Centre for Ecology and Hydrology jointly funded this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bayraktarov, E.; Ehmke, G.; O'Connor, J.; Burns, E.L.; Nguyen, H.A.; McRae, L.; Possingham, H.P.; Lindenmayer, D.B. Do big unstructured biodiversity data mean more knowledge? *Front. Ecol. Environ.* **2019**, *6*, 239. [CrossRef]
2. Bibby, C.J.; Collar, N.J.; Crosby, M.J.; Heath, M.F.; Imboden, C.; Johnson, T.H.; Long, A.J.; Stattersfield, A.J.; Thirgood, S.J. *Putting Biodiversity on the Map: Priority Areas for Global Conservation*; International Council for Bird Preservation: Cambridge, UK, 1992.
3. Allen, H.D. *Mediterranean Ecogeography*; Prentice Hall: Harlow, UK, 2001.
4. Allen, H.D. Vegetation and ecosystem dynamics. In *The Physical Geography of the Mediterranean*; Woodward, J., Ed.; Oxford University Press: Oxford, UK, 2009; pp. 203–227.
5. Regato, P. *Adapting to Global Change: Mediterranean Forests*; World Conservation Union: Malaga, Spain, 2008.
6. Horner, G.J.; Baker, P.J.; Mac Nally, R.; Cunningham, S.C.; Thomson, J.R.; Hamilton, F. Forest structure, habitat and carbon benefits from thinning floodplain forests: Managing early stand density makes a difference. *For. Ecol. Manag.* **2010**, *259*, 286–293. [CrossRef]
7. Noss, R.F. Assessing and monitoring forest biodiversity: A suggested framework and indicators. *For. Ecol. Manag.* **1999**, *115*, 135–146. [CrossRef]
8. Vellend, M.; Verheyen, K.; Flinn, K.M.; Jacquemyn, H.; Kolb, A.; VAN Calster, H.; Peterken, G.; Graae, B.J.; Bellemare, J.; Honnay, O.; et al. Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *J. Ecol.* **2007**, *95*, 565–573. [CrossRef]
9. McElhinny, C.; Gibbons, P.; Brack, C.; Bauhus, J. Forest and woodland stand structural complexity: Its definition and measurement. *For. Ecol. Manag.* **2005**, *218*, 1–24. [CrossRef]
10. Borrass, L. Varying practices of implementing the Habitats Directive in German and British forests. *For. Policy Econ.* **2014**, *38*, 151–160. [CrossRef]
11. Cantarello, E.; Newton, A.C. Identifying cost-effective indicators to assess the conservation status of forested habitats in Natura 2000 sites. *For. Ecol. Manag.* **2008**, *256*, 815–826. [CrossRef]
12. Baccini, A.; Friedl, M.; Woodcock, C.; Zhu, Z. Scaling Field Data to Calibrate and Validate Moderate Spatial Resolution Remote Sensing Models. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 945–954. [CrossRef]
13. Wang, Y.; Weinacker, H.; Koch, B. A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis and 3D Single Tree Modelling in Forest. *Sensors* **2008**, *8*, 3938–3951. [CrossRef] [PubMed]
14. Bässler, C.; Stadler, J.; Müller, J.; Förster, B.; Göttlein, A.; Brandl, R. LiDAR as a rapid tool to predict forest habitat types in Natura 2000 networks. *Biodivers. Conserv.* **2010**, *20*, 465–481. [CrossRef]
15. Simonson, W.D.; Allen, H.D.; Coomes, D.A. Remotely sensed indicators of forest conservation status: Case study from a Natura 2000 site in southern Portugal. *Ecol. Indic.* **2013**, *24*, 636–647. [CrossRef]
16. Ichter, J.; Evans, D.; Richard, D. Terrestrial habitat mapping in Europe: An overview. In *EEA Technical Report No 1/2014*; European Environment Agency: Copenhagen, Denmark, 2014.

17. Zlinszky, A.; Schroiff, A.; Kania, A.; Deák, B.; Mücke, W.; Vári, Á.; Székely, B.; Pfeifer, N. Categorizing grassland vegetation with full-waveform airborne laser scanning: A feasibility study for detecting Natura 2000 habitat types. *Remote Sens.* **2014**, *6*, 8056–8087. [CrossRef]
18. Corbane, C.; Lang, S.; Pipkins, K.; Alleaume, S.; Deshayes, M.; Millán, V.E.G.; Strasser, T.; Borre, J.V.; Toon, S.; Michael, F. Remote sensing for mapping natural habitats and their conservation status–New opportunities and challenges. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *37*, 7–16. [CrossRef]
19. Nagendra, H.; Lucas, R.; Honrado, J.P.; Jongman, R.H.; Tarantino, C.; Adamo, M.; Mairota, P. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Indic.* **2012**, *33*, 45–59. [CrossRef]
20. Smith, A.M.; Kolden, C.A.; Tinkham, W.T.; Talhelm, A.F.; Marshall, J.D.; Hudak, A.T.; Boschetti, L.; Falkowski, M.J.; Greenberg, J.A.; Anderson, J.W.; et al. Remote sensing the vulnerability of vegetation in natural terrestrial ecosystems. *Remote Sens. Environ.* **2014**, *154*, 322–337. [CrossRef]
21. Lawley, V.; Lewis, M.; Clarke, K.; Ostendorf, B. Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. *Ecol. Indic.* **2016**, *60*, 1273–1283. [CrossRef]
22. Tehrany, M.S.; Kumar, L.; Drielsma, M.J. Review of native vegetation condition assessment concepts, methods and future trends. *J. Nat. Conserv.* **2017**, *40*, 12–23. [CrossRef]
23. Morales, R.M.; Miura, T.; Idol, T. An assessment of Hawaiian dry forest condition with fine resolution remote sensing. *For. Ecol. Manag.* **2008**, *255*, 2524–2532. [CrossRef]
24. Cunningham, S.C.; Griffioen, P.; White, M.D.; Mac Nally, R. Assessment of ecosystems: A system for rigorous and rapid mapping of floodplain forest condition for Australia's most important river. *Land Degrad. Dev.* **2017**, *29*, 127–137. [CrossRef]
25. Simonson, W.D.; Allen, H.D.; Coomes, D.A. Use of an Airborne Lidar System to Model Plant Species Composition and Diversity of Mediterranean Oak Forests. *Conserv. Biol.* **2012**, *26*, 840–850. [CrossRef]
26. Chraibi, E.; Arnold, H.; Luque, S.; Deacon, A.; Magurran, A.; Féret, J.-B. A Remote Sensing Approach to Understanding Patterns of Secondary Succession in Tropical Forest. *Remote Sens.* **2021**, *13*, 2148. [CrossRef]
27. Chen, L.; Ren, C.; Zhang, B.; Wang, Z.; Wang, Y. Mapping Spatial Variations of Structure and Function Parameters for Forest Condition Assessment of the Changbai Mountain National Nature Reserve. *Remote Sens.* **2019**, *11*, 3004. [CrossRef]
28. Sampson, P.H.; Mohammed, G.H.; Noland, T.L.; Irving, D.; Colombo, S.J.; Zarco-Tejada, P.J.; Miller, J.R.; Freemantle, J.; Treitz, P.M. The Bioindicators of Forest Condition Project: A physiological, remote sensing approach. *For. Chron.* **2000**, *76*, 941–952. [CrossRef]
29. Sampson, P.H.; Zarco-Tejada, P.J.; Mohammed, G.H.; Miller, J.R.; Noland, T.L. Hyperspectral remote sensing of forest condi-tion: Estimating chlorophyll content in tolerant hardwoods. *For. Sci.* **2003**, *49*, 381–391.
30. Sampson, P.; Treitz, P.; Mohammed, G. Remote Sensing of Forest Condition in Tolerant Hardwoods: An Examination of Spatial Scale, Structure and Function. *Can. J. Remote Sens.* **2001**, *27*, 232–246. [CrossRef]
31. Evans, B.J.; Lyons, T.; Barber, P.; Stone, C.; Hardy, G. Enhancing a eucalypt crown condition indicator driven by high spatial and spectral resolution remote sensing imagery. *J. Appl. Remote Sens.* **2012**, *6*, 063605. [CrossRef]
32. Ochtyra, A.; Zagajewski, B.; Kozlowska, A.; Marcinkowska-Ochtyra, A.; Jarocinska, A. Assessment of the condition of forests in the Tatra National Park using decision tree method and multispectral Landsat TM satellite images. *Sylwan* **2016**, *160*, 256–264.
33. Michez, A.; Piégay, H.; Lisein, J.; Claessens, H.; Lejeune, P. Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environ. Monit. Assess.* **2016**, *188*, 146. [CrossRef]
34. Pasquarella, V.J.; Mickley, J.G.; Plotkin, A.B.; MacLean, R.G.; Anderson, R.M.; Brown, L.M.; Wagner, D.L.; Singer, M.S.; Bagchi, R. Predicting defoliator abundance and defoliation measurements using Landsat-based condition scores. *Remote Sens. Ecol. Conserv.* **2021**, *7*, 592–609. [CrossRef]
35. Buras, A.; Rammig, A.; Zang, C.S. The European Forest Condition Monitor: Using Remotely Sensed Forest Greenness to Identify Hot Spots of Forest Decline. *Front. Plant Sci.* **2021**, *12*, 689220. [CrossRef]
36. Kovalev, A.V.; Voronin, V.I.; Oskolkov, V.A.; Sukhovolskiy, V.G. Analysis of forest condition based on MODIS remote-sensing data. *Contemp. Probl. Ecol.* **2021**, *14*, 717–722. [CrossRef]
37. Shapiro, A.C.; Grantham, H.S.; Aguilar-Amuchastegui, N.; Murray, N.J.; Gond, V.; Bonfils, D.; Rickenbach, O. Forest condition in the Congo Basin for the assessment of ecosystem conservation status. *Ecol. Indic.* **2020**, *122*, 107268. [CrossRef]
38. Coops, N.C.; Shang, C.; Wulder, M.A.; White, J.C.; Hermosilla, T. Change in forest condition: Characterizing non-stand re-placing disturbances using time series satellite imagery. *For. Ecol. Manag.* **2020**, *474*, 118370. [CrossRef]
39. Lentile, L.B.; Smith, A.M.S.; Hudak, A.T.; Morgan, P.; Bobbitt, M.J.; Lewis, S.A.; Robichaud, P.R. Remote sensing for prediction of 1-year post-fire ecosystem condition. *Int. J. Wildland Fire* **2009**, *18*, 594–608. [CrossRef]
40. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. LiDAR remote sensing for ecosystem studies. *Bioscience* **2002**, *52*, 19–30. [CrossRef]
41. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2001**, *80*, 88–99. [CrossRef]
42. Lim, K.; Treitz, P.; Baldwin, K.; Morrison, I.; Green, J. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Can. J. Remote Sens.* **2003**, *29*, 658–678. [CrossRef]
43. Maltamo, M.; Packalén, P.; Yu, X.; Eerikäinen, K.; Hyyppä, J.; Pitkänen, J. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *For. Ecol. Manag.* **2005**, *216*, 41–50. [CrossRef]

44. Anderson, J.E.; Plourde, L.C.; Martin, M.E.; Braswell, B.; Smith, M.-L.; Dubayah, R.O.; Hofton, M.A.; Blair, J.B. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sens. Environ.* **2008**, *112*, 1856–1870. [CrossRef]
45. Lefsky, M.A.; Harding, D.J.; Keller, M.; Cohen, W.B.; Carabajal, C.C.; Del Bom Espirito-Santo, F.; Hunter, M.O.; de Oliveira, R., Jr. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.* **2005**, *32*, L22S02. [CrossRef]
46. Lefsky, M.A.; Cohen, W.B.; Spies, T.A. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. For. Res.* **2001**, *31*, 78–87. [CrossRef]
47. Evans, J.S.; Hudak, A.T.; Faux, R.; Smith, A.M.S. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sens.* **2009**, *1*, 776–794. [CrossRef]
48. Anderson, E.S.; Thompson, J.A.; Austin, R. LIDAR density and linear interpolator effects on elevation estimates. *Int. J. Remote Sens.* **2005**, *26*, 3889–3900. [CrossRef]
49. Mascaro, J.; Detto, M.; Asner, G.P.; Muller-Landau, H.C. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sens. Environ.* **2011**, *115*, 3770–3774. [CrossRef]
50. Plowright, A.A.; Coops, N.C.; Eskelson, B.N.; Sheppard, S.R.; Aven, N.W. Assessing urban tree condition using airborne light detection and ranging. *Urban For. Urban Green.* **2016**, *19*, 140–150. [CrossRef]
51. Goetz, S.; Steinberg, D.; Dubayah, R.; Blair, B. Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sens. Environ.* **2007**, *108*, 254–263. [CrossRef]
52. Muller, J.; Brandl, R. Assessing biodiversity by remote sensing in mountainous terrain: The potential of ALS to predict forest beetle assemblages. *J. Appl. Ecol.* **2009**, *46*, 897–905. [CrossRef]
53. Vierling, K.T.; Bassler, C.; Brandl, R.; Vierling, L.A.; Weib, I.; Muller, J. Spinning a laser web: Predicting spider distributions using ALS. *Ecol. Appl.* **2011**, *21*, 577–588. [CrossRef] [PubMed]
54. Leutner, B.F.; Reineking, B.; Müller, J.; Bachmann, M.; Beierkuhnlein, C.; Dech, S.; Wegmann, M. Modelling forest α-diversity and floristic composition—On the added value of ALS plus hyperspectral remote sensing. *Remote Sens.* **2012**, *4*, 2818–2845. [CrossRef]
55. Sumnall, M.J.; Hill, R.A.; Hinsley, S.A. Comparison of small-footprint discrete return and full waveform airborne lidar data for estimating multiple forest variables. *Remote Sens. Environ.* **2016**, *173*, 214–223. [CrossRef]
56. Korpela, I.; Ørka, H.O.; Hyyppä, J.; Heikkinen, V.; Tokola, T. Range and AGC normalization in airborne discrete-return ALS intensity data for forest canopies. *ISPRS J. Photogramm.* **2010**, *65*, 369–379. [CrossRef]
57. Morsdorf, F.; Mårell, A.; Koetz, B.; Cassagne, N.; Pimont, F.; Rigolot, E.; Allgöwer, B. Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. *Remote Sens. Environ.* **2010**, *114*, 1403–1415. [CrossRef]
58. Korpela, I.; Hovi, A.; Morsdorf, F. Understory trees in airborne LiDAR data—Selective mapping due to transmission losses and echo-triggering mechanisms. *Remote Sens. Environ.* **2012**, *119*, 92–104. [CrossRef]
59. Hopkinson, C.; Chasmer, L.; Gynan, C.; Mahoney, C.; Sitar, M. Multisensor and Multispectral LiDAR Characterization and Classification of a Forest Environment. *Can. J. Remote Sens.* **2016**, *42*, 501–520. [CrossRef]
60. Zhang, C.; Qiu, F. Mapping individual tree species in an urban forest using airborne ALS data and hyperspectral imagery. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1079–1087. [CrossRef]
61. Zhao, Y.; Zeng, Y.; Zheng, Z.; Dong, W.; Zhao, D.; Wu, B.; Zhao, Q. Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. *Remote Sens. Environ.* **2018**, *213*, 104–114. [CrossRef]
62. Antonarakis, A.S.; Richards, K.S.; Brasington, J. Object-based land cover classification using airborne ALS. *Remote Sens. Environ.* **2008**, *112*, 2988–2998. [CrossRef]
63. Simonson, W.D.; Allen, H.D.; Coomes, D.A. Applications of airborne lidar for the assessment of animal species diversity. *Methods Ecol. Evol.* **2014**, *5*, 719–729. [CrossRef]
64. Tew, E.R.; Conway, G.J.; Henderson, I.G.; Milodowski, D.T.; Swinfield, T.; Sutherland, W.J. Recommendations to enhance breeding bird diversity in managed plantation forests determined using LiDAR. *Ecol. Appl.* **2022**, *32*, e2678. [CrossRef] [PubMed]
65. Villikka, M.; Packalén, P.; Maltamo, M. The suitability of leaf-off airborne laser scanning data in an area-based forest inventory of coniferous and deciduous trees. *Silva Fenn.* **2012**, *46*, 99–110. [CrossRef]
66. Hill, R.; Broughton, R. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 223–233. [CrossRef]
67. Wasser, L.; Day, R.; Chasmer, L.; Taylor, A. Influence of Vegetation Structure on Lidar-derived Canopy Height and Fractional Cover in Forested Riparian Buffers During Leaf-Off and Leaf-On Conditions. *PLoS ONE* **2013**, *8*, e54776. [CrossRef] [PubMed]
68. Kim, S.; McGaughey, R.J.; Andersen, H.-E.; Schreuder, G. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sens. Environ.* **2009**, *113*, 1575–1586. [CrossRef]
69. Anderson, R.S.; Bolstad, P.V. Estimating aboveground biomass and average annual wood biomass increment with airborne leaf-on and leaf-off ALS in Great Lakes forest types. *North J. Appl. For.* **2013**, *30*, 16–22. [CrossRef]
70. Davison, S.; Donoghue, D.N.; Galiatsatos, N. The effect of leaf-on and leaf-off forest canopy conditions on LiDAR derived estimations of forest structural diversity. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102160. [CrossRef]
71. Brovkina, O.; Navrátilová, B.; Novotný, J.; Albert, J.; Slezák, L.; Cienciala, E. Influences of vegetation, model, and data pa-rameters on forest aboveground biomass assessment using an area-based approach. *Ecol. Inform.* **2022**, *70*, 101754. [CrossRef]

72. Alexander, C.; Tansey, K.; Kaduk, J.; Holland, D.; Tate, N.J. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 423–432. [CrossRef]
73. Mallet, C.; Bretar, F. Full-waveform topographic lidar: State-of-the-art. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 1–16. [CrossRef]
74. Wagner, W.; Ullrich, A.; Melzer, T.; Briese, C.; Kraus, K. From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges. *Int. Arch. Photogramm.* **2004**, *35*, 201–206.
75. Chauve, A.; Vega, C.; Durrieu, S.; Bretar, F.; Allouis, T.; Pierrot Deseilligny, M.; Puech, W. Advanced full-waveform ALS data echo detection: Assessing quality of derived terrain and tree height models in an alpine coniferous forest. *Int. J. Remote Sens.* **2009**, *30*, 5211–5228. [CrossRef]
76. Moskal, L.M.; Erdody, T.; Kato, A.; Richardson, J.; Zheng, G.; Briggs, D. Lidar applications in precision forestry. In Proceedings of the Silvilaser 2009, College Station, TX, USA, 14–16 October 2009; Texas A&M University: College Station, TX, USA, 2009.
77. Reitberger, J.; Krzystek, P.; Stilla, U. Combined tree segmentation and stem detection using full waveform ALS data. *Int. Arch. Photogramm.* **2007**, *36*, 332–337.
78. Tubbs, C.R. *The New Forest: History, Ecology, and Conservation*; New Forest Ninth Centenary Trust: Lyndhurst, Hampshire, 2001.
79. Newton, A.; Cantarello, E.; Myers, G.; Douglas, S.J.; Tejedor, N. The condition and dynamics of New Forest woodlands. In *Biodiversity in the New Forest*; Newton, A.C., Ed.; Pisces Publications: Newbury, UK, 2010; pp. 132–148.
80. Ferretti, M.; Petriccione, B.; Fabbio, G.; Bussotti, F. (Eds.) Aspects of biodiversity in selected forest ecosystems in Italy: Status and changes over the period 1996–2003. Third report of the Task Force on Integrated and Combined (I & C) evaluation of the CONECOFOR programme. *Annali CRA-Ist. Sper. Selv. Arezzo.* **2006**, *30* (Suppl. S2), 112.
81. Keddy, P.A.; Drummond, C.G. Ecological Properties for the Evaluation, Management, and Restoration of Temperate Deciduous Forest Ecosystems. *Ecol. Appl.* **1996**, *6*, 748–762. [CrossRef]
82. Mountford, E.P.; Peterken, G.F.; Edwards, P.J.; Manners, J.G. Long-term change in growth, mortality and regeneration of trees in Denny Wood, an old-growth wood-pasture in the New Forest (UK). *Perspect. Plant Ecol. Evol. Syst.* **1999**, *2*, 223–272. [CrossRef]
83. Van den Meersschaut, D.; Vandekerkhove, K. Development of a stand scale forest biodiversity index based on the state forest inventory. In *Integrated Tools for Natural Resources Inventories in the 21st Century: Proceedings Held at Boise Centre on the Grove, Boise, ID, USA, 16–20 August 1998*; Hansen, M., Burk, T., Eds.; General Technical Report nc-212; U.S. Department of Agriculture, Forest Service, North Central Research Station: Washington, DC, USA, 1998.
84. Shannon, C.E. *The Mathematical Theory of Communication*; Shannon, C.E., Weaver, W., Eds.; University of Illinois Press: Urbana, IL, USA, 1948; pp. 29–125.
85. Simpson, E.H. Measurement of diversity. *Nature* **1949**, *163*, 688. [CrossRef]
86. Spies, T.A.; Franklin, J.F.; Thomas, T.B. Coarse Woody Debris in Douglas-Fir Forests of Western Oregon and Washington. *Ecology* **1988**, *69*, 1689–1702. [CrossRef]
87. Bunting, P.; Armston, J.; Lucas, R.M.; Clewley, D. Sorted pulse data (SPD) library. Part I: A generic file format for LiDAR data from pulsed laser systems in terrestrial environments. *Comput. Geosci.* **2013**, *56*, 197–206. [CrossRef]
88. Bunting, P.; Armston, J.D.; Clewley, D.; Lucas, R.M. Sorted pulse data (SPD) library—Part II: A processing framework for LiDAR data from pulsed laser systems in terrestrial environments. *Comput. Geosci.* **2013**, *56*, 207–215. [CrossRef]
89. Zhang, K.Q.; Chen, S.C.; Whitman, D.; Shyu, M.L.; Yan, J.H.; Zhang, C.C. A progressive morphological filter for removing non-ground measurements from airborne LiDAR data. *IEEE T. Geosci. Remote* **2003**, *41*, 872–882. [CrossRef]
90. Gatziolis, D. Dynamic Range-based Intensity Normalization for Airborne, Discrete Return Lidar Data of Forest Canopies. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 251–259. [CrossRef]
91. Tesfamichael, S.; Van Aardt, J.; Ahmed, F. Estimating plot-level tree height and volume of Eucalyptus grandis plantations using small-footprint, discrete return lidar data. *Prog. Phys. Geogr. Earth Environ.* **2010**, *34*, 515–540. [CrossRef]
92. Næsset, E.; Bollandsås, O.M.; Gobakken, T. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sens. Environ.* **2005**, *94*, 541–553. [CrossRef]
93. Laurin, G.V.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Del Frate, F.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with ALS and hyperspectral data. *ISPRS J. Photogramm.* **2014**, *89*, 49–58. [CrossRef]
94. Chen, G.; Hay, G.J. A Support Vector Regression Approach to Estimate Forest Biophysical Parameters at the Object Level Using Airborne Lidar Transects and QuickBird Data. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 733–741. [CrossRef]
95. Kursa, M.B.; Rudnicki, W.R. Feature Selection with theBorutaPackage. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
96. Kuhn, M. caret: Classification and Regression Training. R Package Version 6.0-90. 2021. Available online: https://CRAN.R-project.org/package=caret (accessed on 17 April 2019).
97. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
98. Jackman, S. *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory*; United States Studies Centre; University of Sydney: Sydney, NSW, Australia, R Package Version 1.5.2; 2017. Available online: https://github.com/atahk/pscl/ (accessed on 15 September 2021).
99. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. [CrossRef]
100. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

101. Yu, X.; Hyyppä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37. [CrossRef]
102. Wu, J.; Yao, W.; Choi, S.; Park, T.; Myneni, R.B. A Comparative Study of Predicting DBH and Stem Volume of Individual Trees in a Temperate Forest Using Airborne Waveform LiDAR. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2267–2271. [CrossRef]
103. Silva, C.A.; Klauberg, C.; Hudak, A.T.; Vierling, L.A.; Jaafar, W.S.W.M.; Mohan, M.; Garcia, M.; Ferraz, A.; Cardil, A.; Saatchi, S. Predicting Stem Total and Assortment Volumes in an Industrial Pinus taeda L. Forest Plantation Using Airborne Laser Scanning Data and Random Forest. *Forests* **2017**, *8*, 254. [CrossRef]
104. Timsina, J.; Humphreys, E. Performance of CERES-Rice and CERES-Wheat models in rice–wheat systems: A review. *Agric. Syst.* **2006**, *90*, 5–31. [CrossRef]
105. Zhou, L.; Gu, X.; Cheng, S.; Yang, G.; Shu, M.; Sun, Q. Analysis of plant height changes of lodged maize using UAV-LiDAR data. *Agriculture* **2020**, *10*, 146. [CrossRef]
106. Swets, J.A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285–1293. [CrossRef] [PubMed]
107. Hyyppä, J.; Hyyppä, H.; Litkey, P.; Yu, X.; Haggrén, H.; Rönnholm, P.; Pyysalo, U.; Pitkänen, J.; Maltamo, M. Algorithms and methods of airborne laser scanning for forest measurements. *Int. Arch. Photogramm.* **2004**, *36*, 82–89.
108. Brandtberg, T. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS J. Photogramm. Remote Sens.* **2007**, *61*, 325–340. [CrossRef]
109. White, J.C.; Arnett, J.T.; Wulder, M.A.; Tompalski, P.; Coops, N.C. Evaluating the impact of leaf-on and leaf-off airborne laser scanning data on the estimation of forest inventory attributes with the area-based approach. *Can. J. For. Res.* **2015**, *45*, 1498–1513. [CrossRef]
110. Donoghue, D.N.M.; Watt, P.J. Using LiDAR to compare forest height estimates from IKONOS and Landsat ETM+ data in Sitka spruce plantation forests. *Int. J. Remote Sens.* **2006**, *27*, 2161–2175. [CrossRef]
111. Lee, A.C.; Lucas, R.M. A LiDAR-derived canopy density model for tree stem and crown mapping in Australian forests. *Remote Sens. Environ.* **2007**, *111*, 493–518. [CrossRef]
112. Hawbaker, T.J.; Keuler, N.S.; Lesak, A.A.; Gobakken, T.; Contrucci, K.; Radeloff, V.C. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *J. Geophys. Res.Biogeo.* **2009**, *114*, G00E04. [CrossRef]
113. Gobakken, T.; Næsset, E. Weibull and percentile models for lidar-based estimation of basal area distribution. *Scand. J. For. Res.* **2005**, *20*, 490–502. [CrossRef]
114. Vastaranta, M.; Saarinen, N.; Kankare, V.; Holopainen, M.; Kaartinen, H.; Hyyppä, J.; Hyyppä, H. Multisource Single-Tree Inventory in the Prediction of Tree Quality Variables and Logging Recoveries. *Remote Sens.* **2014**, *6*, 3475–3491. [CrossRef]
115. Jakubowski, M.K.; Li, W.; Guo, Q.; Kelly, M. Delineating Individual Trees from Lidar Data: A Comparison of Vector- and Raster-based Segmentation Approaches. *Remote Sens.* **2013**, *5*, 4163–4186. [CrossRef]
116. Kato, A.; Moskal, L.M.; Schiess, P.; Swanson, M.E.; Calhoun, D.; Stuetzle, W. Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. *Remote Sens. Environ.* **2016**, *113*, 1148–1162. [CrossRef]
117. Luo, L.; Zhai, Q.; Su, Y.; Ma, Q.; Kelly, M.; Guo, Q. Simple method for direct crown base height estimation of individual conifer trees using airborne LiDAR data. *Opt. Express* **2018**, *26*, A562–A578. [CrossRef]
118. Chamberlain, C.P.; Meador, A.J.S.; Thode, A.E. Airborne lidar provides reliable estimates of canopy base height and canopy bulk density in southwestern ponderosa pine forests. *For. Ecol. Manag.* **2020**, *481*, 118695. [CrossRef]
119. Sherrill, K.R.; Lefsky, M.A.; Bradford, J.B.; Ryan, M. Forest structure estimation and pattern exploration from discrete-return lidar in subalpine forests of the central Rockies. *Can. J. For. Res.* **2008**, *38*, 2081–2096. [CrossRef]
120. Pesonen, A.; Maltamo, M.; Eerikäinen, K.; Packalèn, P. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *For. Ecol. Manag.* **2008**, *255*, 3288–3296. [CrossRef]
121. Queiroz, G.L.; McDermid, G.J.; Linke, J.; Hopkinson, C.; Kariyeva, J. Estimating Coarse Woody Debris Volume Using Image Analysis and Multispectral LiDAR. *Forests* **2020**, *11*, 141. [CrossRef]
122. Kim, Y.; Yang, Z.; Cohen, W.B.; Pflugmacher, D.; Lauver, C.L.; Vankat, J.L. Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sens. Environ.* **2009**, *113*, 2499–2510. [CrossRef]
123. Falkowski, M.J.; Evans, J.S.; Martinuzzi, S.; Gessler, P.E.; Hudak, A.T. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sens. Environ.* **2009**, *113*, 946–956. [CrossRef]
124. Höfle, B.; Hollaus, M.; Hagenauer, J. Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 134–147. [CrossRef]
125. Maurer, K.D.; Hardiman, B.; Vogel, C.S.; Bohrer, G. Canopy-structure effects on surface roughness parameters: Observations in a Great Lakes mixed-deciduous forest. *Agric. For. Meteorol.* **2013**, *177*, 24–34. [CrossRef]
126. Li, M.; Im, J.; Quackenbush, L.J.; Liu, T. Forest Biomass and Carbon Stock Quantification Using Airborne LiDAR Data: A Case Study Over Huntington Wildlife Forest in the Adirondack Park. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3143–3156. [CrossRef]
127. Moffiet, T.; Mengersen, K.; Witte, C.; King, R.; Denham, R. Airborne laser scanning: Exploratory data analysis indicates po-tential variables for classification of individual trees or forest stands according to species. *ISPRS J. Photogramm.* **2005**, *59*, 289–309. [CrossRef]

128. Kaartinen, H.; Hyyppä, J.; Yu, X.; Vastaranta, M.; Hyyppä, H.; Kukko, A.; Holopainen, M.; Heipke, C.; Hirschmugl, M.; Morsdorf, F.; et al. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sens.* **2012**, *4*, 950–974. [CrossRef]
129. Reitberger, J.; Krzystek, P.; Stilla, U. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *Int. J. Remote Sens.* **2008**, *29*, 1407–1431. [CrossRef]
130. Mücke, W.; Deák, B.; Schroiff, A.; Hollaus, M.; Pfeifer, N. Detection of fallen trees in forested areas using small footprint airborne laser scanning data. *Can. J. Remote Sens.* **2013**, *39* (Suppl. S1), S32–S40. [CrossRef]
131. Garabedian, J.E.; Moorman, C.E.; Peterson, M.N.; Kilgo, J.C. Use of LiDAR to define habitat thresholds for forest bird con-servation. *For. Ecol. Manag.* **2017**, *399*, 24–36. [CrossRef]
132. Bombi, P.; Gnetti, V.; D'Andrea, E.; De Cinti, B.; Taglianti, A.V.; Bologna, M.A.; Matteucci, G. Identifying priority sites for insect conservation in forest ecosystems at high resolution: The potential of LiDAR data. *J. Insect Conserv.* **2019**, *23*, 689–698. [CrossRef]
133. Shanley, C.S.; Eacker, D.R.; Reynolds, C.P.; Bennetsen, B.M.; Gilbert, S.L. Using LiDAR and Random Forest to improve deer habitat models in a managed forest landscape. *For. Ecol. Manag.* **2021**, *499*, 119580. [CrossRef]
134. Torre-Tojal, L.; Bastarrika, A.; Boyano, A.; Lopez-Guede, J.M.; Graña, M. Above-ground biomass estimation from LiDAR data using random forest algorithms. *J. Comput. Sci.* **2021**, *58*, 101517. [CrossRef]
135. Zhang, H.; Nettleton, D.; Zhu, Z. Regression-enhanced random forests. *arXiv* **2019**, arXiv:1904.10416. [CrossRef]
136. Hoover, C.M.; Ducey, M.J.; Colter, R.A.; Yamasaki, M. Evaluation of alternative approaches for landscape-scale biomass estimation in a mixed-species northern forest. *For. Ecol. Manag.* **2018**, *409*, 552–563. [CrossRef]
137. Coops, N.C.; Tompalski, P.; Goodbody, T.R.; Queinnec, M.; Luther, J.E.; Bolton, D.K.; White, J.C.; Wulder, M.A.; van Lier, O.R.; Hermosilla, T. Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends. *Remote Sens. Environ.* **2021**, *260*, 112477. [CrossRef]
138. Næsset, E. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sens. Environ.* **2005**, *98*, 356–370. [CrossRef]
139. Hopkinson, C. The influence of lidar acquisition settings on canopy penetration and laser pulse return characteristics. In Proceedings of the IEEE International Symposium on Geoscience and Remote Sensing, Denver, CO, USA, 31 July–4 August 2006.
140. Næsset, E. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and bio-physical stand properties derived from small-footprint airborne laser data. *Remote Sens. Environ.* **2009**, *113*, 148–159. [CrossRef]
141. Kamoske, A.G.; Dahlin, K.M.; Stark, S.C.; Serbin, S.P. Leaf area density from airborne LiDAR: Comparing sensors and resolutions in a temperate broadleaf forest ecosystem. *For. Ecol. Manag.* **2018**, *433*, 364–375. [CrossRef]
142. Yu, X.; Kukko, A.; Kaartinen, H.; Wang, Y.; Liang, X.; Matikainen, L.; Hyyppä, J. Comparing features of single and multi-photon lidar in boreal forests. *ISPRS J. Photogramm. Remote Sens.* **2020**, *168*, 268–276. [CrossRef]
143. Marino, E.; Tomé, J.L.; Hernando, C.; Guijarro, M.; Madrigal, J. Transferability of Airborne LiDAR Data for Canopy Fuel Mapping: Effect of Pulse Density and Model Formulation. *Fire* **2022**, *5*, 126. [CrossRef]
144. Goodwin, N.R.; Coops, N.C.; Culvenor, D.S. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sens. Environ.* **2006**, *103*, 140–152. [CrossRef]
145. Shao, G.; Stark, S.C.; de Almeida, D.R.; Smith, M.N. Towards high throughput assessment of canopy dynamics: The estimation of leaf area structure in Amazonian forests with multitemporal multi-sensor airborne lidar. *Remote Sens. Environ.* **2019**, *221*, 1–13. [CrossRef]
146. Van Leeuwen, M.; Nieuwenhuis, M. Retrieval of forest structural parameters using LiDAR remote sensing. *Eur. J. For. Res.* **2010**, *129*, 749–770. [CrossRef]
147. Karjalainen, T.; Korhonen, L.; Packalen, P.; Maltamo, M. The transferability of airborne laser scanning based tree-level models between different inventory areas. *Can. J. For. Res.* **2019**, *49*, 228–236. [CrossRef]
148. Van Ewijk, K.Y.; Treitz, P.M.; Scott, N.A. Characterizing Forest Succession in Central Ontario using Lidar-derived Indices. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 261–269. [CrossRef]
149. Melin, M.; Hinsley, S.A.; Broughton, R.K.; Bellamy, P.; Hill, R.A. Living on the edge: Utilising lidar data to assess the importance of vegetation structure for avian diversity in fragmented woodlands and their edges. *Landsc. Ecol.* **2018**, *33*, 895–910. [CrossRef]
150. Sumnall, M.J.; Trlica, A.; Carter, D.R.; Cook, R.L.; Schulte, M.L.; Campoe, O.C.; Rubilar, R.A.; Wynne, R.H.; Thomas, V.A. Estimating the overstory and understory vertical extents and their leaf area index in intensively managed loblolly pine (*Pinus taeda* L.) plantations using airborne laser scanning. *Remote Sens. Environ.* **2020**, *254*, 112250. [CrossRef]
151. Coops, N.C.; Hilker, T.; Wulder, M.A.; St-Onge, B.; Newnham, G.; Siggins, A.; Trofymow, J.A. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees* **2007**, *21*, 295–310. [CrossRef]
152. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *For. Ecol. Manag.* **2012**, *275*, 117–129. [CrossRef]