University of Ulster

Estimation of forest aboveground biomass using multisource remote sensing data and machine learning regression

██████████████

████████████

Research Project Final Report

MSc Remote Sensing and Geographic Information Systems

School of Geography & Environmental Science

Supervised by Robert McNabb

April 2024

Abstract

Forests play a pivotal role globally in climate change mitigation, biodiversity conservation, and supporting the economies of local communities. Aboveground biomass (AGB) is an essential metric for understanding forest dynamics. This study seeks to address challenges in estimating AGB by synthesising the best practices from prior research. It aimed to map AGB in Västra Götaland, southwestern Sweden, utilising open-access data. Earth Engine (GEE) was used to process and extract synthetic aperture radar (SAR) data from Sentinel-1 (S-1), optical imagery from Sentinel-2 (S-2), and topographic data. Using Random Forest (RF) regression, this data was then used to develop a forest AGB model with moderate accuracy (R^2 = 0.50, RMSE% = 46.83%). Following that, AGB models incorporating forest height information from three different sources - a global canopy height dataset, airborne LiDAR, and a RF mean height (MH) estimation model - were compared. The RF estimation of MH achieved only moderate accuracy (R^2 = 0.52, RMSE% = 30.18%), but its incorporation into the AGB model improved estimated performance (R^2 = 0.60, RMSE% = 42.00%). This model was subsequently chosen to map AGB. The model's tendency to underestimate high AGB values was evident when mapping AGB, with mean uncertainty 38.90% of the mean predicted biomass. Nevertheless, this study demonstrates an approach of data fusion and algorithm optimisation, including iterative feature elimination and hyperparameter tuning in the RF model to achieve superior model performance. This approach provides a foundational framework for estimating AGB at a regional scale, offering a scalable and cost-effective method for potential long-term monitoring of forest biomass dynamics. While the current model accuracy does not fully support widespread application, it has yielded important insights and established a solid basis for future enhancements.

1. Introduction

Forests provide key ecosystem services, contain a wide variety of resources, and are the world's primary biodiversity hotspots. They are essential for climate change mitigation through absorption of CO_2 from the atmosphere and storing it in their biomass – a process known as carbon sequestration [1,2]. Beyond visible wood and foliage, the biomass below ground and underlying soil are also major carbon sinks, making forests the largest terrestrial ecosystems for carbon storage by a large margin [3,4]. The loss of forests is a major driver of climate change, with greenhouse gases emitted through the burning and decomposition of removed biomass, and the reduction of the carbon sequestration capacity of the

underlying soil [5–7]. In addition to the carbon cycle, forests are further interlinked with the global climate system through regulation of atmospheric water and energy [8].

Human activity is the major driver on changes in forest ecosystems [9]. Direct changes come from removal of forest for agriculture, pasture, or wood harvesting [10]. There are also indirect impacts from anthropogenic activity. Climate change has been shown to have several harmful long-term impacts on forest ecosystems, such as degradation through conversion to vegetation that is drier and of lower biomass, reduction of suitable growing area by shifting boreal forests northwards and forcing montane forests onto higher slopes, and increasing the outbreaks of disease and harmful insect pests [5,11,12]. Effective understanding and management of forests depend on accurate and current data across both national and international contexts [7,11,13]

Besides their carbon sequestration potential, forests are habitats for at least half of the earth's plants, animals, and fungi [9,14]. The function of these ecosystems is strongly dependent on biodiversity, which is positively correlated with biomass [15–17]. It is not only the removal of forests that threatens biodiversity, but also the ecosystem fragmentation and loss of interconnectedness [18]. A further consideration is the importance of forests to those who live in and around them. Forests hold huge cultural and societal value to the people indigenous to them, as well as providing food, fuel, water, construction material, medicine, livelihood, and other benefits [19–21]

Forest aboveground biomass (AGB) serves as the key metric in international carbon trading schemes, necessitating reliable long-term information about AGB due to its dynamic nature and potential for change [22]. Furthermore, assigning an economic value to forest carbon is also important for climate policy, assessing the success of forest protection programmes, and for generating collective willingness to fund such efforts [23]. The United Nations Framework Convention on Climate Change (UNFCCC) obliges nations to regularly report emissions and loss of carbon sequestering vegetation through programmes such as the Reducing Emissions From Deforestation And Forest Degradation In Developing Countries (REDD+) scheme [24,25].

However, there are numerous challenges associated with estimation of AGB, particularly for forests with high diversity of tree species. Traditionally, AGB estimation involved the destructive and labour-intensive process of felling trees and weighing them. Following that, allometric equations were developed where parameters such as tree height and diameter at breast height (DBH) were related to biomass through regression [26,27]. These equations are specific to both tree type and region, but work on the principle that trees with greater diameter or height hold more biomass [28,29]. Field-based methods are time consuming, expensive, cover only a small area, and cannot be effectively scaled up in forest stands where trees are not of the same age and species [22].

Remote sensing can gather data from large areas at relatively low cost. This approach is often paired with forest inventories taken at carefully selected locations (plots) that are meant to be representative of the area being sampled [30,31]. This information is then used to train and validate models to estimate AGB and other metrics at a broad scale. This approach, while increasingly effective, is not devoid of limitations, and uncertainties persist in the estimation of AGB due to various factors throughout the process [32–34]. The ability to estimate structural metrics from remote sensing data has made them an important component in biomass estimation [35–37]. Estimating tree height has become particularly important, with novel methods of relating it to biomass being developed in response to the limitations in using allometric equations [38–42].

Synthetic Aperture Radar (SAR) sensors (such as Sentinel-1 and ALOS PALSAR) are active remote sensing systems and able to operate irrespective of time of day and can penetrate cloud, smoke, and

haze. The amount of radiation returned to the receiver (backscatter) is dependent on the wavelength of the emitted microwave radiation, and the size and electromagnetic properties of the object with which it is interacting. Longer wavelengths (L-band) tend to produce better estimations of forest biomass than shorter (C -band) due to their interaction with larger canopy elements (such as branches and stem) that tend to comprise the majority of AGB [43–45]. Sentinel-1 (S-1), first launched in 2014, is a C-band mission that operates at 5.405GHz with a revisit time of 12 days. Because of the limitations of C-band radar in AGB estimation, S-1 tends to be used in conjunction with other platforms [46,47].

Optical space-borne remote sensing platforms (such as Sentinel-2, Landsat, and Hyperion) have been in operation for many decades and are widely used in AGB studies [48,49]. However, data collection is affected by cloud cover and can only be undertaken during the daytime. In forest AGB studies, optical remote sensing is limited by signal saturation at high levels of biomass due to canopy closure [50,51]. Sentinel-2 (S-2) is a dual platform, super-spectral satellite mission following a polar orbit. The first satellite was launched in 2015, and the second in 2017. S-2 has a high likelihood of finding cloud-free imagery due to it being a dual platform mission with a revisit time of 5 days, and has been effectively used to study AGB in varying forest environments [52–54]. Monitoring vegetation dynamics and land cover characteristics is a key intention of the S-2 project [55]. This includes the generation of land-cover maps and vegetation parameters known as biophysical variables (BPVs) at regional, continental, and global scales.

Multi-sensor synergy allows researchers to overcome the limitations inherent in platforms [56]. AGB studies of this nature have become increasingly popular with the increasing number of satellite different platforms and the advancement of UAV technology in the last two decades [57–61]. The Sentinel missions in particular were designed to be used in a complementary manner [62]. Since the launch of S-2 in 2015, studies have effectively combined S-1 and S-2 data for a variety of biomass estimation applications such as wheat, mangroves, and a variety of forest types [63–65]. Combining data from both optical and SAR platforms has shown to be able to overcome the saturation problem that occurs in both sensor types when studying forests with high-density [66,67]. Ghosh and Behera [68] showed that combining S-1 texture images and vegetation indices from S-2 data, they were able to predict tropical forest biomass with a high accuracy.

Light Detection and Ranging (LiDAR), also known as laser scanning, is a comparatively recent technology in remote sensing able to provide very high-resolution three-dimensional imagery. Airborne LiDAR instruments are able to derive direct information about the structural characteristics of forests such as tree heights, canopy density, and tree species, which can be accurately related to forest biomass [69–72]. LiDAR's ability to gather precise readings of forest structure means that is often used as a reference when using satellite data across a larger area [73–75].

Google Earth Engine (GEE) is a free, cloud-based platform for "petabyte-scale" processing, analysis, and visualisation of geospatial data [76]. Its public data archive includes a wide variety of historical and real-time satellite imagery, topography data, land cover layers, demographic and infrastructure maps, vector shapefiles, and climate datasets [77]. In addition, there is also a collection of community-created datasets that are available for use under open licenses [78]. It enables high-performance computing without the need for expensive hardware or complex software, democratising access to advanced remote sensing methods and facilitating advanced analysis at scale [79].

Machine learning (ML) algorithms are adept at applying a flexible non-linear models to find relationships between input and output data in a wide variety of applications [80]. The random forest (RF) machine learning algorithm is an ensemble non-parametric method that can be used for both classification and regression tasks [81]. It uses numerous decision trees - a forest - and for each tree

randomly selects subsets of training data. For each split in the tree, it randomly selects a subset of predictor variables to use. When the predictions of these diverse trees are averaged together, they produce a robust model that is resistant to overfitting. It is a popular choice due to it having a wide variety of applications, being fairly simple to train and use, and having the ability to capture complex interactions in data while being resistant to noise or outliers [82–85]. RF is able to statistically measure the importance and weight of variables, detect outliers, and impute missing values [86]. It has also been shown to be a suitable model for estimating forest AGB in numerous studies [50,87–90].

Boreal forests are the world's largest land biome. They provide vital ecosystem services and hold more biomass than any other [91]. Recent research has suggested that younger northern temperate and boreal forests are playing a crucial role in global carbon sequestration, helping to offset the emissions caused by tropical forest loss [92]. Within Europe, the forests of the boreal-dominated Fennoscandia region in the north are the most biomass-dense, with the majority of this being held in southern Sweden and Finland [93]. The forestry industry in these countries goes back centuries, and as a result they have long histories of forest management [94]. However, recent literature has found that the largest losses of forest extent in Europe between 2001 and 2021 occurred in this region [95].

Sweden is 70% forested, with its total forest stock doubling in the past century [96]. The forestry industry is of major economic importance to the country, being the world's 4th largest exporter of forestry products and employing over 115 000 people [97,98]. 80% of forested land in Sweden is defined as being in "active use", with 1% being felled every year [99]. Keeping track of forest stock is important for managing timber harvests, maintaining a steady long-term supply, and tracking targets for programmes such as REDD+ [100]. Timber industry plantations tend to be fast growing monocrops which are eventually harvested, causing biomass to fluctuate [101]. Intense harvests also have implications for soil nutrients, water chemistry, and biodiversity [102]. In addition to the traditional products of pulp, paper, and sawn timber, wood is now being seen as a sustainable building material and energy source [103]. For this to become a reality, information about the abundance, quality, and distribution is required [104].

Past studies on AGB estimation in Sweden demonstrate a diverse array of approaches and underscore the complexity of accurately assessing forest biomass in northern regions. Jalkanen et al. [105] evaluated the accuracy of direct biomass estimation methods and found that southern spruce-dominated and young forests have the highest uncertainty, indicating the significant impact of seasonality and stand development on estimation accuracy. Shendryk et al. [106] developed a method for estimating AGB in conifer-dominated forests using a fusion of low-density LiDAR and SPOT-5 imagery, demonstrating the integration's effectiveness with an R^2 of 0.80. Askne and Santoro [107] demonstrated a method for estimating AGB from TanDEM-X data without reliance on training samples, achieving a relative RMSE of 16% to 32% across 202 forest stands. Persson et al. [108] used TanDEM-X data for generating large-scale, high-resolution maps of AGB and forest stem volume across the country, highlighting the effectiveness of SAR data in forest mapping, alongside challenges such as the complexity of processing extensive datasets and the need for robust algorithms to handle the variability within boreal forest environments. Persson and Fransson [40] compared the accuracy of TanDEM-X and LiDAR data for estimating AGB and tree height in boreal forests, finding comparable accuracies at the stand level, with a particular emphasis on the potential of InSAR data to approximate LiDAR-derived metrics like the 99th percentile height and vegetation ratio. Persson et al. [109] demonstrated the effectiveness of combining TanDEM-X and S-2 data with field inventory data for large-area, species-wise prediction of forest biomass and volume, achieving a relative RMSE of 29.1% for AGB, while also successfully classifying the dominant tree species with an overall accuracy of 77%. Majasalmi and Rautiainen [110] conducted a simulation study on the potential of S-2 data for estimating biophysical variables in boreal forests,

revealing that while specific band combinations showed promise for estimation of some biophysical variables, they observed weak correlations between S-2 band combinations and above-ground biomass (AGB) estimation.

This study addresses the challenges in estimating biomass within Sweden's economically and ecologically important forests, focusing on the need for advanced remote sensing techniques to overcome issues like signal saturation in dense canopies, diversity of forest structures, and seasonal variations. It attempts to employ a novel approach to AGB estimation in the region by utilising machine learning to process multivariate and multitemporal data collected in GEE. Drawing on established methodologies from the scientific literature and adapting them to this specific context, the research aims to contribute to understanding of northern forest biomass dynamics and forest management practices.

2. Data and Methods

2.1. Study Area

The county of Västra Götaland lies in the south-west of Sweden, covering 28 772 km^2 (Figure 1). It is situated in the hemi-boreal transition zone, where temperate and boreal forest biomes overlap, featuring a mix of coniferous and deciduous trees.
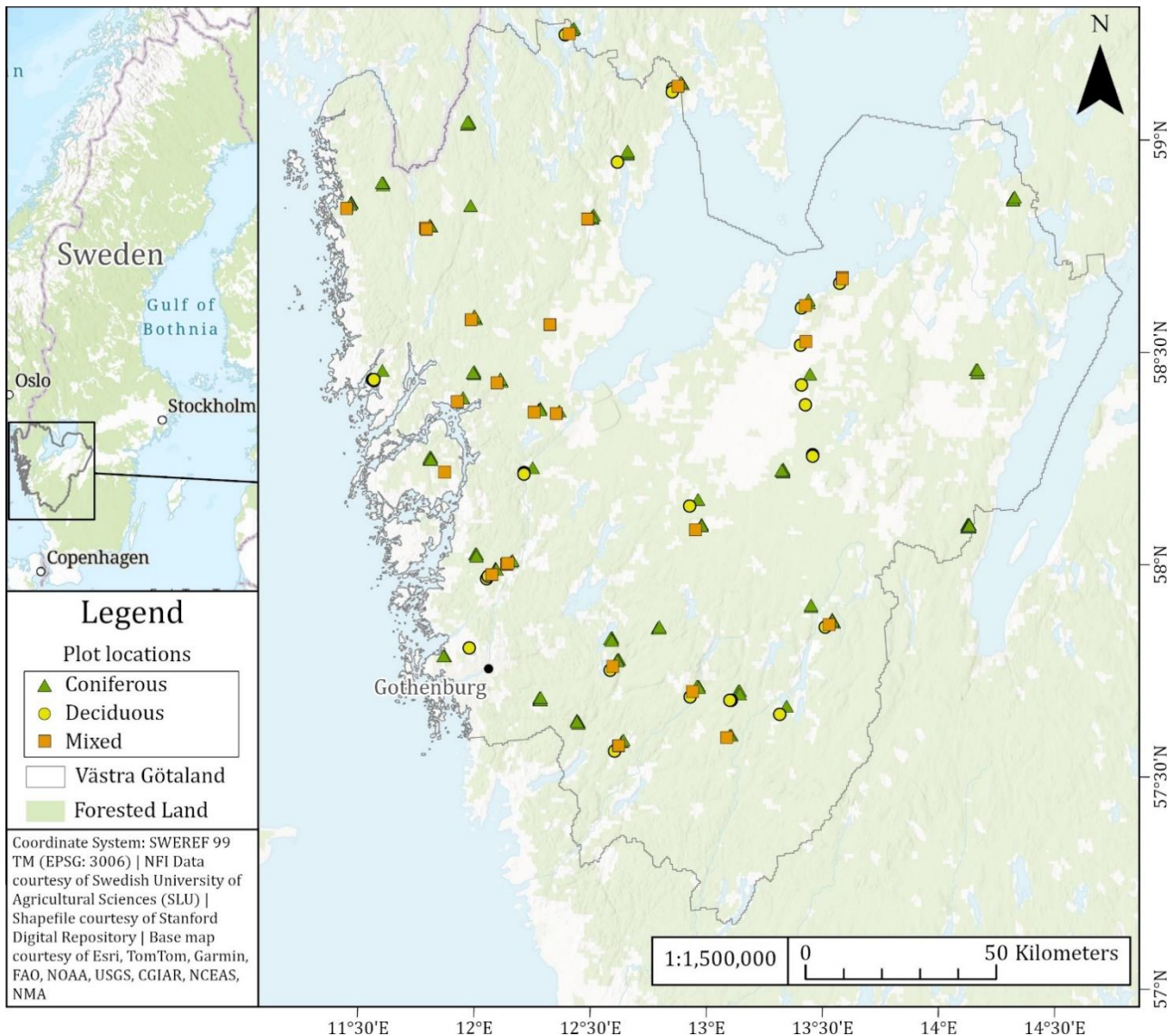

Figure 1. Map of study area, with plot locations by dominant tree type and hillshade to show elevation changes.

Västra Götaland is characterised by coastal lowlands and interior highlands. The climate is temperate, with varied seasons. The warmest month is July with an average temperature of 17.3°C. January and February are the coldest months, with an average temperature of -0.3°C. Precipitation is distributed throughout the year with wet summers and winter snowfall common in elevated areas.

2.2. Forest Inventory Data

The Swedish National Forest Inventory (NFI) is carried out by annually sampling the entirety of Sweden with a systematic network of inventory tracts. AGB is determined by measuring the diameter at breast height (DBH) and height of all trees within these plots, recording dead wood and understorey vegetation, and then applying allometric equations using these dimensions. The inclusion of dead wood means that biomass is still present at sites that have been recently cut and as a result are assigned a mean height (MH) of zero. MH is calculated as the average height of the trees in the plot, weighted by basal area [111]. Plots are circular with 7 m radius. All plot data for this study is from the 2021 inventory. The 147 plots located in Västra Götaland used after pre-processing and outlier removal (detailed below) are summarised by tree type in Table 1. Plots that are deciduous-dominated (more than two-thirds of the total plot biomass) made up 17% of the total, while coniferous-dominated make up 67%, and mixed plots the remaining 16%.

Table 1. Summary of aboveground biomass (AGB) and mean height (MH) by dominant tree type for NFI plots used in Västra Götaland.

		Coniferous	Deciduous	Mixed	All
	Count (n)	99	25	23	147
Aboveground Biomass (Mg.ha^{-1})	Minimum	5.48	5.30	11.42	5.30
	Maximum	433.54	399.30	384.55	433.54
	Mean	139.08	130.50	129.25	136.08
	Standard Deviation	92.08	102.01	80.96	92.33
Mean Height (m)	Minimum	0.00	0.00	3.30	0.00
	Maximum	29.70	28.20	24.00	29.70
	Mean	15.77	15.76	14.86	15.63
	Standard Deviation	6.84	8.34	5.72	6.97

2.3. Satellite Data Collection and Pre-processing

All collection, pre-processing, and extraction of data at plot level was done using Google Earth Engine. See Figure 2 for an overview of the methodological approach.

Unforested land was masked out using the ESA 2021 Worldcover dataset to limit soil background effects and minimise computational requirements. This dataset is developed from both S-1 and S-2 data and available at 10m resolution [112]. Given S-2 data's sensitivity to cloud coverage, which influences the usability of study plots, it was prioritised in the data collection sequence.

Sentinel-2

S-2 Level-2A data were used, these images have undergone orthorectification and atmospheric correction to surface reflectance values [113]. In total, 10 optical bands were used to derive 29 spectral indices (SIs) and 5 biophysical variables (BPVs) (Table 2). The selection of these indices and BPVs was informed by their relevance to vegetation monitoring [54,64,66,68,114,115].
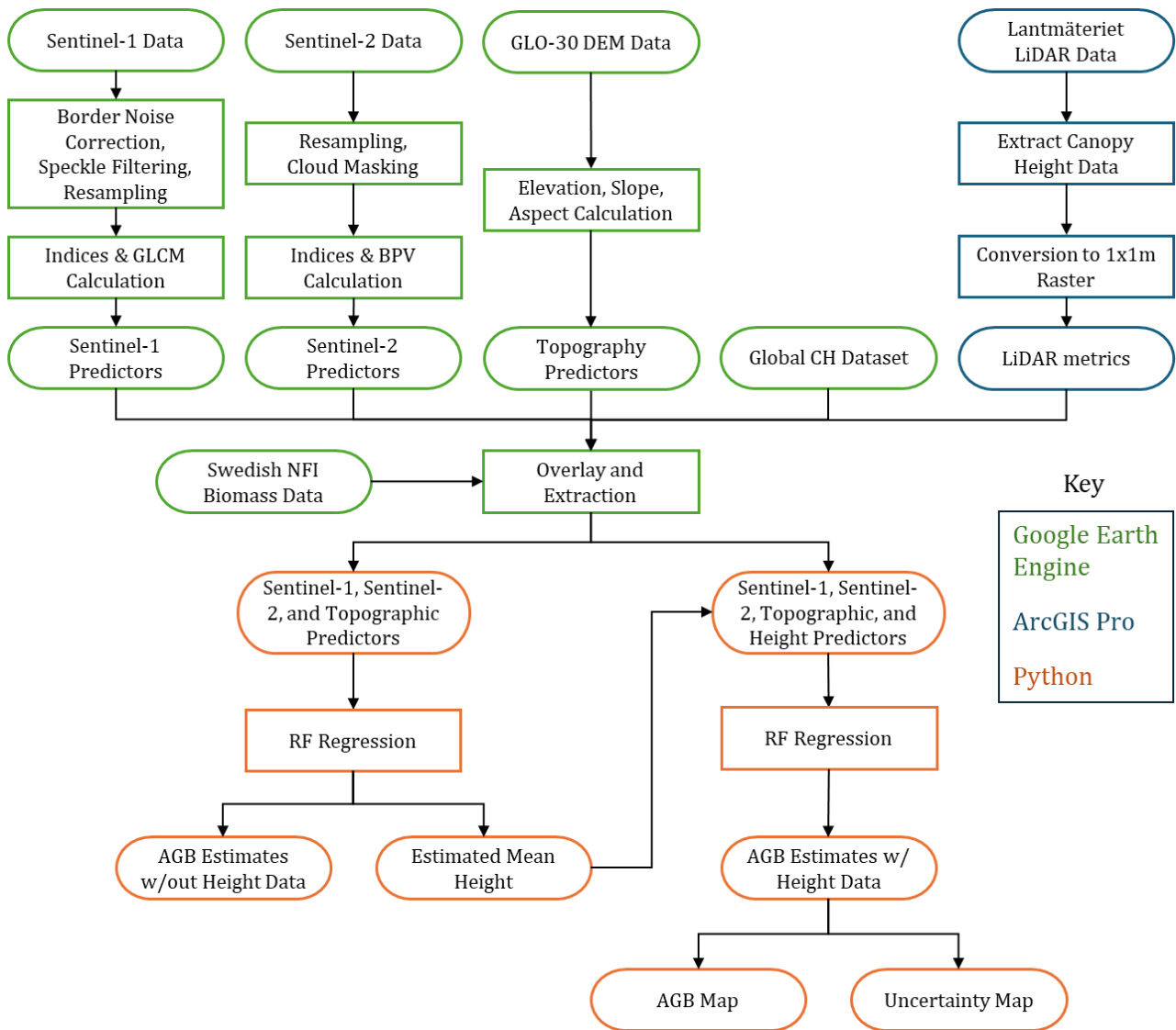
Figure 2. Overview of Methodology. For a detailed breakdown of the RF model, see Figure 4.

Monthly composites from May to September 2021 were calculated to capture the distinct stages of growth in vegetation from late spring to early autumn. These months offer the highest probability of producing suitable imagery, benefitting from prolonged daylight hours – a crucial factor in Scandinavia where daylight is limited in winter. Additionally, using composites rather than single scene images increased the odds of finding clear imagery in a region characterised by the frequent cloud cover. The median is chosen as the method of averaging to account for the potential presence of extreme or outlier values, and a common technique used in studies of this nature [88,116,117].

Clouds were masked using the cloud displacement index (S-2c), Cloud Probability Map (MSK), and cloud probability band created for GEE (Prob) [118–120]. Initial cloud mask probability thresholds (MSK: 50%, S-2c: 60%, Prob: 20%) were established based on visual inspection of the processed imagery. Snow masking was applied using the S-2 cloud probability band and a normalised difference snow index (NDSI) to account for the possible presence of snow in the highland areas [113]. Visual inspection was again used to set the snow probability band at threshold 30% and the NDSI threshold at 25%. 20 plots were removed from the dataset due to being clouded over. For every remaining plot, 44 predictors were derived for each month, totalling 220 across the 5 months studied.

Table 2. Predictors extracted and derived from Sentinel-2 (S-2) data.

Product	Band Number (Name)	Identifier	Central Wavelength	Spatial Resolution	Reference
Bands	2 (Blue)	B2	490 nm	10m	
	3 (Green)	B3	560 nm	10m	
	4 (Red)	B4	665 nm	10m	
	5 (Red edge 1)	B5	705 nm	20m	
	6 (Red edge 2)	B6	749 nm	20m	[55]
	7 (Red edge 3)	B7	783 nm	20m	
	8 (NIR)	B8	842 nm	10m	
	8A (NIR)	B8A	865 nm	20m	
	11 (SWIR 1)	B11	1610 nm	20m	
	12 (SWIR 2)	B12	2190 nm	20m	

	Name	Identifier	Formula	Reference
Indices	Difference Vegetation Index	DVI	B8 – B4	[121]
	Enhanced Vegetation Index	EVI	2.5 * ((B8 – B4) / (B8 + 6 * B4 – 7.5 * B2 + 1))	[122]
	Enhanced Vegetation Index 2	EVI2	2.5 * (B8 – B4 / B8 + 2.4 * B4+ 1)	[123]
	Green Normalized Difference Vegetation Index	GNDVI	(B8 − B3) / (B8 + B3)	[124]
	Green-Red Vegetation Index	GRVI	(B3 – B4) / (B3 + B4)	[125]
	Inverted Red-Edge Chlorophyll Index	IRECI	(B8 – B4) / (B5 / B6)	[126]
	Modified Chlorophyll Absorption in Reflectance Index	MCARI	((B5 – B4) − 0.2 (B5 – B3)) * (B5 / B4)	[127]
	Modified Soil-Adjusted Vegetation Index	MSAVI	$(2 * B8 + 1 - \sqrt{(2 * B8 + 1)^2 - 8 * (B8 - B4)}) / 2$	[128]
	Normalised Burn Ratio	NBR	(B8 – B12) / (B8 + B12)	[129]
	Normalised Burn Ratio 2	NBR2	(B11 – B12) / (B11 + B12)	[130]
	Normalized Difference Index 45	NDI45	(B5 – B4) / (B5 + B4)	[131]
	Normalized Difference Infrared Index	NDII	(B11 − B4) / (B11 + B4)	[132]
	Normalized Difference Red Edge Index 1	NDRE1	(B7 – B5) / (B7 + B5)	[133]
	Normalized Difference Red Edge Index 2	NDRE2	(B8 – B5) / (B8 + B5)	[133]
	Normalized Difference Vegetation Index 1	NDVI1	(B8 – B4) / (B8 + B4)	[134]
	Normalized Difference Vegetation Index 2	NDVI2	(B8a – B4) / (B8a + B4)	[64]
	Normalized Difference Vegetation Index Red Edge	NDVIRE	(B6 – B4) / (B6 + B4)	[66]
	Optimized Soil Adjusted Vegetation Index	OSAVI	(1.16) * (B8 – B4) / (B8 + B4 + 0.16)	[135]
	Pigment Specific Simple Ratio - Chlorophyll A	PSSRa	B7 / B4	[136]
	Ratio Vegetation Index	RVI	B8 / B4	[125]
	Red-Edge Simple Ratio Vegetation Index	RERVI	B8a / B5	[137]
	Red-Edge Chlorophyll Index	CIre	B8a / (B5 – 1)	[138]
	Red-Edge Re-Normalized Difference Vegetation Index	RERDVI	$(B8a − B5) / \sqrt{(B8a/B5)}$	[139]
	Renormalized Difference Vegetation Index	RDVI	$(B8 – B4) / \sqrt{(B8 + B4)}$	[140]
	Soil-Adjusted Vegetation Index	SAVI	(B8 – B4) / (B8 + B4 + L) * (1.0 + L), L = 0.5	[141]
	Stress Related Vegetation Index 1	STVI1	(B11 * B4) / B8	[142]
	Stress Related Vegetation Index 2	STVI2	B8 / (B4 * B12)	[142]
	Stress Related Vegetation Index 3	STVI3	B8 / (B4 * B11)	[142]
	Transformed Chlorophyll in Reflectance Index	TCARI	3 * ((B5 − B4) − 0.2 * (B5 − B3) * (B5/B4))	[143]

	Name	Identifier	Description	Reference
Biophysical Variables	Leaf Area Index	LAI	Area of leaves per unit area of ground, indicating the density of vegetation.	[144]
	Fraction of Vegetation Cover	FCOVER	Proportion of ground covered by vegetation, reflecting vegetation density and distribution.	[145]
	Fraction of Absorbed Photosynthetically Active Radiation	FAPAR	Fraction of solar radiation in the photosynthetically active wavelengths (400-700 nm) absorbed by the vegetation canopy.	[146]
	Leaf Chlorophyll Content	CAB	Amount of chlorophyll in leaves, indicating the health and productivity of plants.	[147]
	Canopy Water Content	CWC	Total amount of water held in the vegetation canopy, reflecting vegetation health and stress.	[148]

Sensitivity analysis was performed to find the pre-processing settings that produced the least uncertainty (lowest RMSE and highest R^2) for predicting AGB. This was done iteratively by producing a raster image using variations of the described parameters, extracting the metrics at each plot, using the extracted information to train and test the random forest model (described below), and assessing the accuracy of the results.

The first step performed was to compare the effects of the pixel size of the produced raster image. Images were produced at 10 and 20m spatial resolution using the default bilinear resampling method and the results compared. At the cloud masking stage, higher and lower thresholds from those initially set were tested in incremental steps ($\pm$ 5 percentage points for all three settings). This step was performed in a grid analysis comparing bilinear and bicubic methods of resampling. A grid method allows for a more comprehensive exploration of the interdependent nature of parameters in image pre-processing through testing all possible combinations of two parameters. Once the optimal settings for this dataset were found and the final set of study plots established (Table 6), testing of S-1 followed.

Sentinel-1

For S-1, VV and VH backscatter metrics were collected and used to derive indices (Table 3). Interferometric wide swath mode data was used. These images are available at 5 x 20m resolution. S-1 data undergoes orbit file application, border noise removal, thermal noise removal, radiometric calibration, and orthorectification before inclusion in GEE [149,150]. Data was further pre-processed – additional border noise correction and speckle filtering - using the method developed by Mullissa et al. [151].

Monthly median averages for S-1 data were calculated to assess vegetation at distinct growth stages across the entire year. Recent studies of forest biomass have found that some SAR backscatter data coefficients for AGB were highest in the winter when deciduous trees were missing their leaves [46,152]. Seasonality has also been shown to be important for SAR data used in other biomass studies [66,153].

Various indices that have been used in recent literature were also derived from these backscatter polarisations. The Grey Level Co-occurrence Matrix (GLCM) is a statistical method of analysing texture that considers the spatial relationship between pixels [154]. Previous studies have shown GLCM texture metrics calculated from SAR data to be useful in biomass estimation in tropical [155,156], mediterranean [157], and sub-tropical settings [158]. The contrast, entropy, sum entropy, difference entropy, correlation, angular second moment, variance, sum variance, difference variance, inverse difference moment, dissimilarity, cluster shade, and sum average were derived for both VV and VH polarisations. 37 predictors were derived for each month, totalling 444 across the 12 months.

Pre-processing sensitivity analysis for S-1 data began with the speckle filtering step to find the method, window size, and number of images that produced the highest sensitivity. Initially speckle filter window size, number of speckle filter images, and the GLCM window size were all set to 3 for these tests. The Refined Lee, Lee Sigma, and Gamma Map speckle filtering methods were compared. Following that, a grid analysis of the speckle filter window size and number of images was performed. The terrain flattening shadow buffer was then investigated, using 25 m increments from 0 to 100 m. The next step was to determine best spatial resolution to resample the bands to (10 or 20 m). Finally, the window size for the GLCM was then tested at 3x3, 5x5, 7x7, and 9x9.

Iterations from the final step of the S-1 fine-tuning that showed promising results were then assessed in combination with the S-2 data to find the most suitable for capturing the non-linear variable interactions that are inherent in the RF model. The combination that produced the highest estimation accuracy was used in the analysis (Table 7).

Table 3. Predictors extracted and derived from Sentinel-1 (S-1) data.

Product	Name	Identifier	Details	Reference

Back-scatter	Vertical Transmit – Vertical Receive	VV	Return signal when the radar transmits and receives in the vertical polarisation	
	Vertical Transmit - Horizontal Receive	VH	Return signal when the radar transmits in the vertical polarisation and receives in the horizontal polarisation	[159]

	Name	Identifier	Formula	Reference
Indices	Backscatter Ratio	VVVH_ratio	VV / VH	-
	SAR Simple Difference Index	sarSDI	VH - VV	[152]
	SAR Simple Additive Index	sarSAI	VH + VV	
	SAR Multiplication Index	sarMI	VH × VV	
	SAR Ratio Index 1	sarRI1	VH / (VH × VV)	[115]
	SAR Ratio Index 2	sarRI2	(VH + VV) / (VH × VV)	
	SAR Square Difference Index	sarSqDI	(VH × VH) – (VV × VV)	
	Normalized Difference SAR Index	sarNDI	(VH – VV) / (VH + VV)	-

	Name	Identifier	Details	Reference
Gray-Level Co-occurrence Matrix (GLCM)	Contrast	GLCM_VV_contrast / GLCM_VH_contrast	Measures the local variations in the Gray-level co-occurrence matrix.	
	Entropy	GLCM_VV_ent / GLCM_VH_ent	Measures the randomness in the image texture, indicating complexity.	
	Sum Entropy	GLCM_VV_sent / GLCM_VH_sent	Calculates the entropy of the sum of pairs of pixels, representing the complexity in the sum distribution of the GLCM.	
	Difference Entropy	GLCM_VV_dent / GLCM_VH_dent	Measures the entropy in the difference distribution of pairs of pixels, representing complexity in contrast variation.	
	Correlation	GLCM_VV_corr / GLCM_VH_corr	Indicates how correlated a pixel is to its neighbours over the whole image.	
	Angular Second Moment	GLCM_VV_asm / GLCM_VH_asm	Measures the uniformity of the GLCM.	
	Variance	GLCM_VV_var / GLCM_VH_var	Represents the variance (spread) from the mean in the GLCM.	[154,161]
	Sum Variance	GLCM_VV_svar / GLCM_VH_svar	Measures the variance of the sum of pixel values.	
	Difference Variance	GLCM_VV_dvar / GLCM_VH_dvar	Measures the variance of the difference in pixel values.	
	Inverse Difference Moment	GLCM_VV_idm / GLCM_VH_idm	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal	
	Dissimilarity	GLCM_VV_diss / GLCM_VH_diss	Measures how dissimilar the pixel pairs are in the GLCM.	
	Cluster Shade	GLCM_VV_shade / GLCM_VH_shade	Measures the asymmetry of the GLCM, indicating skewness and uniformity in the pixel values distribution.	
	Sum Average	GLCM_VV_savg / GLCM_VH_savg	Calculates the average of the sum of pixel values.	

2.3.1.Outlier Detection

Following the collection of satellite data, the data was inspected for outliers and plots that the forest mask may have erroneously classified. A scatterplot of NFI MH vs AGB was inspected to identify points that lie far outside of the cluster of points (Figure 3). This method was chosen to maintain the variance in the dataset and not lose the values that would be classified as outliers when using methods such as interquartile range or Z-score. Potential outlier plots were inspected on high resolution Google Earth imagery to determine if they were anomalous. A single plot, consisting of a single large tree in an open field, was deemed to be not representative of a forest according to the Food and Agriculture Organization's definition [14]. The final sample size after outlier removal was 147 plots.
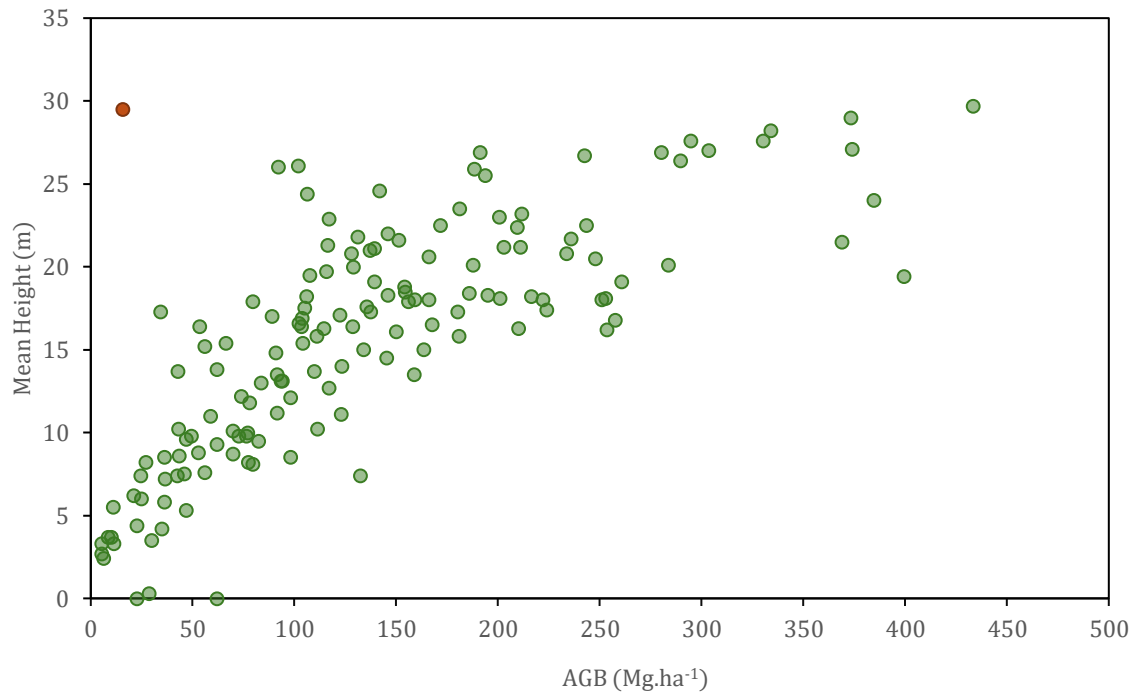
Figure 3. Mean Height (MH) vs AGB from Swedish NFI data for study area. The single outlier plot removed is shown in orange. Points with height values of zero indicate plots with only dead biomass.

2.4. DEM Data

Elevation, slope, and aspect, derived from the Copernicus Global 30m Digital Elevation Model (GLO-30 DEM) [162] were included as predictor variables (Table 4), as these metrics have been shown to have a correlation with biomass [163,164]. Castillo et al. [165] found that the inclusion of elevation data aided improved the estimates of AGB when using S-1 and S-2 data. Zhao et al. [158] found that accounting for slope aspect improved estimation of AGB.

Table 4. Predictors extracted and derived from Copernicus Global 30m Digital Elevation Model (GLO-30 DEM).

Data product	Identifier	Details	Reference
Elevation	Elev	Height above sea level	
Slope	Slope	Slope steepness in degrees	[162]
Aspect	Aspect	Direction of slope face	

2.5. Tree Height Data

Incorporating tree height parameters is an effective method to reduce error in estimations of AGB [36,88,166,167]. The relationship between the observed AGB and MH values of the study plots was investigated to establish a quantitative basis for the use of height as a predictor variable. After finding a significant relationship between the AGB and MH, three height datasets were acquired.

First, an open source global canopy dataset published by Lang et al. (2023), available through the GEE community catalogue (Identifier: Global_CH). It is derived from 2020 S-2 data and available at 10m resolution. Second, LiDAR data acquired from the Swedish Mapping, Cadastral, and Land Registration Authority (Lantmäteriet). The organisation performs LiDAR surveys in 4-year cycles. The dates of capture for the study area ranged from 28/11/2019 – 27/04/2023. Using the LAStools add-on for ArcGIS Pro (version 3.2.2), canopy metrics (Table 5) were extracted and converted to raster format with 1m spatial resolution [169,170]. The minimum cutoff height, determined by the range of MH values for the study plots, was set at 0m (ground level).

Table 5. LiDAR metrics calculated for use in analysis.

Canopy Metric	Identifier	Description
Average	lidar_avg	Average height of lidar points
Kurtosis	lidar_kur	Bulge or tailedness of point distribution
Maximum	lidar_max	Maximum point height
Minimum	lidar_min	Minimum point height
P01	lidar_p01	1st height percentile
P05	lidar_p05	5th height percentile
P10	lidar_p10	10th height percentile
P25	lidar_p25	25th height percentile
P50	lidar_p50	50th height percentile
P75	lidar_p75	75th height percentile
P90	lidar_p90	90th height percentile
P95	lidar_p95	95th height percentile
P99	lidar_p99	99th height percentile
SKE	lidar_ske	Skewness of point distribution
STD	lidar_std	Standard deviation of points

Third, MH was estimated using the NFI data as a reference. Given the time constraints of the project, the pre-existing RF model - incorporating S-1, S-2, and topographic predictors - was employed to estimate MH (Identifier: Estimated_MH) rather than developing a new model.

These height datasets were then incorporated into the RF model for AGB estimation, and their impact on prediction accuracy was evaluated through changes in RMSE and increase in R^2. The model using Estimated_MH was selected for mapping AGB. This decision was influenced by its comparative accuracy to the model using LiDAR data, the intensive data processing demands associated with LiDAR, and the study's focus on utilising satellite data.

2.6. AGB Modelling

2.6.1.Overlay and Extraction

Prepared rasters were merged to a single layer, upon which a shapefile of plot locations was overlayed. At each plot location, the average of the pixels intersecting the footprint of the circular field plot (r = 7m) was averaged, with values being adjusted for the proportion of the pixel that overlapped with the plot area.

2.7. Random Forest Regression

Once data was processed in GEE and predictor values were extracted at each plot, Python 3.8.18 [171] was used for analysis with the following libraries: Pandas (version 2.0.3) for data manipulation [172], Scikit-learn (version 1.3.0) for the random forest algorithm [173], Matplotlib (version 3.7.2) for data visualisation [174], Numpy (version 1.24.4) for calculation [175], and IPython (version 8.12.2) for interactive analysis [176].

The RF model's built-in variable importance ranking for feature elimination was a key consideration in designing this study. Through iterative feature elimination, the model is able to reduce a very large number of predictors to a small optimised set [177]. This allows the study to exploit the initial richness of information provided by multitemporal [115,152] and multivariate data [54,114], with the final model only making use of the most influential predictors. The feature removal loop can be seen in Figure 4. In a study area that sees large seasonal and topographic variations, this is beneficial for both AGB and MH estimation, which are influenced by different predictors.
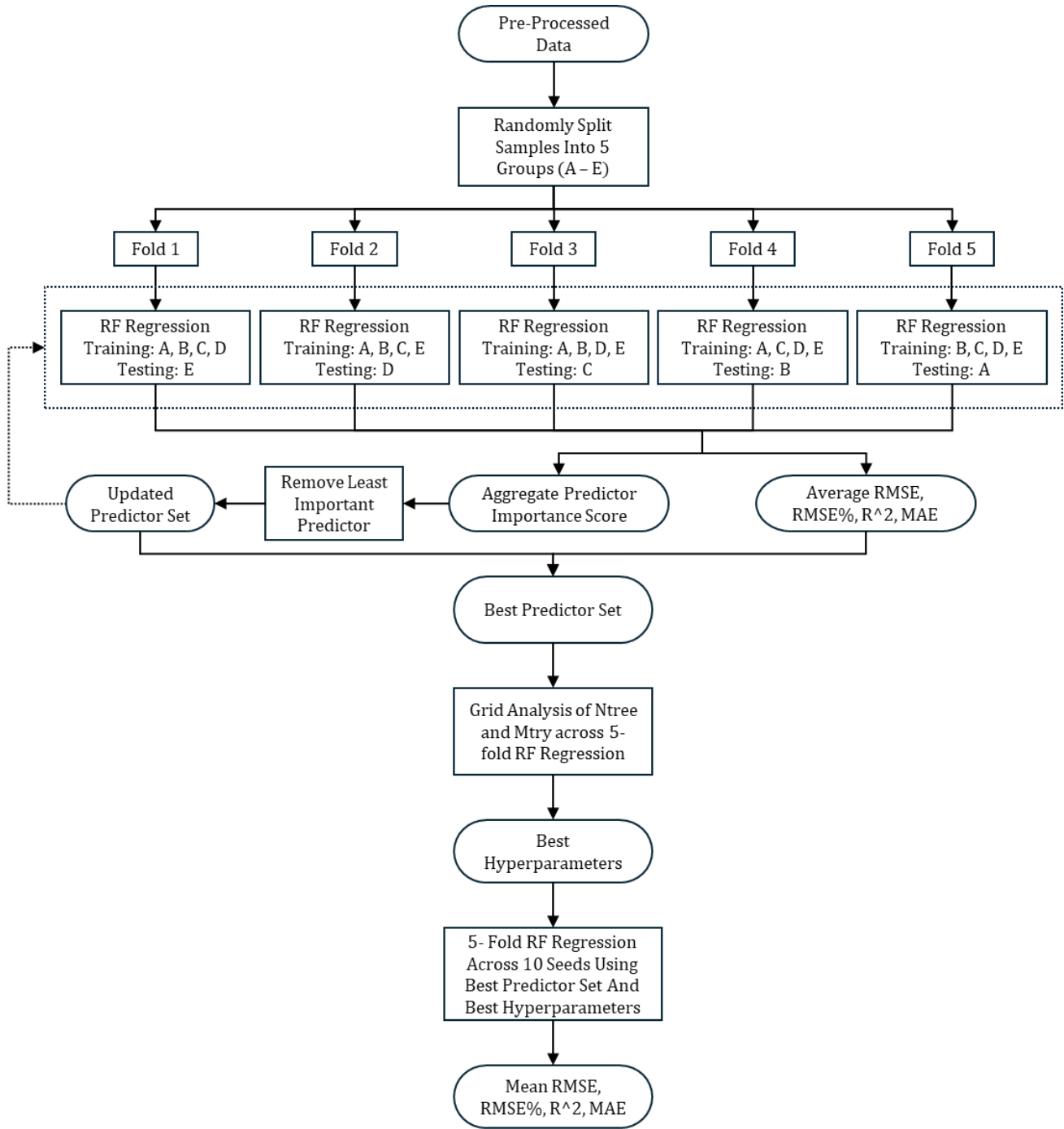
Figure 4. Flowchart demonstrating the random forest model used for analysis with 5-fold cross-validation feature removal loop and hyperparameter optimisation.

The accuracy metrics chosen were Root Mean Square Error (RMSE), Root Mean Square Error Percentage (RMSE%), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). RMSE aggregates the magnitudes of errors in predictions into a single measurement (Equation (1)), and is a standard metric used for assessing the error of regression models. RMSE% is a normalised version of RMSE that allows for comparison between different datasets (Equation (2)). R^2 is a measure of the proportion of variance in the dependent variable that is explained by the variance in the independent variables (Equation (3)). MAE measures the magnitude of the errors in a set of predictors (Equation (4)). Lower RMSE and MAE values, and higher R^2 values indicate a better regression model.

$$RMSE \ = \ \sqrt{\frac{1}{n} \sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \qquad (1)$$

$$RMSE\% \ = \ \left(\frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n}(\hat{y}_i - y_i)^2}}{\bar{y}} \right) \times 100 \qquad (2)$$

$$R^2 \ = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n}|\hat{y}_i - y_i| \qquad (4)$$

Where:

- $\hat{y}_i$ is the predicted value,
- y_i is the observed value,
- $\bar{y}$ is the average of actual values y_i,
- n is the number of observations.

For initial hyperparameter settings, the number of decision trees (N_{tree}) was set to 500, and the number of predictors per tree (M_{try}) was initially set to the default of the square root of the number of input predictors, as per the recommendations of Belgiu and Drăguţ [178].

A five-fold cross-validation was chosen for the modelling approach to enhance the robustness and predictive accuracy of AGB estimation. By partitioning the dataset into five distinct groups, this technique ensures comprehensive training and validation across the dataset, using each segment as a testing set in successive folds. This method not only maximises the utilisation of available data but also aligns with findings by Fassnacht et al. [179], who underscored the resilience of Random Forest model diagnostics to variations in validation settings, particularly emphasising the stability of RMSE. For the feature removal loop, the predictor importance score was averaged across all folds, and the lowest ranked predictor then removed from the dataset. The process is then repeated until all predictors have been removed. The optimum combination of predictor variables was deemed to be that which rendered the lowest RMSE value after the process had completed.

Once the combination of predictors that produced the lowest RMSE was found, hyperparameter tuning was conducted using a grid approach. N_{tree} ranged from 100 – 1000 in increments of 100, and M_{try} ranged from 1-10. Again, the best results were determined by the lowest RMSE values. Following that the optimum set of predictor variables and hyperparameters were used to determine the overall accuracy of the model by performing five-fold cross validation across 10 different seeds. By averaging the accuracy metrics over 50 total folds, a robust estimate of the models prediction accuracy is established [115].

The optimal predictor set and hyperparameter settings were then used to map AGB across the study area. Standard deviation (SD) for each pixel was computed by aggregating predictions from each tree in the RF regressor, illustrating the prediction uncertainty spatially.

3. Results

3.1. Pre-processing Settings

The combination of settings that produced the lowest RMSE for S-2 data can be seen in Table 6.

Table 6. Optimal pre-processing settings for S-2 data

Spatial Resolution	10m
Resampling Method	Bilinear
Cloud mask thresholds	MSK: 55, S-2c: 65, Prob: 25

The combination of settings that produced the lowest RMSE for S-1 data can be seen in Table 7.

Table 7. Optimal pre-processing settings for S-1 data.

Speckle Filter Method	Lee Sigma
Speckle Filter Window Size	3x3
Speckle Filter Number of Images	5
Terrain Flattening Shadow Buffer	0m
Spatial Resolution	10m
GLCM Window Size	5x5

3.2. AGB Estimation Without Height Metrics

The model using S-2 predictors produced superior results to the model using S-1 predictors (Table 8). Combining both S-1 and S-2 data yielded higher accuracy of AGB estimation than when using either platform on its own. The inclusion of topographic predictor variables produced slightly more reliable results for AGB prediction and a smaller set of best predictors, however none of them made up the set of best predictors.

Table 8. Results for AGB estimation accuracy for models excluding height metrics.

	S-1	S-2	S-1 + S-2	S-2 + S-1 + Topographic
RMSE	75.665	67.756	64.751	63.721
RMSE%	55.94%	50.09%	47.58%	46.83%
R²	0.288	0.399	0.491	0.503
MAE	59.728	52.344	52.020	51.370
Best Predictor Set	sarSAI_Jan, GLCM_VV_contrast_Jan, GLCM_VV_dvar_Jan, GLCM_VV_dent_Jan, GLCM_VV_diss_Jan, GLCM_VV_shade_Jan, VH_Feb, sarMI_Feb, sarRI2_Feb, sarRDI_Mar, sarRI1_Mar, GLCM_VH_asm_Apr, GLCM_VV_corr_Jun, GLCM_VV_shade_Jun, VV_Sep, sarRI1_Sep, sarSqDI_Sep, GLCM_VV_corr_Sep, GLCM_VV_savg_Sep, GLCM_VH_asm_Oct, GLCM_VV_shade_Dec	B5_May, CAB_May, B3_Jun, B5_Jun, B12_Jun, CAB_Jun, TCARI_Jun, B3_Jul, B5_Jul, TCARI_Jul	B5_May, CAB_May, B2_Jun, B3_Jun, B4_Jun, B5_Jun, B6_Jun, B8_Jun, B11_Jun, B12_Jun, CAB_Jun, RERDVI_Jun, TCARI_Jun, B2_Jul, B3_Jul, B5_Jul, B6_Jul, CAB_Jul, MCARI_Jul, TCARI_Jul, B5_Aug, MCARI_Aug, NDI45_Aug, VH_Jan, sarMI_Jan, sarRI2_Jan, GLCM_VV_contrast_Jan, GLCM_VV_dvar_Jan, GLCM_VV_dent_Jan, GLCM_VV_diss_Jan, GLCM_VV_shade_Jan, VH_Feb, sarMI_Feb, sarRI2_Feb, VV_Mar, sarRI1_Mar, VH_Apr, GLCM_VV_shade_Jun, VV_Sep, sarMI_Sep, sarRI2_Sep, sarSqDI_Sep, GLCM_VH_corr_Sep	CAB_May, B3_Jun, B5_Jun, B12_Jun, TCARI_Jun, B3_Jul, B5_Jul, TCARI_Jul, MCARI_Aug, sarRI2_Jan, GLCM_VV_contrast_Jan, GLCM_VV_dvar_Jan, VH_Feb, sarMI_Feb, sarRI2_Feb, sarRI1_Mar, sarSqDI_Sep

The optimisation process of the RF model is illustrated in Figure 5, refining the number of predictors from 667 to 17.
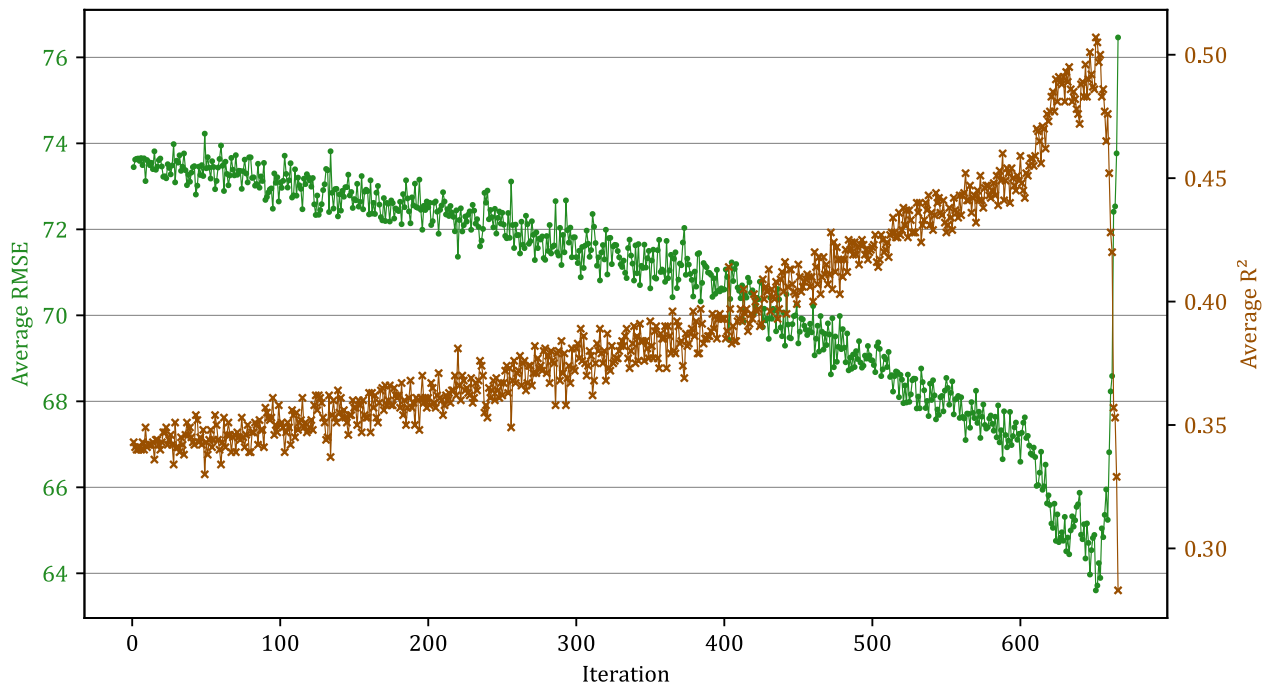

Figure 5. RF model training to determine the best predictor set for AGB estimation using S-1, S-2, and topographic variables. The lowest RMSE value determines the best predictor set.

3.3. Tree Height Data

Analysing the log-transformed relationship between MH and AGB for the inventory plot data produced a Pearson correlation coefficient of 0.821 ($p < 0.001$) and a coefficient of determination (R^2) value of 0.675, of indicating a strong and statistically significant relationship (Figure 6).
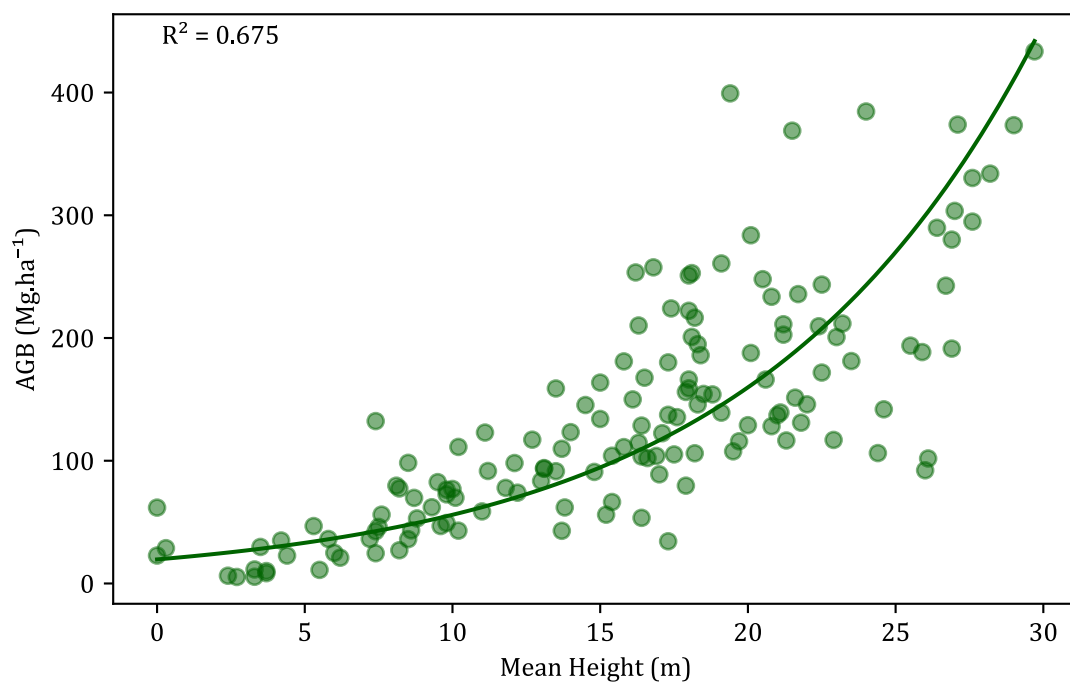
Figure 6. Relationship between MH and AGB for Swedish NFI plot data.

3.3.1. Mean Height Estimation

Using the random forest model to predict MH at plot level from S-1, S-2, and Topographic predictors produced results of RMSE: 4.716, RMSE%: 30.18%, R^2: 0.522, MAE: 3.845. The scatterplot of estimated vs. observed MH reveals that the model tends to underestimate the largest height values and overestimate the smallest (Figure 7).
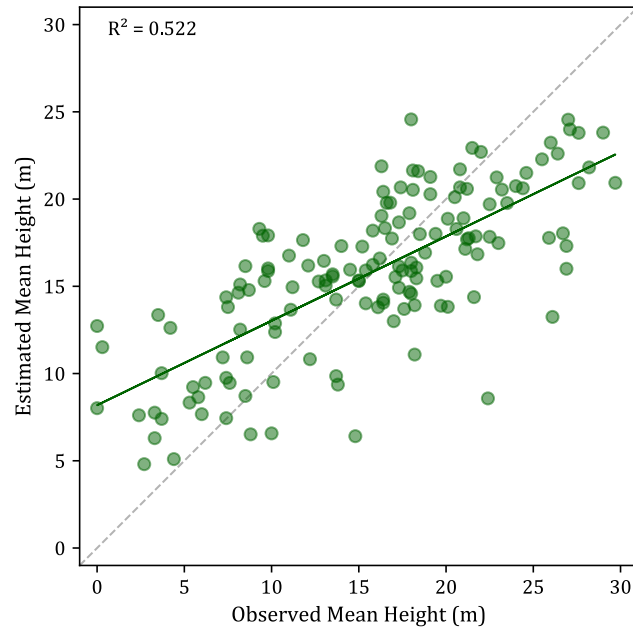

Figure 7. Scatterplot of estimated vs observed MH.

When examining the distribution and importance of the predictors, it can be seen that optical (S-2) predictors from June and July are dominant, with band 5 (B5), band 3 (B3), and the TCARI index being the most important (Figure 8). Overall, S-2 predictors account for almost two-thirds of the importance score (0.664). A selection of backscatter, indices, and texture metrics make up the S-1 predictor subset. Nine are from winter or early spring (January – March), while two are from autumn (October). The remaining two are from early in the peak growth period (June).
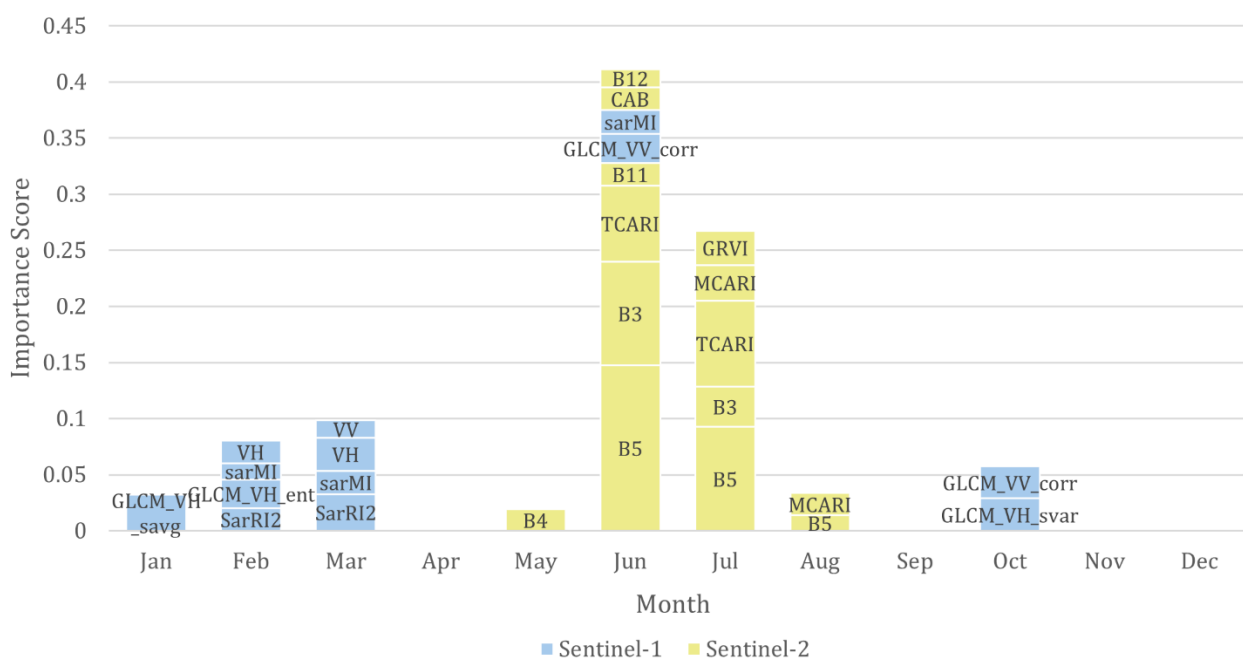

Figure 8. Distribution and relative importance of predictors by month and platform for MH estimation.

3.4. AGB Estimation with Height Data

The inclusion of height data leads to improved estimation of AGB for all height datasets (Table 9). The global canopy height dataset (Global_CH) provides only a small improvement, while the model using LiDAR produces the greatest improvement. For the model that includes LiDAR data, all but one of the LiDAR metrics - kurtosis (lidar_kur) - make up the best predictor set, with 28 in total. The optimal predictor set for the model using Estimated_MH uses 12 predictors to produce the lowest RMSE value.

Table 9. Results for AGB estimation using height metrics.

	S-2 + S-1 + Topo + Global_CH	S-2 + S-1 + Topo + LiDAR	S-2 + S-1 + Topo + Estimated_MH
RMSE	63.118	50.895	57.159
RMSE%	46.38%	37.40%	42.00%
R^2	0.511	0.675	0.598
MAE	51.030	37.769	43.199
Best Predictor Set	CAB_May, B3_Jun, B5_Jun, B12_Jun, TCARI_Jun, B3_Jul, B5_Jul, TCARI_Jul, MCARI_Aug, sarRI2_Jan, GLCM_VV_contrast_Jan, VH_Feb, sarMI_Feb, sarRI2_Feb, sarSqDI_Sep, Global_CH	CAB_May, B3_Jun, B5_Jun, TCARI_Jun, B3_Jul, B5_Jul, MCARI_Aug, NDI45_Aug, GLCM_VV_dent_Jan, GLCM_VV_diss_Jan, GLCM_VH_ent_Apr, GLCM_VH_shade_Jul, GLCM_VV_contrast_Sep, GLCM_VH_asm_Dec, lidar_avg, lidar_max, lidar_min, lidar_p01, lidar_p05, lidar_p10, lidar_p25, lidar_p50, lidar_p75, lidar_p90, lidar_p95, lidar_p99, lidar_ske, lidar_std	CAB_May, B3_Jun, B5_Jun, TCARI_Jun, B3_Jul, B5_Jul, TCARI_Jul, MCARI_Aug, GLCM_VV_dent_Jan, sarRI2_Feb, sarSqDI_Sep, Estimated_MH

The model incorporating LiDAR data tends to underestimate the AGB values above 200 Mg.ha^{-1} (Figure 9a). This phenomenon is even more pronounced for the model that incorporates Estimated_MH (Figure 9b).
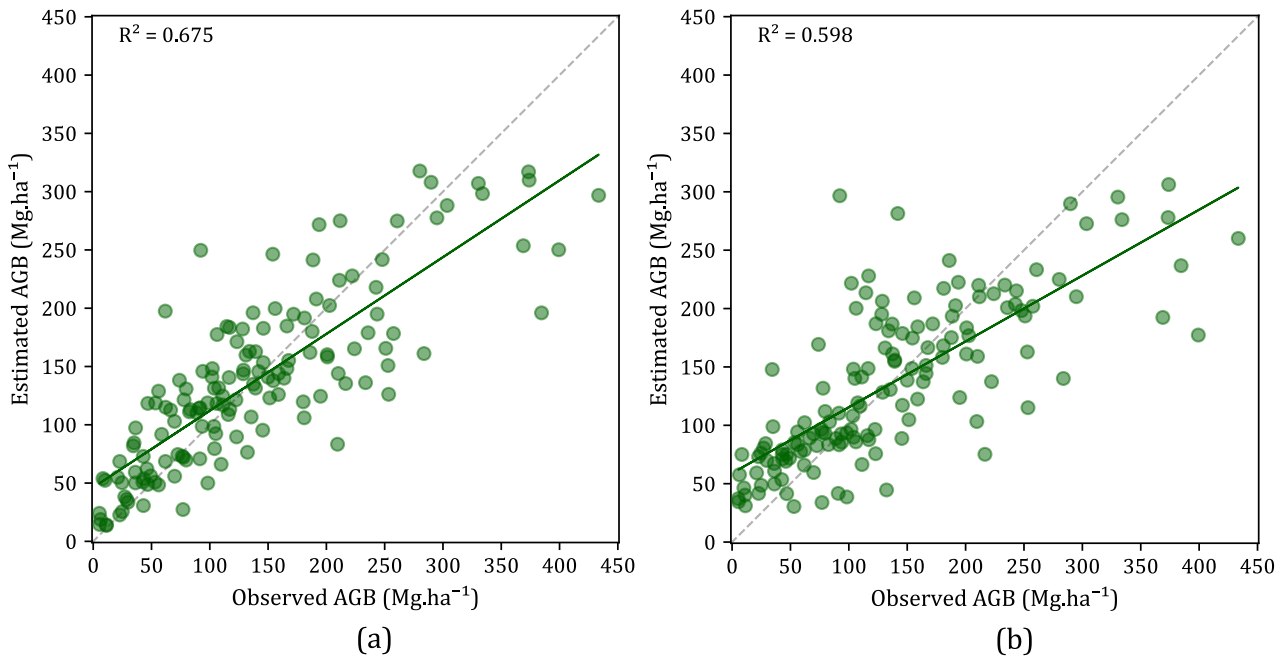

Figure 9. a) Estimated vs. observed AGB for model using S-2 + S-1 + Topo + LiDAR predictors. b) Estimated vs. observed AGB for model using S-2 + S-1 + Topo + Estimated_MH predictors.

For the model chosen to map AGB, the Estimated_MH variable is the most important predictor, accounting for over half of the total importance score (0.545) – see Figure 10. The rest of the predictors are dominated by those derived from the S-2 platform (0.343), with S-1 accounting for the lowest

portion (0.112). Two of the S-1 predictors are from winter (January and February), while the other is from mid-autumn (September).
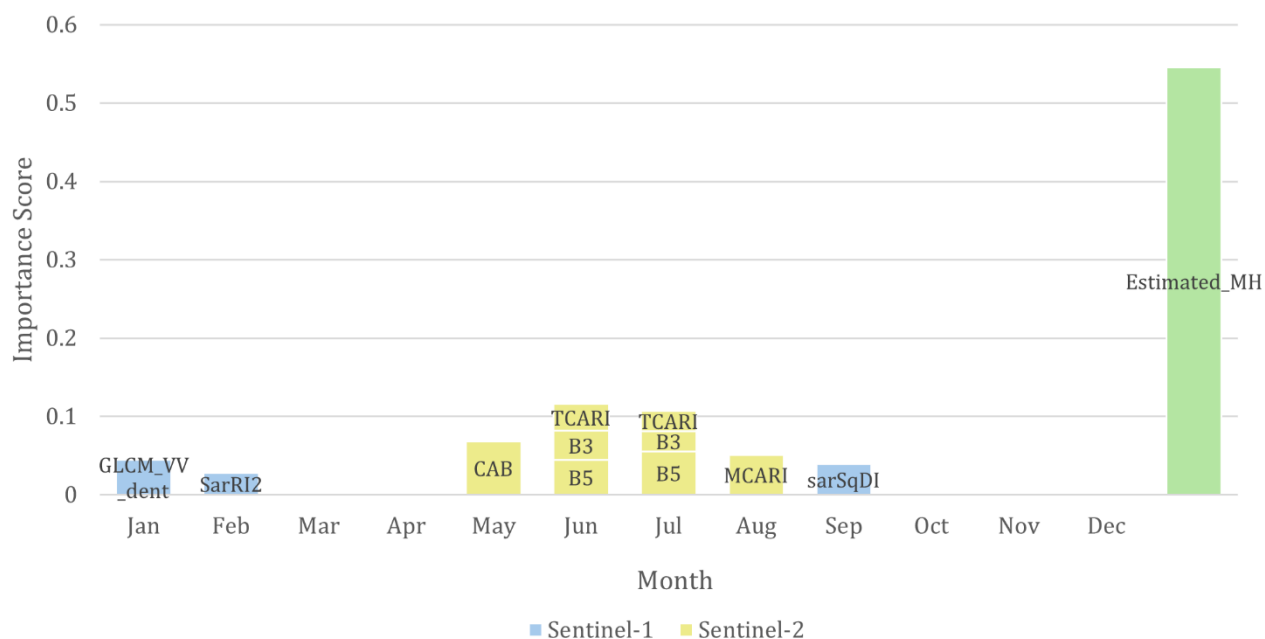

Figure 10. Distribution and relative importance of predictors by month and platform for AGB prediction for the model incorporating Estimated_MH.

Examining the predictions of the model by dominant tree type at plot level reveals that the correlation for the coniferous dominated plots ($R^2 = 0.601$) closely matches the overall correlation (Figure 11a). The deciduous plots show the weakest correlation (Figure 11b), while the mixed plots show the strongest (Figure 11c). The underestimation of values above 200 Mg.ha^{-1} is evident in all plot types.
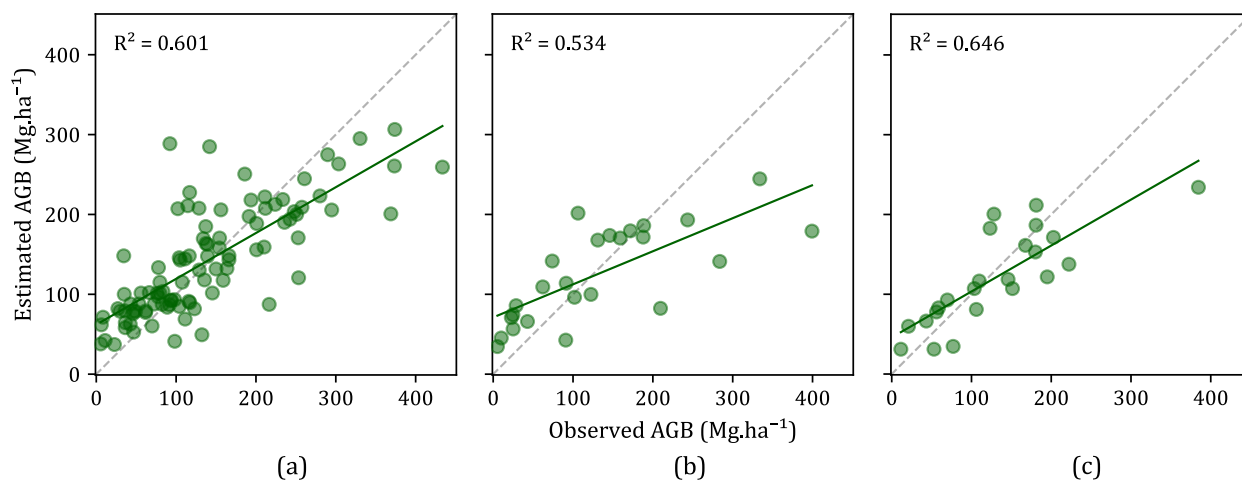

Figure 11. Estimated vs. observed AGB for a) coniferous plots, b) deciduous plots, c) mixed plots.

3.5. AGB Mapping

An AGB map of the study area was produced at 10m resolution using 12 predictor variables (Figure 12). The values of AGB ranged from 36.96 Mg.ha^{-1} to 367.44 Mg.ha^{-1}, with a mean value of 253.73 Mg.ha^{-1}. The western coast of the county can be seen to have low to moderate biomass, while interior highland areas tend to have higher biomass.
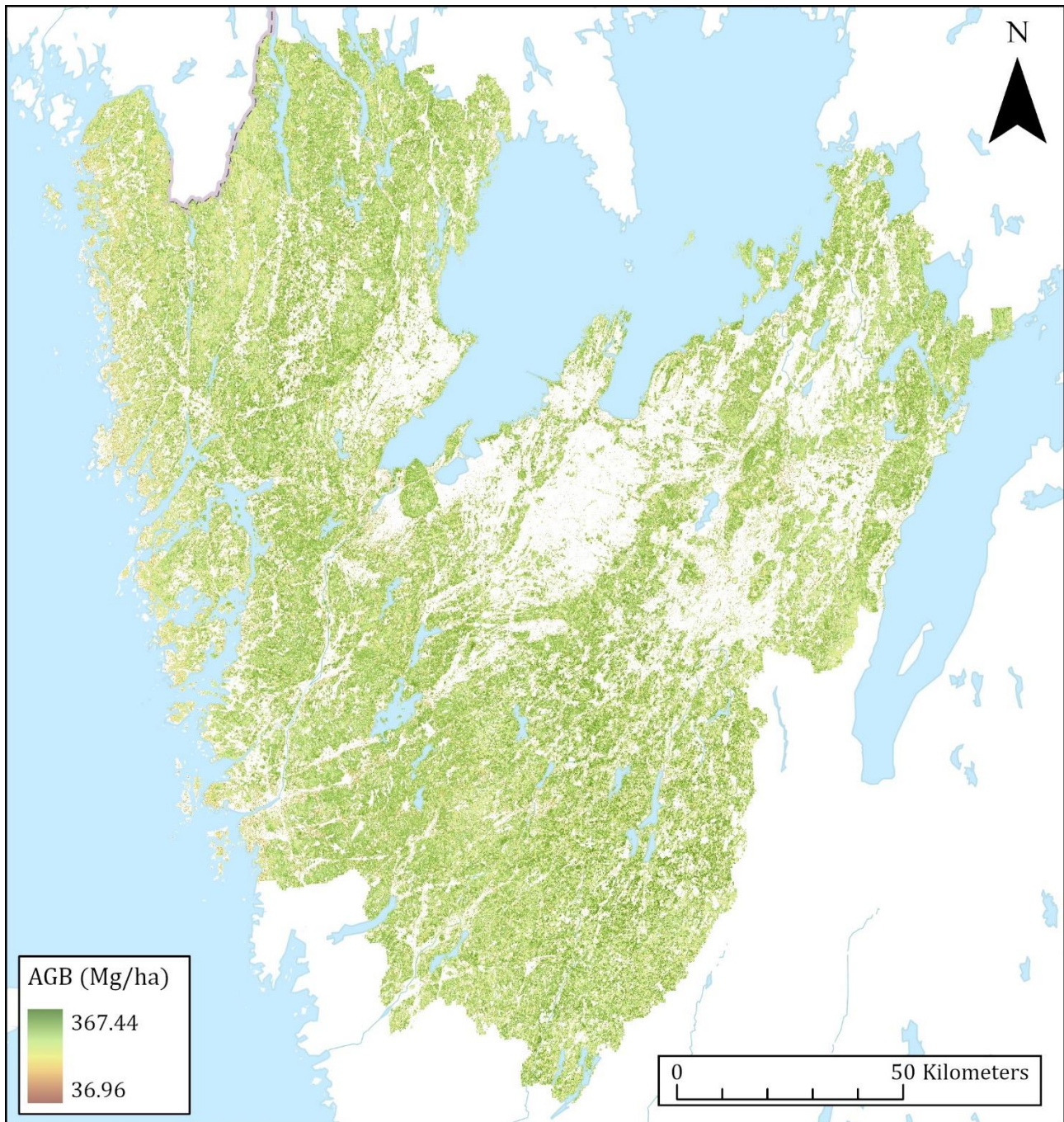
Figure 12. Spatial distribution of AGB across Västra Götaland.

3.5.1.Uncertainty Map of AGB

The examination of uncertainty in biomass estimation revealed an average of 98.71 Mg.ha^{-1}, with the standard deviation (SD) values ranging from 43.24 Mg.ha^{-1} to 136.23 Mg.ha^{-1} (Figure 13). The mean uncertainty constitutes approximately 38.90 % of the mean predicted biomass. Comparison of the two maps reveals higher AGB areas having higher uncertainty, while conversely lower uncertainty is seen in lower AGB areas such as the western coast.
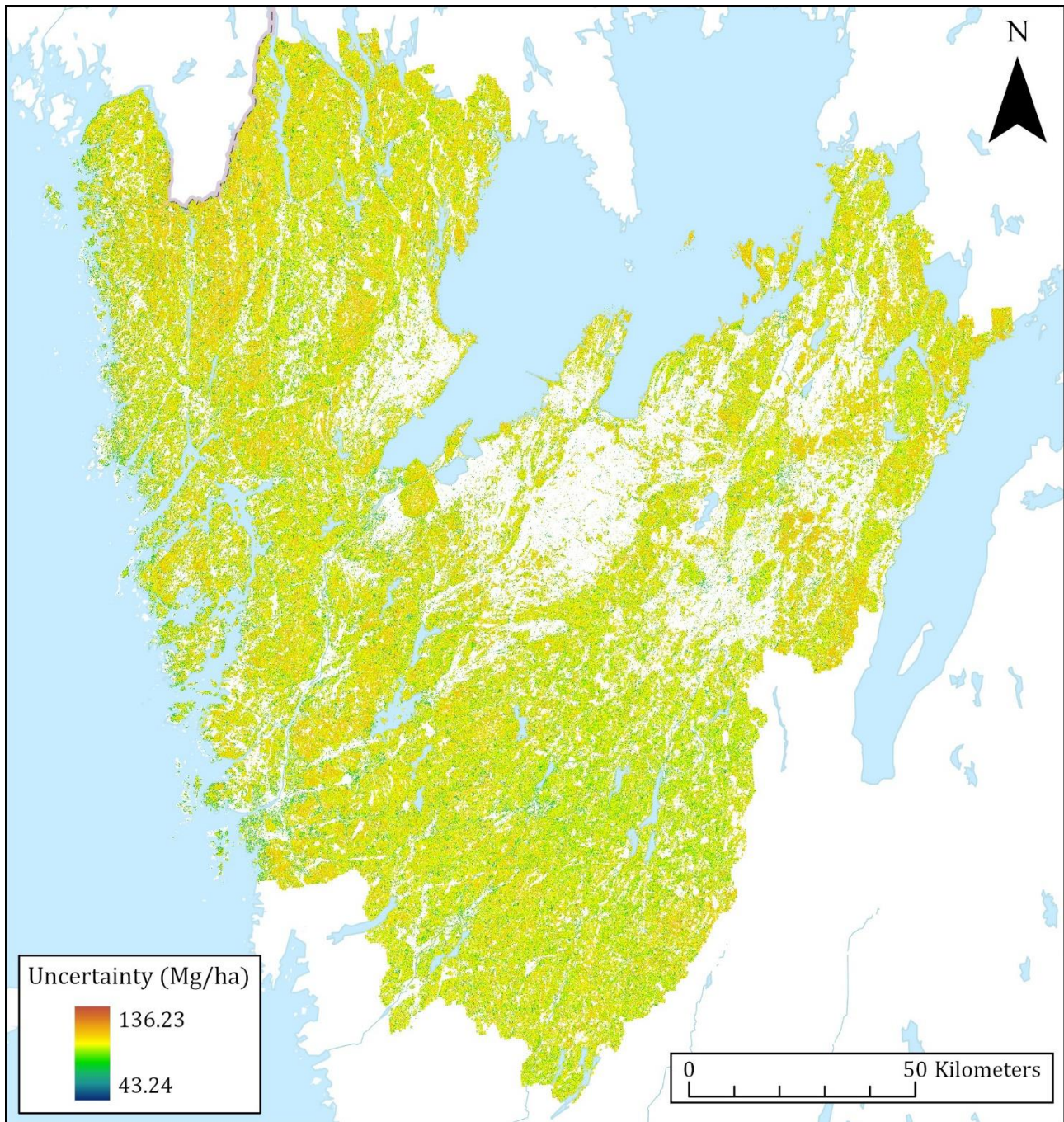
Figure 13. Standard Deviation of predicted AGB representing uncertainty

4. Discussion

Pre-processing Settings for S-2 and S-1 data

The advantage of resampling all S-2 bands to 10m spatial resolution for AGB estimation is likely due to retaining the native resolution of the 10m bands – of which B3 (Green) and B4 (Red) were shown to be important predictors. A finer resolution reduces the incidence of mixed pixels and improves the delineation of small-scale features in forested landscapes such as canopy gaps and heterogenous tree types. Spatial resolution of pixels is of particular relevance given the small size of the circular forest inventory plots. Additionally, matching the resolution of the 10m ESA Worldcover dataset allows for more accurate definition of forest edges. Bilinear resampling proved superior to the bicubic method,

likely due to its ability to smoothly interpolate pixel values, which better preserves the integrity and continuity of forest data while also reducing noise.

The optimal S-1 settings collectively suggest an approach that balances noise reduction, spatial resolution, and the extraction of detailed texture and structure information from SAR data. The Lee Sigma filter balances the reduction of speckle noise against the preservation of crucial forest structure details. A speckle filter window size of 3x3 and the use of five images for filtering effectively minimised noise while preserving spatial and temporal details. Opting for a 0m shadow buffer suggests a minimal influence of topography on the SAR signal. The selected 10m spatial resolution was found to be optimal for capturing forest structural characteristics pertinent to biomass, such as canopy gaps and stand density. Finally, the choice of a 5x5 GLCM window size allowed for the extraction of detailed texture features indicative of spatial variations in AGB density.

AGB Estimation Without Height Metrics

The model leveraging only S-2 predictors outperforms the S-1 exclusive model in terms of predictive accuracy, yet neither reaches a satisfactory level of performance. This outcome, particularly the limited effectiveness of S-1, aligns with expectations due to the known saturation limitations of both platforms. [51,66,67]. By integrating S-1 and S-2 data, however the model achieves an enhanced predictive accuracy, surpassing the capabilities of using either dataset in isolation. This improvement through the combined use of S-1 and S-2 data corroborates similar findings from previous research, underscoring the benefits of data fusion in enhancing model accuracy [63,180].

The integration of topographic data with S-1 and S-2 datasets led to a refined selection of predictors and a slight improvement in predictive accuracy for AGB estimation. While topographic variables themselves were not seen in the best predictor set, their inclusion likely contributed to a more streamlined set of predictors by introducing a layer of ecological and physical information that overlaps with signals already captured by the remote sensing predictors. This outcome suggests a complex interplay between predictors in the RF algorithm to achieve a more accurate and efficient AGB estimation model.

Mean Height Estimation

The establishment of a strong and statistically significant relationship between MH and AGB within the study plots reaffirms the crucial link between tree height and AGB [29,35,39]. The moderate accuracy of the RF model when predicting MH reveals limitations in the approach used, primarily a result of the model initially being designed for AGB estimation and subsequently adapted for this purpose. Much of the error in estimates appears to stem from the tendency to underestimate high values. Coupled with overestimation of the lowest height values, it indicates the model's bias towards the mean and reveals an inadequacy in capturing the extremes of the height distribution. This tendency may be attributable to the training set being unequally distributed, predominantly mid-range values.

The dominance of S-2 over S-1 predictors when estimating MH matches the findings of Silveira et al. [181] using comparable data and methods. Of these, the prevalence of predictors from June and July suggest that optical data collected early-to-mid summer is the most reliable indicator of forest structure in this region. The importance of the green (B3), and near-infrared (B5) bands, and the TCARI index can be attributed to their direct relationship with vegetative characteristics that are indicative of forest structure. The green band (B3) captures chlorophyll absorption, offering insights into the health and density of the canopy. Meanwhile, the near-infrared band (B5) is highly responsive to the structure of the vegetation canopy, as NIR light is strongly reflected by healthy vegetation. This makes B5 particularly useful in distinguishing between vegetative densities and heights, as taller, denser canopies reflect more

NIR light. The TCARI index, which leverages both these bands, is designed to minimise the influence of soil brightness and enhance the sensitivity to chlorophyll concentration in leaves, providing a robust indicator of vegetative vigour and density. Future research should explore the application of indices utilising the green (B3) and near-infrared (B5) bands captured during early summer. Such indices could provide a more detailed understanding of vegetative dynamics in hemi-boreal ecosystems.

The diverse array of backscatter, indices, and texture metrics from S-1 data in the subset of predictors, and the fairly equal magnitude of their importance, underscores the complexity in capturing forest structure. The S-1 backscatter metrics, particularly VH, provide insights into forest canopy and understorey interactions due to their sensitivity to the structural components of the forest. The indices enrich this information by exploiting specific relationships between the polarisations. The texture metrics derived from the Gray-Level Co-occurrence Matrix (GLCM) enhance the model's capacity to interpret the spatial complexity and heterogeneity of forest stands, offering details beyond what is possible with backscatter and indices alone. This blend of backscatter and texture variables from S-1 data enables a more comprehensive analysis of forest structure, enhancing the model's ability to estimate forest MH. The importance of winter and early spring information in leveraging SAR data for forest height estimation aligns with findings by Olesk et al. [182] in a similar forest environment. These findings underscore the importance of seasonality and a variety of data products in capturing the structural detail of forests.

AGB Estimation with Height Data

The minimal improvement in Above Ground Biomass (AGB) estimation accuracy observed when incorporating the global canopy height dataset as a predictor may reflect limitations inherent to the dataset's production methodology. Notably, the dataset leverages canopy height values from the Global Ecosystem Dynamics Investigation (GEDI) as reference data. However, GEDI's coverage is limited, extending only up to 51.6 degrees north. Consequently, for areas beyond this latitude, the dataset relies on extrapolations from S-2 imagery alone. This reliance on optical data for extrapolation might introduce inaccuracies in canopy height predictions, particularly given the distinct ecological and structural characteristics of vegetation types at higher latitudes. These potential inaccuracies could thus limit the effectiveness of the global canopy height dataset for enhancing AGB estimation for the study area. Additionally, the discrepancy in temporal coverage between the canopy height dataset, derived from 2020 data, and the field inventory data collected in 2021 presents another potential source of error in accurately estimating AGB. This time gap could have implications where forest management activities such as logging have altered the inventory plot sites in this period. For such plots, this results in canopy height values representing an established forest, while the inventory and data gathered from the S-1 and S-2 platforms reflects that of a recently cut stand.

This time gap phenomena also likely affects the AGB estimation results of the model incorporating LiDAR - for some plots, LiDAR data is over two years out of date. The accuracy results for biomass estimation are towards the lower end of the range of what can be expected from airborne LiDAR [183]. For future studies in this region, methods for adjusting or normalising temporally mis-matched LiDAR should be considered. This could be done using by scaling LiDAR data according to observed changes in height from field inventory plots, using time-series spaceborne LiDAR [184], or a combination of forest removal maps and remote sensing data [95].

Despite the moderate accuracy in the Estimated_MH values, the model incorporating this metric shows fairly similar results to the model incorporating the LiDAR metrics. This suggests that temporal alignment is an important factor for structural predictor variables. The model's performance is also somewhat comparable to other models using multiple sources of data with an RF to model AGB [46,185].

For the model incorporating Estimated_MH that was used to map AGB across the study area, the tendency to underestimate AGB values above 200 Mg.ha^{-1} is attributable to several compounding factors. The first being the inherent saturation limits of optical (~200 Mg.ha^{-1}) and C-band SAR (~100 Mg.ha^{-1}) platforms [51,67,158]. This echoes the findings of Agata et al. [186], who observed a similar saturation point in their study using S-1 and S-2 data for AGB estimation in a temperate European forest. Second, similar to the MH model, the AGB model is trained on a dataset that is unequally distributed and predominantly features mid-range values. The importance of the Estimated_MH predictor, which tends to be underestimated in high-range values, likely exacerbates this bias. Huang et al. [185] found height measurement inaccuracies to be a major source of uncertainty in AGB estimation. Finally, the inclusion of dead biomass in plots with a height of zero would likely bias the training of the model further towards the lower end of the biomass range. These findings highlight a limitation of using height as a predictor variable and underline the need for more accurate techniques to predict height values.

Nevertheless, the Estimated_MH predictor is still of primary importance. This is attributable to two key factors: firstly, its ability to consolidate essential vegetation structure and biomass-related information - which results in a streamlined predictor set compared to that for AGB estimation without height data. Secondly, the superiority of height metrics over optical and radar remote sensing predictors in estimating AGB, as documented in other studies [171]. Concerning the S-1 predictors, the finding that two of the three are from the winter period compares with the findings of Laurin et al. [152], who attributed their findings to the loss of leaves on broadleaf deciduous trees. Consequently, the low proportion of deciduous plots could explain the overall low contribution of S-1 predictors and suggest the methods employed using C-band SAR data are unsuited to directly estimating AGB in this region. Other studies have shown that lower frequency interferometric SAR imagery may be more suitable for northern forests [108,109].

The model's overall accuracy aligns closely with the accuracy observed in coniferous plots, which constitute two-thirds of all samples. Given the lower accuracy recorded for deciduous plots compared to coniferous ones, it is reasonable to expect that the accuracy for mixed plots falls somewhere between these two extremes. However, the unexpected result of the highest accuracy for AGB estimation in mixed plots can be attributed to the presence of only one plot with an observed AGB value exceeding 250 Mg.ha^{-1}, which led to a lower average residual, particularly as the model tends to underestimate values above 200 Mg.ha^{-1}.

AGB Mapping

The maximum values from the AGB map exceed those of the largest estimated values for the training plots but fall short of the largest observed values. This is expected given the model's tendency to underestimate high AGB values and suggests that the training dataset may not be fully capturing the spatial variability or range of conditions present across the entire region.

Quantifying uncertainty is an important component of evaluating model predictive performance. The variability in SD values indicates a substantial range of uncertainty in the AGB estimates, reflecting the complexities and challenges in precise biomass prediction in this environment. The average uncertainty accounts for nearly 38.90% of the mean predicted AGB, indicating a substantial proportion of the predicted values are encompassed by uncertainty. The tendency of the model to underestimate high AGB values is evident in the spatial distribution, with the densest forest areas having the highest uncertainty. The high uncertainty of the predictions may be partly explained by the choice of 10m spatial resolution. According to Ehlers et al. [46], AGB mapping accuracy over large areas benefits from a coarser spatial resolution due to the vegetation cover effect.

It is evident that the incorporation of time-matched height metrics into the predictor dataset is a viable method of enhancing AGB estimation accuracy. However, to reduce uncertainty in AGB predictions, more effective methods for height estimation need to be explored. Effective techniques using comparable data have been demonstrated [88,187,188], while interferometric SAR data based methods have also shown promise [38–40] and should be considered when undertaking such work in the future.

Enhanced modelling techniques are required to overcome the underestimation of the largest AGB values. This may include efforts to estimate proxies such as basal area or stem volume, which could provide ways to circumvent the limitations imposed by signal saturation [43,155,189,190]. Sampling design is also an important consideration, and so the issue of proportionate representation across the distribution of training data samples for the area being mapped should also be investigated [178,191]. The differences in predictive accuracy for different tree types also warrants further investigation. The seasonality of deciduous trees [152], and inter-species differences in tree structure and allometric relationships [27,192] suggest that methods to classify tree type before estimation and then apply species specific models may improve the outcomes.

5. Conclusion

This study addressed the complex challenge of estimating forest AGB in Sweden's hemi-boreal forests through the processing capabilities of GEE, multisource remote sensing data, and machine learning regression techniques. It emphasised the importance of optimising predictors and incorporating tree height metrics for improving AGB estimation accuracy. Building upon previous findings, this research highlighted the value of combining optical and SAR remote sensing data to overcome individual sensor limitations such as signal saturation in dense canopy environments. Despite the data fusion technique employed, this research observed suboptimal performance at higher AGB values, reaffirming the inherent limitations of current sensor technology in such conditions.

In spite of moderate model performance, a valuable contribution of this research is the implementation of GEE and a machine learning approach that capitalises on multivariate and multitemporal data. The feature selection process in the RF algorithm, which was key to enhancing model performance, allowing large datasets to be optimised to improve AGB estimation accuracy. This approach also underscored the critical role of seasonality in the timing of data collection, which is essential for capturing accurate forest structure and dynamics at different periods of the vegetative cycle.

Accurate height data emerged as a vital component for refining AGB estimation. The precision of height measurements strongly influences the accuracy of AGB predictions, underlining the need for reliable height information from remote sensing technologies. The application of the RF model showcases the potential of machine learning to enhance AGB estimation models to achieve higher accuracy by effectively leveraging crucial predictors.

These insights lay the groundwork for a scalable and cost-effective method to estimate AGB and monitor forest biomass dynamics over time. Additionally, this study highlights the importance of data preprocessing to enhance the predictive performance of remote sensing-based AGB estimation models.

However, the challenges of AGB underestimation in high-density areas and heterogeneity of tree types underscore the need for ongoing advancements in remote sensing technology and data processing techniques. Addressing these challenges is crucial to enhance the accuracy and reliability of forest AGB estimation. This research contributes knowledge to the fields of forestry and climate change mitigation, offering practical insights for policymakers, forest managers, and researchers. Future work should continue to focus on addressing these challenges to further improve forest AGB estimation.

Source code and data available via Github: https://github.com/RM-Ulster/MSc_Project/tree/Report-Submission

References

1. Kindermann, G.E.; McCallum, I.; Fritz, S.; Obersteiner, M. A Global Forest Growing Stock, Biomass and Carbon Map Based on FAO Statistics. *Silva Fennica* **2008**, *42*.
2. Dixon, R.K.; Solomon, A.M.; Brown, S.; Houghton, R.A.; Trexier, M.C.; Wisniewski, J. Carbon Pools and Flux of Global Forest Ecosystems. *Science* **1994**, *263*, 185–190, doi:10.1126/science.263.5144.185.
3. Clemmensen, K.E.; Finlay, R.D.; Dahlberg, A.; Stenlid, J.; Wardle, D.A.; Lindahl, B.D. Carbon Sequestration Is Related to Mycorrhizal Fungal Community Shifts during Long-Term Succession in Boreal Forests. *New Phytologist* **2015**, *205*, 1525–1536, doi:10.1111/nph.13208.
4. Pan, Y.; Birdsey, R.A.; Fang, J.; Houghton, R.; Kauppi, P.E.; Kurz, W.A.; Phillips, O.L.; Shvidenko, A.; Lewis, S.L.; Canadell, J.G.; et al. A Large and Persistent Carbon Sink in the World's Forests. *Science* **2011**, *333*, 988–993, doi:10.1126/science.1201609.
5. IPCC *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; 1st ed.; Cambridge University Press, 2023; ISBN 978-1-00-932584-4.
6. Raiesi, F.; Beheshti, A. Evaluating Forest Soil Quality after Deforestation and Loss of Ecosystem Services Using Network Analysis and Factor Analysis Techniques. *CATENA* **2022**, *208*, 105778, doi:10.1016/j.catena.2021.105778.
7. Herold, M.; Carter, S.; Avitabile, V.; Espejo, A.B.; Jonckheere, I.; Lucas, R.; McRoberts, R.E.; Næsset, E.; Nightingale, J.; Petersen, R.; et al. The Role and Need for Space-Based Forest Biomass-Related Measurements in Environmental Management and Policy. *Surv Geophys* **2019**, *40*, 757–778, doi:10.1007/s10712-019-09510-6.
8. Lorenz, R.; Pitman, A.J. Effect of Land-Atmosphere Coupling Strength on Impacts from Amazonian Deforestation. *Geophysical Research Letters* **2014**, *41*, 5987–5995, doi:10.1002/2014GL061017.
9. Wright, S.J. Tropical Forests in a Changing Environment. *Trends in Ecology & Evolution* **2005**, *20*, 553–560, doi:10.1016/j.tree.2005.07.009.
10. Phillips, H.R.P.; Newbold, T.; Purvis, A. Land-Use Effects on Local Biodiversity in Tropical Forests Vary between Continents. *Biodivers Conserv* **2017**, *26*, 2251–2270, doi:10.1007/s10531-017-1356-2.
11. Pearson, T.R.H.; Brown, S.; Murray, L.; Sidman, G. Greenhouse Gas Emissions from Tropical Forest Degradation: An Underestimated Source. *Carbon Balance Manage* **2017**, *12*, 3, doi:10.1186/s13021-017-0072-2.
12. Main-Knorn, M.; Moisen, G.G.; Healey, S.P.; Keeton, W.S.; Freeman, E.A.; Hostert, P. Evaluating the Remote Sensing and Inventory-Based Estimation of Biomass in the Western Carpathians. *Remote Sensing* **2011**, *3*, 1427–1446, doi:10.3390/rs3071427.
13. Nesha, K.; Herold, M.; De Sy, V.; de Bruin, S.; Araza, A.; Málaga, N.; Gamarra, J.G.P.; Hergoualc'h, K.; Pekkarinen, A.; Ramirez, C.; et al. Exploring Characteristics of National Forest Inventories for Integration with Global Space-Based Forest Biomass Data. *Science of The Total Environment* **2022**, *850*, 157788, doi:10.1016/j.scitotenv.2022.157788.
14. FAO *Global Forest Resources Assessment 2020*; FAO, 2020; ISBN 978-92-5-132974-0.
15. de Lima, R.A.F.; Oliveira, A.A.; Pitta, G.R.; de Gasper, A.L.; Vibrans, A.C.; Chave, J.; ter Steege, H.; Prado, P.I. The Erosion of Biodiversity and Biomass in the Atlantic Forest Biodiversity Hotspot. *Nat Commun* **2020**, *11*, 6347, doi:10.1038/s41467-020-20217-w.
16. IPBES Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. **2019**, doi:10.5281/zenodo.6417333.
17. Tilman, D.; Isbell, F.; Cowles, J.M. Biodiversity and Ecosystem Functioning. *Annu. Rev. Ecol. Evol. Syst.* **2014**, *45*, 471–493, doi:10.1146/annurev-ecolsys-120213-091917.

18. Hill, S.L.L.; Arnell, A.; Maney, C.; Butchart, S.H.M.; Hilton-Taylor, C.; Ciciarelli, C.; Davis, C.; Dinerstein, E.; Purvis, A.; Burgess, N.D. Measuring Forest Biodiversity Status and Changes Globally. *Frontiers in Forests and Global Change* **2019**, *2*.

19. Angelsen, A.; Jagger, P.; Babigumira, R.; Belcher, B.; Hogarth, N.J.; Bauch, S.; Börner, J.; Smith-Hall, C.; Wunder, S. Environmental Income and Rural Livelihoods: A Global-Comparative Analysis. *World Development* **2014**, *64*, S12–S28, doi:10.1016/j.worlddev.2014.03.006.

20. Zhang, D.; Pearse, P.H. *Forest Economics*; UBCPress: Vancouver; Toronto, 2011; ISBN 978-0-7748-2153-7.

21. Lewis, J.L.; Sheppard, S.R.J. Ancient Values, New Challenges: Indigenous Spiritual Perceptions of Landscapes and Forest Management. *Society & Natural Resources* **2005**, *18*, 907–920, doi:10.1080/08941920500205533.

22. Kumar, L.; Mutanga, O. Remote Sensing of Above-Ground Biomass. *Remote Sensing* **2017**, *9*, 935, doi:10.3390/rs9090935.

23. Raihan, A.; Ara Begum, R.; Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia; Mohd Said, M.N.; Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia A Meta-Analysis of the Economic Value of Forest Carbon Stock. *GEOGRAFIA* **2021**, *17*, doi:10.17576/geo-2021-1704-22.

24. United Nations United Nations Framework Convention on Climate Change 1992.

25. UNFCCC *Key Decisions Relevant for Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD+) - Decision Booklet REDD+*; 2016;

26. Montès, N.; Gauquelin, T.; Badri, W.; Bertaudière, V.; Zaoui, E.H. A Non-Destructive Method for Estimating above-Ground Forest Biomass in Threatened Woodlands. *Forest Ecology and Management* **2000**, *130*, 37–46, doi:10.1016/S0378-1127(99)00188-7.

27. Berner, L.T.; Alexander, H.D.; Loranty, M.M.; Ganzlin, P.; Mack, M.C.; Davydov, S.P.; Goetz, S.J. Biomass Allometry for Alder, Dwarf Birch, and Willow in Boreal Forest and Tundra Ecosystems of Far Northeastern Siberia and North-Central Alaska. *Forest Ecology and Management* **2015**, *337*, 110–118, doi:10.1016/j.foreco.2014.10.027.

28. Basuki, T.M.; van Laake, P.E.; Skidmore, A.K.; Hussin, Y.A. Allometric Equations for Estimating the Above-Ground Biomass in Tropical Lowland *Dipterocarp* Forests. *Forest Ecology and Management* **2009**, *257*, 1684–1694, doi:10.1016/j.foreco.2009.01.027.

29. Xiao, C.-W.; Ceulemans, R. Allometric Relationships for Below- and Aboveground Biomass of Young Scots Pines. *Forest Ecology and Management* **2004**, *203*, 177–186, doi:10.1016/j.foreco.2004.07.062.

30. Réjou-Méchain, M.; Barbier, N.; Couteron, P.; Ploton, P.; Vincent, G.; Herold, M.; Mermoz, S.; Saatchi, S.; Chave, J.; de Boissieu, F.; et al. Upscaling Forest Biomass from Field to Satellite Measurements: Sources of Errors and Ways to Reduce Them. *Surv Geophys* **2019**, *40*, 881–911, doi:10.1007/s10712-019-09532-0.

31. Dube, T.; Onisimo, M.; Cletah, S.; Adelabu, S.; Tsitsi, B. Remote Sensing of Aboveground Forest Biomass: A Review. *Tropical Ecology* **2016**, *57*, 125–132.

32. Persson, H.J.; Ståhl, G. Characterizing Uncertainty in Forest Remote Sensing Studies. *Remote Sensing* **2020**, *12*, 505, doi:10.3390/rs12030505.

33. Rodríguez-Veiga, P.; Wheeler, J.; Louis, V.; Tansey, K.; Balzter, H. Quantifying Forest Biomass Carbon Stocks From Space. *Curr Forestry Rep* **2017**, *3*, 1–18, doi:10.1007/s40725-017-0052-5.

34. Mitchard, E.T.A.; Feldpausch, T.R.; Brienen, R.J.W.; Lopez-Gonzalez, G.; Monteagudo, A.; Baker, T.R.; Lewis, S.L.; Lloyd, J.; Quesada, C.A.; Gloor, M.; et al. Markedly Divergent Estimates of Amazon Forest Carbon Density from Ground Plots and Satellites. *Global Ecology and Biogeography* **2014**, *23*, 935–946, doi:10.1111/geb.12168.

35. Sullivan, M.J.P.; Lewis, S.L.; Hubau, W.; Qie, L.; Baker, T.R.; Banin, L.F.; Chave, J.; Cuni-Sanchez, A.; Feldpausch, T.R.; Lopez-Gonzalez, G.; et al. Field Methods for Sampling Tree Height for Tropical Forest Biomass Estimation. *Methods in Ecology and Evolution* **2018**, *9*, 1179–1189, doi:10.1111/2041-210X.12962.

36. Hunter, M.O.; Keller, M.; Victoria, D.; Morton, D.C. Tree Height and Tropical Forest Biomass Estimation. *Biogeosciences* **2013**, *10*, 8385–8399, doi:10.5194/bg-10-8385-2013.

37. Badreldin, N.; Sanchez-Azofeifa, A. Estimating Forest Biomass Dynamics by Integrating Multi-Temporal Landsat Satellite Images with Ground and Airborne LiDAR Data in the Coal Valley Mine, Alberta, Canada. *Remote Sensing* **2015**, *7*, 2832–2849, doi:10.3390/rs70302832.

38. Berninger, A.; Lohberger, S.; Zhang, D.; Siegert, F. Canopy Height and Above-Ground Biomass Retrieval in Tropical Forests Using Multi-Pass X- and C-Band Pol-InSAR Data. *Remote Sensing* **2019**, *11*, 2105, doi:10.3390/rs11182105.

39. Sadeghi, Y.; St-Onge, B.; Leblon, B.; Prieur, J.-F.; Simard, M. Mapping Boreal Forest Biomass from a SRTM and TanDEM-X Based on Canopy Height Model and Landsat Spectral Indices. *International Journal of Applied Earth Observation and Geoinformation* **2018**, *68*, 202–213, doi:10.1016/j.jag.2017.12.004.

40. Persson, H.J.; Fransson, J.E.S. Comparison between TanDEM-X- and ALS-Based Estimation of Aboveground Biomass and Tree Height in Boreal Forests. *Scandinavian Journal of Forest Research* **2017**, *32*, 306–319, doi:10.1080/02827581.2016.1220618.

41. Feldpausch, T.R.; Banin, L.; Phillips, O.L.; Baker, T.R.; Lewis, S.L.; Quesada, C.A.; Affum-Baffoe, K.; Arets, E.J.M.M.; Berry, N.J.; Bird, M.; et al. Height-Diameter Allometry of Tropical Forest Trees. *Biogeosciences* **2011**, *8*, 1081–1106, doi:10.5194/bg-8-1081-2011.

42. Ma, Y.; Zhang, L.; Im, J.; Zhao, Y.; Zhen, Z. Novel Features of Canopy Height Distribution for Aboveground Biomass Estimation Using Machine Learning: A Case Study in Natural Secondary Forests. *Remote Sensing* **2023**, *15*, 4364, doi:10.3390/rs15184364.

43. Fransson, J.E.S. Estimation of Stem Volume in Boreal Forests Using ERS-1 C- and JERS-1 L-Band SAR Data. *International Journal of Remote Sensing* **1999**, *20*, 123–137, doi:10.1080/014311699213640.

44. Le Toan, T.; Quegan, S.; Woodward, I.; Lomas, M.; Delbart, N.; Picard, G. Relating Radar Remote Sensing of Biomass to Modelling of Forest Carbon Budgets. *Climatic Change* **2004**, *67*, 379–402, doi:10.1007/s10584-004-3155-5.

45. Mitchard, E.T.A.; Saatchi, S.S.; Woodhouse, I.H.; Nangendo, G.; Ribeiro, N.S.; Williams, M.; Ryan, C.M.; Lewis, S.L.; Feldpausch, T.R.; Meir, P. Using Satellite Radar Backscatter to Predict Above-Ground Woody Biomass: A Consistent Relationship across Four Different African Landscapes. *Geophysical Research Letters* **2009**, *36*, doi:10.1029/2009GL040692.

46. Ehlers, D.; Wang, C.; Coulston, J.; Zhang, Y.; Pavelsky, T.; Frankenberg, E.; Woodcock, C.; Song, C. Mapping Forest Aboveground Biomass Using Multisource Remotely Sensed Data. *Remote Sensing* **2022**, *14*, 1115, doi:10.3390/rs14051115.

47. Guerra-Hernández, J.; Narine, L.L.; Pascual, A.; Gonzalez-Ferreiro, E.; Botequim, B.; Malambo, L.; Neuenschwander, A.; Popescu, S.C.; Godinho, S. Aboveground Biomass Mapping by Integrating ICESat-2, SENTINEL-1, SENTINEL-2, ALOS2/PALSAR2, and Topographic Information in Mediterranean Forests. *GIScience & Remote Sensing* **2022**, *59*, 1509–1533, doi:10.1080/15481603.2022.2115599.

48. Condés, S.; McRoberts, R.E. Updating National Forest Inventory Estimates of Growing Stock Volume Using Hybrid Inference. *Forest Ecology and Management* **2017**, *400*, 48–57, doi:10.1016/j.foreco.2017.04.046.

49. Lu, D. Aboveground Biomass Estimation Using Landsat TM Data in the Brazilian Amazon. *International Journal of Remote Sensing* **2005**, *26*, 2509–2525, doi:10.1080/01431160500142145.

50. Zhang, L.; Shao, Z.; Liu, J.; Cheng, Q. Deep Learning Based Retrieval of Forest Aboveground Biomass from Combined LiDAR and Landsat 8 Data. *Remote Sensing* **2019**, *11*, 1459, doi:10.3390/rs11121459.

51. Rodríguez-Veiga, P.; Saatchi, S.; Tansey, K.; Balzter, H. Magnitude, Spatial Distribution and Uncertainty of Forest Biomass Stocks in Mexico. *Remote Sensing of Environment* **2016**, *183*, 265–281, doi:10.1016/j.rse.2016.06.004.

52. Puliti, S.; Breidenbach, J.; Schumacher, J.; Hauglin, M.; Klingenberg, T.F.; Astrup, R. Above-Ground Biomass Change Estimation Using National Forest Inventory Data with Sentinel-2 and Landsat. *Remote Sensing of Environment* **2021**, *265*, 112644, doi:10.1016/j.rse.2021.112644.

53. Grabska, E.; Frantz, D.; Ostapowicz, K. Evaluation of Machine Learning Algorithms for Forest Stand Species Mapping Using Sentinel-2 Imagery and Environmental Data in the Polish Carpathians. *Remote Sensing of Environment* **2020**, *251*, 112103, doi:10.1016/j.rse.2020.112103.

54. Dang, A.T.N.; Nandy, S.; Srinet, R.; Luong, N.V.; Ghosh, S.; Senthil Kumar, A. Forest Aboveground Biomass Estimation Using Machine Learning Regression Algorithm in Yok Don National Park, Vietnam. *Ecological Informatics* **2019**, *50*, 24–32, doi:10.1016/j.ecoinf.2018.12.010.

55. European Space Agency *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services*; ESA SP; ESA Communications: Noordwijk, 2012; ISBN 978-92-9221-419-7.

56. Goetz, S.J.; Baccini, A.; Laporte, N.T.; Johns, T.; Walker, W.; Kellndorfer, J.; Houghton, R.A.; Sun, M. Mapping and Monitoring Carbon Stocks with Satellite Observations: A Comparison of Methods. *Carbon Balance and Management* **2009**, *4*, 2, doi:10.1186/1750-0680-4-2.

57. Pham, T.D.; Yoshino, K.; Le, N.N.; Bui, D.T. Estimating Aboveground Biomass of a Mangrove Plantation on the Northern Coast of Vietnam Using Machine Learning Techniques with an Integration of ALOS-2 PALSAR-2 and Sentinel-2A Data. *International Journal of Remote Sensing* **2018**, *39*, 7761–7788, doi:10.1080/01431161.2018.1471544.

58. Vafaei, S.; Soosani, J.; Adeli, K.; Fadaei, H.; Naghavi, H.; Pham, T.D.; Tien Bui, D. Improving Accuracy Estimation of Forest Aboveground Biomass Based on Incorporation of ALOS-2 PALSAR-2 and Sentinel-2A Imagery and Machine Learning: A Case Study of the Hyrcanian Forest Area (Iran). *Remote Sensing* **2018**, *10*, 172, doi:10.3390/rs10020172.

59. Saarela, S.; Holm, S.; Grafström, A.; Schnell, S.; Næsset, E.; Gregoire, T.G.; Nelson, R.F.; Ståhl, G. Hierarchical Model-Based Inference for Forest Inventory Utilizing Three Sources of Information. *Annals of Forest Science* **2016**, *73*, 895–910, doi:10.1007/s13595-016-0590-1.

60. Shao, Z.; Zhang, L. Estimating Forest Aboveground Biomass by Combining Optical and SAR Data: A Case Study in Genhe, Inner Mongolia, China. *Sensors* **2016**, *16*, 834, doi:10.3390/s16060834.

61. Saatchi, S.S.; Harris, N.L.; Brown, S.; Lefsky, M.; Mitchard, E.T.A.; Salas, W.; Zutta, B.R.; Buermann, W.; Lewis, S.L.; Hagen, S.; et al. Benchmark Map of Forest Carbon Stocks in Tropical Regions across Three Continents. *Proceedings of the National Academy of Sciences* **2011**, *108*, 9899–9904, doi:10.1073/pnas.1019576108.

62. Malenovský, Z.; Rott, H.; Cihlar, J.; Schaepman, M.E.; García-Santos, G.; Fernandes, R.; Berger, M. Sentinels for Science: Potential of Sentinel-1, -2, and -3 Missions for Scientific Observations of Ocean, Cryosphere, and Land. *Remote Sensing of Environment* **2012**, *120*, 91–101, doi:10.1016/j.rse.2011.09.026.

63. David, R.M.; Rosser, N.J.; Donoghue, D.N.M. Improving above Ground Biomass Estimates of Southern Africa Dryland Forests by Combining Sentinel-1 SAR and Sentinel-2 Multispectral Imagery. *Remote Sensing of Environment* **2022**, *282*, 113232, doi:10.1016/j.rse.2022.113232.

64. Navarro, J.A.; Algeet, N.; Fernández-Landa, A.; Esteban, J.; Rodríguez-Noriega, P.; Guillén-Climent, M.L. Integration of UAV, Sentinel-1, and Sentinel-2 Data for Mangrove Plantation Aboveground Biomass Monitoring in Senegal. *Remote Sensing* **2019**, *11*, 77, doi:10.3390/rs11010077.

65. Wang, L.; Zhou, X.; Zhu, X.; Dong, Z.; Guo, W. Estimation of Biomass in Wheat Using Random Forest Regression Algorithm and Remote Sensing Data. *The Crop Journal* **2016**, *4*, 212–219, doi:10.1016/j.cj.2016.01.008.

66. Forkuor, G.; Benewinde Zoungrana, J.-B.; Dimobe, K.; Ouattara, B.; Vadrevu, K.P.; Tondoh, J.E. Above-Ground Biomass Mapping in West African Dryland Forest Using Sentinel-1 and 2 Datasets - A Case Study. *Remote Sensing of Environment* **2020**, *236*, 111496, doi:10.1016/j.rse.2019.111496.

67. Koch, B. Status and Future of Laser Scanning, Synthetic Aperture Radar and Hyperspectral Remote Sensing Data for Forest Biomass Assessment. *ISPRS Journal of Photogrammetry and Remote Sensing* **2010**, *65*, 581–590, doi:10.1016/j.isprsjprs.2010.09.001.

68. Ghosh, S.M.; Behera, M.D. Aboveground Biomass Estimation Using Multi-Sensor Data Synergy and Machine Learning Algorithms in a Dense Tropical Forest. *Applied Geography* **2018**, *96*, 29–40, doi:10.1016/j.apgeog.2018.05.011.

69. George-Chacon, S.P.; Dupuy, J.M.; Peduzzi, A.; Hernandez-Stefanoni, J.L. Combining High Resolution Satellite Imagery and Lidar Data to Model Woody Species Diversity of Tropical Dry Forests. *Ecological Indicators* **2019**, *101*, 975–984, doi:10.1016/j.ecolind.2019.02.015.

70. Vaglio Laurin, G.; Puletti, N.; Chen, Q.; Corona, P.; Papale, D.; Valentini, R. Above Ground Biomass and Tree Species Richness Estimation with Airborne Lidar in Tropical Ghana Forests. *International Journal of Applied Earth Observation and Geoinformation* **2016**, *52*, 371–379, doi:10.1016/j.jag.2016.07.008.

71. Duncanson, L.I.; Dubayah, R.O.; Cook, B.D.; Rosette, J.; Parker, G. The Importance of Spatial Detail: Assessing the Utility of Individual Crown Information and Scaling Approaches for Lidar-Based Biomass Density Estimation. *Remote Sensing of Environment* **2015**, *168*, 102–112, doi:10.1016/j.rse.2015.06.021.

72. Asner, G.P.; Mascaro, J. Mapping Tropical Forest Carbon: Calibrating Plot Estimates to a Simple LiDAR Metric. *Remote Sensing of Environment* **2014**, *140*, 614–624, doi:10.1016/j.rse.2013.09.023.

73. Stojanova, D.; Panov, P.; Gjorgjioski, V.; Kobler, A.; Džeroski, S. Estimating Vegetation Height and Canopy Cover from Remotely Sensed Data with Machine Learning. *Ecological Informatics* **2010**, *5*, 256–266, doi:10.1016/j.ecoinf.2010.03.004.

74. Henn, K.A.; Peduzzi, A. Biomass Estimation of Urban Forests Using LiDAR and High-Resolution Aerial Imagery in Athens–Clarke County, GA. *Forests* **2023**, *14*, 1064, doi:10.3390/f14051064.

75. Montesano, P.M.; Nelson, R.F.; Dubayah, R.O.; Sun, G.; Cook, B.D.; Ranson, K.J.R.; Næsset, E.; Kharuk, V. The Uncertainty of Biomass Estimates from LiDAR and SAR across a Boreal Forest Structure Gradient. *Remote Sensing of Environment* **2014**, *154*, 398–407, doi:10.1016/j.rse.2014.01.027.

76. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sensing of Environment* **2017**, *202*, 18–27, doi:10.1016/j.rse.2017.06.031.

77. Google Earth Engine Earth Engine Data Catalog Available online: https://developers.google.com/earth-engine/datasets/catalog (accessed on 20 March 2024).

78. Google Earth Engine Awesome GEE Community Catalog | Earth Engine Data Catalog Available online: https://developers.google.com/earth-engine/datasets/community/sat-io (accessed on 20 March 2024).

79. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sensing* **2018**, *10*, 1509, doi:10.3390/rs10101509.

80. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine Learning Regression Algorithms for Biophysical Parameter Retrieval: Opportunities for Sentinel-2 and -3. *Remote Sensing of Environment* **2012**, *118*, 127–139, doi:10.1016/j.rse.2011.11.002.

81. Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5–32, doi:10.1023/A:1010933404324.

82. Dahms, T.; Seissiger, S.; Borg, E.; Vajen, H.; Fichtelmann, B.; Conrad, C. Important Variables of a RapidEye Time Series for Modelling Biophysical Parameters of Winter Wheat. *Photogrammetrie - Fernerkundung - Geoinformation* **2016**, 285–299, doi:10.1127/pfg/2016/0303.

83. Hastie, T.; Tibshirani, R.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer series in statistics; 2nd ed.; Springer: New York, NY, 2009; ISBN 978-0-387-84857-0.

84. Cutler, D.R.; Edwards Jr., T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random Forests for Classification in Ecology. *Ecology* **2007**, *88*, 2783–2792, doi:10.1890/07-0539.1.

85. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems* **2006**, *9*, 181–199, doi:10.1007/s10021-005-0054-1.

86. Cutler, A.; Cutler, D.R.; Stevens, J.R. Random Forests. In *Ensemble Machine Learning: Methods and Applications*; Zhang, C., Ma, Y., Eds.; Springer: New York, NY, 2012; pp. 157–175 ISBN 978-1-4419-9326-7.

87. Argamosa, R.J.L.; Blanco, A.C.; Baloloy, A.B.; Candido, C.G.; Dumalag, J.B.L.C.; Dimapilis, L.L.C.; Paringit, E.C. MODELLING ABOVE GROUND BIOMASS OF MANGROVE FOREST USING SENTINEL-1 IMAGERY. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2018**, *IV–3*, 13–20, doi:10.5194/isprs-annals-IV-3-13-2018.

88. Nandy, S.; Srinet, R.; Padalia, H. Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India. *Geophysical Research Letters* **2021**, *48*, e2021GL093799, doi:10.1029/2021GL093799.

89. Silveira, E.M.O.; Silva, S.H.G.; Acerbi-Junior, F.W.; Carvalho, M.C.; Carvalho, L.M.T.; Scolforo, J.R.S.; Wulder, M.A. Object-Based Random Forest Modelling of Aboveground Forest Biomass Outperforms a Pixel-Based Approach in a Heterogeneous and Mountain Tropical Environment. *International Journal of Applied Earth Observation and Geoinformation* **2019**, *78*, 175–188, doi:10.1016/j.jag.2019.02.004.

90. Boisvenue, C.; Smiley, B.P.; White, J.C.; Kurz, W.A.; Wulder, M.A. Integration of Landsat Time Series and Field Plots for Forest Productivity Estimates in Decision Support Models. *Forest Ecology and Management* **2016**, *376*, 284–297, doi:10.1016/j.foreco.2016.06.022.

91. DeAngelis, D.L. Boreal Forest. In *Encyclopedia of Ecology*; Jørgensen, S.E., Fath, B.D., Eds.; Academic Press: Oxford, 2008; pp. 493–495 ISBN 978-0-08-045405-4.

92. Yang, H.; Ciais, P.; Frappart, F.; Li, X.; Brandt, M.; Fensholt, R.; Fan, L.; Saatchi, S.; Besnard, S.; Deng, Z.; et al. Global Increase in Biomass Carbon Stock Dominated by Growth of Northern Young Forests over Past Decade. *Nat. Geosci.* **2023**, *16*, 886–892, doi:10.1038/s41561-023-01274-4.

93. Verkerk, P.J.; Fitzgerald, J.B.; Datta, P.; Dees, M.; Hengeveld, G.M.; Lindner, M.; Zudin, S. Spatial Distribution of the Potential Forest Biomass Availability in Europe. *For. Ecosyst.* **2019**, *6*, 5, doi:10.1186/s40663-019-0163-5.

94. Kangas, A.; Astrup, R.; Breidenbach, J.; Fridman, J.; Gobakken, T.; Korhonen, K.T.; Maltamo, M.; Nilsson, M.; Nord-Larsen, T.; Næsset, E.; et al. Remote Sensing and Forest Inventories in Nordic Countries – Roadmap for the Future. *Scandinavian Journal of Forest Research* **2018**, *33*, 397–412, doi:10.1080/02827581.2017.1416666.

95. Turubanova, S.; Potapov, P.; Hansen, M.C.; Li, X.; Tyukavina, A.; Pickens, A.H.; Hernandez-Serna, A.; Arranz, A.P.; Guerra-Hernández, J.; Senf, C.; et al. Tree Canopy Extent and Height Change in Europe, 2001–2021, Quantified Using Landsat Data Archive. *Remote Sensing of Environment* **2023**, *298*, 113797, doi:10.1016/j.rse.2023.113797.

96. Swedish Forest Industry Federation *Swedish Forests*; 2016;

97. Swedish Forest Industry Federation *Economic Importance*; 2023;

98. Swedish Forest Industry Federation *Global Forest Industry*; 2016;

99. Swedish Forest Industry Federation Facts & Figures Available online: https://www.forestindustries.se/forest-industry/statistics/facts-and-figures/ (accessed on 11 March 2024).

100. Titus, B.D.; Brown, K.; Helmisaari, H.-S.; Vanguelova, E.; Stupak, I.; Evans, A.; Clarke, N.; Guidi, C.; Bruckman, V.J.; Varnagiryte-Kabasinskiene, I.; et al. Sustainable Forest Biomass: A Review of Current Residue Harvesting Guidelines. *Energy, Sustainability and Society* **2021**, *11*, 10, doi:10.1186/s13705-021-00281-w.

101. Clarke, N.; Kiær, L.P.; Janne Kjønaas, O.; Bárcena, T.G.; Vesterdal, L.; Stupak, I.; Finér, L.; Jacobson, S.; Armolaitis, K.; Lazdina, D.; et al. Effects of Intensive Biomass Harvesting on Forest Soils in the Nordic Countries and the UK: A Meta-Analysis. *Forest Ecology and Management* **2021**, *482*, 118877, doi:10.1016/j.foreco.2020.118877.

102. de Jong, J.; Akselsson, C.; Egnell, G.; Löfgren, S.; Olsson, B.A. Realizing the Energy Potential of Forest Biomass in Sweden – How Much Is Environmentally Sustainable? *Forest Ecology and Management* **2017**, *383*, 3–16, doi:10.1016/j.foreco.2016.06.028.

103. Kumar, A.; Adamopoulos, S.; Jones, D.; Amiandamhen, S.O. Forest Biomass Availability and Utilization Potential in Sweden: A Review. *Waste Biomass Valor* **2021**, *12*, 65–80, doi:10.1007/s12649-020-00947-0.

104. Kumar, L.; Sinha, P.; Taylor, S.; Alqurashi, A.F. Review of the Use of Remote Sensing for Biomass Estimation to Support Renewable Energy Generation. *JARS* **2015**, *9*, 097696, doi:10.1117/1.JRS.9.097696.

105. Jalkanen, A.; Mäkipää, R.; Ståhl, G.; Lehtonen, A.; Petersson, H. Estimation of the Biomass Stock of Trees in Sweden: Comparison of Biomass Equations and Age-Dependent Biomass Expansion Factors. *Ann. For. Sci.* **2005**, *62*, 845–851, doi:10.1051/forest:2005075.

106. Shendryk, I.; Hellström, M.; Klemedtsson, L.; Kljun, N. Low-Density LiDAR and Optical Imagery for Biomass Estimation over Boreal Forest in Sweden. *Forests* **2014**, *5*, 992–1010, doi:10.3390/f5050992.

107. Askne, J.I.H.; Santoro, M. On the Estimation of Boreal Forest Biomass From TanDEM-X Data Without Training Samples. *IEEE Geoscience and Remote Sensing Letters* **2015**, *12*, 771–775, doi:10.1109/LGRS.2014.2361393.

108. Persson, H.J.; Olsson, H.; Soja, M.J.; Ulander, L.M.H.; Fransson, J.E.S. Experiences from Large-Scale Forest Mapping of Sweden Using TanDEM-X Data. *Remote Sensing* **2017**, *9*, 1253, doi:10.3390/rs9121253.

109. Persson, H.J.; Jonzén, J.; Nilsson, M. Combining TanDEM-X and Sentinel-2 for Large-Area Species-Wise Prediction of Forest Biomass and Volume. *International Journal of Applied Earth Observation and Geoinformation* **2021**, *96*, 102275, doi:10.1016/j.jag.2020.102275.

110. Majasalmi, T.; Rautiainen, M. The Potential of Sentinel-2 Data for Estimating Biophysical Variables in a Boreal Forest: A Simulation Study. *Remote Sensing Letters* **2016**, *7*, 427–436, doi:10.1080/2150704X.2016.1149251.

111. Swedish University of Agricultural Sciences Field Work Instructions 2021 - Swedish National Forest Inventory | Swedish Soil Inventory 2021.

112. Zanaga, D.; Van De Kerchove, R.; Daems, D.; De Keersmaecker, W.; Brockmann, C.; Kirches, G.; Wevers, J.; Cartus, O.; Santoro, M.; Fritz, S.; et al. ESA WorldCover 10 m 2021 V200 2022.

113. European Space Agency MSI Level-2A BOA Reflectance Product. Collection 1 Available online: https://copernicus.eu/sentinel-data-access/sentinel-products/sentinel-2-data-products/collection-1-level-2a (accessed on 14 March 2024).

114. Xi, Z.; Xu, H.; Xing, Y.; Gong, W.; Chen, G.; Yang, S. Forest Canopy Height Mapping by Synergizing ICESat-2, Sentinel-1, Sentinel-2 and Topographic Information Based on Machine Learning Methods. *Remote Sensing* **2022**, *14*, 364, doi:10.3390/rs14020364.

115. Xu, C.; Ding, Y.; Zheng, X.; Wang, Y.; Zhang, R.; Zhang, H.; Dai, Z.; Xie, Q. A Comprehensive Comparison of Machine Learning and Feature Selection Methods for Maize Biomass Estimation Using Sentinel-1 SAR, Sentinel-2 Vegetation Indices, and Biophysical Variables. *Remote Sensing* **2022**, *14*, 4083, doi:10.3390/rs14164083.

116. Luo, C.; Liu, H.; Lu, L.; Liu, Z.; Kong, F.; Zhang, X. Monthly Composites from Sentinel-1 and Sentinel-2 Images for Regional Major Crop Mapping with Google Earth Engine. *Journal of Integrative Agriculture* **2021**, *20*, 1944–1957, doi:10.1016/S2095-3119(20)63329-9.

117. Li, W.; Niu, Z.; Shang, R.; Qin, Y.; Wang, L.; Chen, H. High-Resolution Mapping of Forest Canopy Height Using Machine Learning by Coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 Data. *International Journal of Applied Earth Observation and Geoinformation* **2020**, *92*, 102163, doi:10.1016/j.jag.2020.102163.

118. Braaten, J. Sentinel-2 Cloud Masking with S2cloudless | Google Earth Engine Available online: https://developers.google.com/earth-engine/tutorials/community/sentinel-2-s2cloudless (accessed on 30 November 2023).

119. Google Earth Engine Sentinel-2: Cloud Probability | Earth Engine Data Catalog Available online: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_CLOUD_PROBABILITY (accessed on 30 November 2023).

120. Frantz, D.; Haß, E.; Uhl, A.; Stoffels, J.; Hill, J. Improvement of the Fmask Algorithm for Sentinel-2 Images: Separating Clouds from Bright Surfaces Based on Parallax Effects. *Remote Sensing of Environment* **2018**, *215*, 471–481, doi:10.1016/j.rse.2018.04.046.

121. Richardson, A.J.; Everitt, J.H. Using Spectral Vegetation Indices to Estimate Rangeland Productivity. *Geocarto International* **1992**, *7*, 63–69, doi:10.1080/10106049209354353.

122. Liu, H.Q.; Huete, A. A Feedback Based Modification of the NDVI to Minimize Canopy Background and Atmospheric Noise. *IEEE Transactions on Geoscience and Remote Sensing* **1995**, *33*, 457–465, doi:10.1109/TGRS.1995.8746027.

123. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a Two-Band Enhanced Vegetation Index without a Blue Band. *Remote Sensing of Environment* **2008**, *112*, 3833–3845, doi:10.1016/j.rse.2008.06.006.

124. Gitelson, A.A.; Merzlyak, M.N. Remote Sensing of Chlorophyll Concentration in Higher Plant Leaves. *Advances in Space Research* **1998**, *22*, 689–692, doi:10.1016/S0273-1177(97)01133-2.

125. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sensing of Environment* **1979**, *8*, 127–150, doi:10.1016/0034-4257(79)90013-0.

126. Frampton, W.; Dash, J.; Watmough, G.; Milton, E. Evaluating the Capabilities of Sentinel-2 for Quantitative Estimation of Biophysical Variables in Vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing* **2013**, *82*, 83–92, doi:10.1016/j.isprsjprs.2013.04.007.

127. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; de Colstoun, E.B.; McMurtrey, J.E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment* **2000**, *74*, 229–239, doi:10.1016/S0034-4257(00)00113-9.

128. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A Modified Soil Adjusted Vegetation Index. *Remote Sensing of Environment* **1994**, *48*, 119–126, doi:10.1016/0034-4257(94)90134-1.

129. Key, C.; Benson, N.; Ohlen, D.; Howard, S.; Zhu, Z. The Normalized Burn Ratio and Relationships to Burn Severity: Ecology, Remote Sensing and Implementation. In Proceedings of the Proceedings of the Ninth Forest Service Remote Sensing Applications Conference. American Society for Photogrammetry and Remote Sensing, Bethesda, MD; 2002.

130. USGS Landsat Normalized Burn Ratio 2 Available online: https://www.usgs.gov/landsat-missions/landsat-normalized-burn-ratio-2 (accessed on 25 October 2023).

131. Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors* **2011**, *11*, 7063–7081, doi:10.3390/s110707063.

132. Buschmann, C. Fernerkundung von Pflanzen. *Naturwissenschaften* **1993**, *80*, 439–453, doi:10.1007/BF01136034.

133. Barnes, E.; Clarke, T.R.; Richards, S.E.; Colaizzi, P.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T.L. Coincident Detection of Crop Water Stress, Nitrogen Status, and Canopy Density Using Ground Based Multispectral Data. **2000**.

134. Rouse, J.W.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; 1974;

135. Rondeaux, G.; Steven, M.; Baret, F. Optimization of Soil-Adjusted Vegetation Indices. *Remote Sensing of Environment* **1996**, *55*, 95–107, doi:10.1016/0034-4257(95)00186-7.

136. Blackburn, G.A. Spectral Indices for Estimating Photosynthetic Pigment Concentrations: A Test Using Senescent Tree Leaves. *International Journal of Remote Sensing* **1998**, doi:10.1080/014311698215919.

137. CARTER, G.A. Ratios of Leaf Reflectances in Narrow Wavebands as Indicators of Plant Stress. *International Journal of Remote Sensing* **1994**, *15*, 697–703, doi:10.1080/01431169408954109.

138. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote Estimation of Canopy Chlorophyll Content in Crops. *Geophysical Research Letters* **2005**, *32*, doi:10.1029/2005GL022688.

139. Cao, Q.; Miao, Y.; Shen, J.; Yu, W.; Yuan, F.; Cheng, S.; Huang, S.; Wang, H.; Yang, W.; Liu, F. Improving In-Season Estimation of Rice Yield Potential and Responsiveness to Topdressing Nitrogen Application with Crop Circle Active Crop Canopy Sensor. *Precision Agric* **2016**, *17*, 136–154, doi:10.1007/s11119-015-9412-y.

140. Roujean, J.-L.; Breon, F.-M. Estimating PAR Absorbed by Vegetation from Bidirectional Reflectance Measurements. *Remote Sensing of Environment* **1995**, *51*, 375–384, doi:10.1016/0034-4257(94)00114-3.

141. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment* **1988**, *25*, 295–309, doi:10.1016/0034-4257(88)90106-X.

142. Thenkabail, P.S.; Ward, A.D.; Lyon, J.; Merry, C.J. Thematic Mapper Vegetation Indices for Determining Soybean and Corn Growth Parameters. **1994**, *60*.

143. Haboudane, D.; Miller, J.; Tremblay, N.; Zarco-Tejada, P.; Dextraze, L. Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture. *Remote Sensing of Environment* **2002**, *81*, 416–426, doi:10.1016/S0034-4257(02)00018-4.

144. Watson, D.J. Comparative Physiological Studies on the Growth of Field Crops: I. Variation in Net Assimilation Rate and Leaf Area between Species and Varieties, and within and between Years. *Annals of Botany* **1947**, *11*, 41–76, doi:10.1093/oxfordjournals.aob.a083148.

145. Gutman, G.; Ignatov, A. The Derivation of the Green Vegetation Fraction from NOAA/AVHRR Data for Use in Numerical Weather Prediction Models. *International Journal of Remote Sensing* **1998**, *19*, 1533–1543, doi:10.1080/014311698215333.

146. Myneni, R.B.; Hall, F.; Sellers, P.; Marshak, A. The Interpretation of Spectral Vegetation Indexes. IEEE Trans. Geosci. Remote Sens. *Geoscience and Remote Sensing, IEEE Transactions on* **1995**, *33*, 481–486, doi:10.1109/36.377948.

147. Gitelson, A.A.; Merzlyak, M.N. Remote Estimation of Chlorophyll Content in Higher Plant Leaves. *International Journal of Remote Sensing* **1997**, *18*, 2691–2697, doi:10.1080/014311697217558.

148. Ceccato, P.; Gobron, N.; Flasse, S.; Pinty, B.; Tarantola, S. Designing a Spectral Index to Estimate Vegetation Water Content from Remote Sensing Data: Part 1: Theoretical Approach. *Remote Sensing of Environment* **2002**, *82*, 188–197, doi:10.1016/S0034-4257(02)00037-8.
149. Google Earth Engine Sentinel-1 Algorithms | Google Earth Engine Available online: https://developers.google.com/earth-engine/guides/sentinel1 (accessed on 26 October 2023).
150. European Space Agency Sentinel-1 Product Definition 2016.
151. Mullissa, A.; Vollrath, A.; Odongo-Braun, C.; Slagter, B.; Balling, J.; Gou, Y.; Gorelick, N.; Reiche, J. Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine. *Remote Sensing* **2021**, *13*, 1–8, doi:10.3390/rs13101954.
152. Laurin, G.V.; Balling, J.; Corona, P.; Mattioli, W.; Papale, D.; Puletti, N.; Rizzo, M.; Truckenbrodt, J.; Urban, M. Above-Ground Biomass Prediction by Sentinel-1 Multitemporal Data in Central Italy with Integration of ALOS2 and Sentinel-2 Data. *JARS* **2018**, *12*, 016008, doi:10.1117/1.JRS.12.016008.
153. Nguyen, L.V.; Tateishi, R.; Nguyen, H.T.; Sharma, R.C.; To, T.T.; Le, S.M. Estimation of Tropical Forest Structural Characteristics Using ALOS-2 SAR Data. *Advances in Remote Sensing* **2016**, *5*, 131–144, doi:10.4236/ars.2016.52011.
154. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics* **1973**, *SMC-3*, 610–621, doi:10.1109/TSMC.1973.4309314.
155. Wijaya, A.; Kusnadi, S.; Gloaguen, R.; Heilmeier, H. Improved Strategy for Estimating Stem Volume and Forest Biomass Using Moderate Resolution Remote Sensing Data and GIS. *Journal of Forestry Research* **2010**, *21*, 1–12, doi:10.1007/s11676-010-0001-7.
156. Debastiani, A.B.; Sanquetta, C.R.; Corte, A.P.D.; Pinto, N.S.; Rex, F.E. Evaluating SAR-Optical Sensor Fusion for Aboveground Biomass Estimation in a Brazilian Tropical Forest. *Annals of Forest Research* **2009**, 109–122, doi:10.15287/afr.2018.1267.
157. Lourenço, P.; Godinho, S.; Sousa, A.; Gonçalves, A.C. Estimating Tree Aboveground Biomass Using Multispectral Satellite-Based Data in Mediterranean Agroforestry System Using Random Forest Algorithm. *Remote Sensing Applications: Society and Environment* **2021**, *23*, 100560, doi:10.1016/j.rsase.2021.100560.
158. Zhao, P.; Lu, D.; Wang, G.; Wu, C.; Huang, Y.; Yu, S. Examining Spectral Reflectance Saturation in Landsat Imagery and Corresponding Solutions to Improve Forest Aboveground Biomass Estimation. *Remote Sensing* **2016**, *8*, 469, doi:10.3390/rs8060469.
159. European Space Agency *Sentinel-1: ESA's Radar Observatory Mission for GMES Operational Services*; ESA SP; ESA communications: Noordwijk (NL), 2012; ISBN 978-92-9221-418-0.
160. Kim, Y.; van Zyl, J. Comparison of Forest Parameter Estimation Techniques Using SAR Data. In Proceedings of the IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217); July 2001; Vol. 3, pp. 1395–1397 vol.3.
161. Google Earth Engine Ee.Image.glcmTexture Available online: https://developers.google.com/earth-engine/apidocs/ee-image-glcmtexture (accessed on 14 March 2024).
162. European Space Agency Copernicus Digital Elevation Model Product Handbook 2022.
163. Dutcă, I.; Cernat, A.; Stăncioiu, P.T.; Ioraș, F.; Niță, M.D. Does Slope Aspect Affect the Aboveground Tree Shape and Volume Allometry of European Beech (Fagus Sylvatica L.) Trees? *Forests* **2022**, *13*, 1071, doi:10.3390/f13071071.
164. Singh, S. Understanding the Role of Slope Aspect in Shaping the Vegetation Attributes and Soil Properties in Montane Ecosystems. *Tropical Ecology* **2018**, *59*, 417–430.
165. Castillo, J.A.A.; Apan, A.A.; Maraseni, T.N.; Salmo, S.G. Estimation and Mapping of Above-Ground Biomass of Mangrove Forests and Their Replacement Land Uses in the Philippines Using Sentinel Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **2017**, *134*, 70–85, doi:10.1016/j.isprsjprs.2017.10.016.
166. Feldpausch, T.R.; Lloyd, J.; Lewis, S.L.; Brienen, R.J.W.; Gloor, M.; Monteagudo Mendoza, A.; Lopez-Gonzalez, G.; Banin, L.; Abu Salim, K.; Affum-Baffoe, K.; et al. Tree Height Integrated into Pantropical Forest Biomass Estimates. *Biogeosciences* **2012**, *9*, 3381–3403, doi:10.5194/bg-9-3381-2012.

167. Campbell, M.J.; Dennison, P.E.; Kerr, K.L.; Brewer, S.C.; Anderegg, W.R.L. Scaled Biomass Estimation in Woodland Ecosystems: Testing the Individual and Combined Capacities of Satellite Multispectral and Lidar Data. *Remote Sensing of Environment* **2021**, *262*, 112511, doi:10.1016/j.rse.2021.112511.

168. Lang, N.; Jetz, W.; Schindler, K.; Wegner, J.D. A High-Resolution Canopy Height Model of the Earth. *Nat Ecol Evol* **2023**, *7*, 1778–1789, doi:10.1038/s41559-023-02206-6.

169. Esri Inc. ArcGIS Pro 2023.

170. rapidlasso GmbH LAStools 2023.

171. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, 2009; ISBN 1-4414-1269-7.

172. McKinney, W.; others Data Structures for Statistical Computing in Python. In Proceedings of the Proceedings of the 9th Python in Science Conference; Austin, TX, 2010; Vol. 445, pp. 51–56.

173. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *Journal of machine learning research* **2011**, *12*, 2825–2830.

174. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Computing in science & engineering* **2007**, *9*, 90–95.

175. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array Programming with NumPy. *Nature* **2020**, *585*, 357–362, doi:10.1038/s41586-020-2649-2.

176. Pérez, F.; Granger, B.E. IPython: A System for Interactive Scientific Computing. *Computing in Science & Engineering* **2007**, *9*.

177. Mutanga, O.; Adam, E.; Cho, M.A. High Density Biomass Estimation for Wetland Vegetation Using WorldView-2 Imagery and Random Forest Regression Algorithm. *International Journal of Applied Earth Observation and Geoinformation* **2012**, *18*, 399–406, doi:10.1016/j.jag.2012.03.012.

178. Belgiu, M.; Drăguţ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS Journal of Photogrammetry and Remote Sensing* **2016**, *114*, 24–31, doi:10.1016/j.isprsjprs.2016.01.011.

179. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; Koch, B. Importance of Sample Size, Data Type and Prediction Method for Remote Sensing-Based Estimations of Aboveground Forest Biomass. *Remote Sensing of Environment* **2014**, *154*, 102–114, doi:10.1016/j.rse.2014.07.028.

180. Malhi, R.K.M.; Anand, A.; Srivastava, P.K.; Chaudhary, S.K.; Pandey, M.K.; Behera, M.D.; Kumar, A.; Singh, P.; Sandhya Kiran, G. Synergistic Evaluation of Sentinel 1 and 2 for Biomass Estimation in a Tropical Forest of India. *Advances in Space Research* **2022**, *69*, 1752–1767, doi:10.1016/j.asr.2021.03.035.

181. Silveira, E.M.O.; Radeloff, V.C.; Martinuzzi, S.; Martinez Pastur, G.J.; Bono, J.; Politi, N.; Lizarraga, L.; Rivera, L.O.; Ciuffoli, L.; Rosas, Y.M.; et al. Nationwide Native Forest Structure Maps for Argentina Based on Forest Inventory Data, SAR Sentinel-1 and Vegetation Metrics from Sentinel-2 Imagery. *Remote Sensing of Environment* **2023**, *285*, 113391, doi:10.1016/j.rse.2022.113391.

182. Olesk, A.; Voormansik, K.; Vain, A.; Noorma, M.; Praks, J. Seasonal Differences in Forest Height Estimation From Interferometric TanDEM-X Coherence Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2015**, *8*, 5565–5572, doi:10.1109/JSTARS.2015.2501648.

183. Coops, N.C.; Tompalski, P.; Goodbody, T.R.H.; 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 Sensing of Environment* **2021**, *260*, 112477, doi:10.1016/j.rse.2021.112477.

184. Sun, T.; Qi, J.; Huang, H. Discovering Forest Height Changes Based on Spaceborne Lidar Data of ICESat-1 in 2005 and ICESat-2 in 2019: A Case Study in the Beijing-Tianjin-Hebei Region of China. *Forest Ecosystems* **2020**, *7*, 53, doi:10.1186/s40663-020-00265-w.

185. Huang, H.; Liu, C.; Wang, X.; Zhou, X.; Gong, P. Integration of Multi-Resource Remotely Sensed Data and Allometric Models for Forest Aboveground Biomass Estimation in China. *Remote Sensing of Environment* **2019**, *221*, 225–234, doi:10.1016/j.rse.2018.11.017.

186. Agata, H.; Aneta, L.; Dariusz, Z.; Krzysztof, S.; Marek, L.; Christiane, S.; Carsten, P. Forest Aboveground Biomass Estimation Using a Combination of Sentinel-1 and Sentinel-2 Data. In Proceedings of the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium; July 2018; pp. 9026–9029.

187. Ghosh, S.M.; Behera, M.D.; Kumar, S.; Das, P.; Prakash, A.J.; Bhaskaran, P.K.; Roy, P.S.; Barik, S.K.; Jeganathan, C.; Srivastava, P.K.; et al. Predicting the Forest Canopy Height from LiDAR and Multi-Sensor Data Using Machine Learning over India. *Remote Sensing* **2022**, *14*, 5968, doi:10.3390/rs14235968.

188. Lei, Y.; Siqueira, P. An Automatic Mosaicking Algorithm for the Generation of a Large-Scale Forest Height Map Using Spaceborne Repeat-Pass InSAR Correlation Magnitude. *Remote Sensing* **2015**, *7*, 5639–5659, doi:10.3390/rs70505639.

189. Georgopoulos, N.; Sotiropoulos, C.; Stefanidou, A.; Gitas, I.Z. Total Stem Biomass Estimation Using Sentinel-1 and -2 Data in a Dense Coniferous Forest of Complex Structure and Terrain. *Forests* **2022**, *13*, 2157, doi:10.3390/f13122157.

190. Lindberg, E.; Hollaus, M. Comparison of Methods for Estimation of Stem Volume, Stem Number and Basal Area from Airborne Laser Scanning Data in a Hemi-Boreal Forest. *Remote Sensing* **2012**, *4*, 1004–1023, doi:10.3390/rs4041004.

191. Valavi, R.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. Modelling Species Presence-Only Data with Random Forests. *Ecography* **2021**, *44*, 1731–1742, doi:10.1111/ecog.05615.

192. Dube, T.; Mutanga, O.; Elhadi, A.; Ismail, R. Intra-and-Inter Species Biomass Prediction in a Plantation Forest: Testing the Utility of High Spatial Resolution Spaceborne Multispectral RapidEye Sensor and Advanced Machine Learning Algorithms. *Sensors* **2014**, *14*, 15348–15370, doi:10.3390/s140815348.