

## Article

# Estimation of Forest Growing Stock Volume with Synthetic Aperture Radar: A Comparison of Model-Fitting Methods

Maurizio Santoro <sup>1,\*</sup>, Oliver Cartus <sup>1</sup>, Oleg Antropov <sup>2</sup> and Jukka Miettinen <sup>2</sup><sup>1</sup> Gamma Remote Sensing, Worbstrasse 225, 3073 Gümligen, Switzerland; cartus@gamma-rs.ch<sup>2</sup> VTT Technical Research Centre of Finland, P.O. Box 1000, 02044 Espoo, Finland; oleg.antropov@vtt.fi (O.A.); jukka.miettinen@vtt.fi (J.M.)

\* Correspondence: santoro@gamma-rs.ch; Tel.: +41-31-9517005

**Abstract:** Satellite-based estimation of forest variables including forest biomass relies on model-based approaches since forest biomass cannot be directly measured from space. Such models require ground reference data to adapt to the local forest structure and acquired satellite data. For wide-area mapping, such reference data are too sparse to train the biomass retrieval model and approaches for calibrating that are independent from training data are sought. In this study, we compare the performance of one such calibration approach with the traditional regression modelling using reference measurements. The performance was evaluated at four sites representative of the major forest biomes in Europe focusing on growing stock volume (GSV) prediction from time series of C-band Sentinel-1 and Advanced Land Observing Satellite Phased Array L-band Synthetic Aperture Radar (ALOS-2 PALSAR-2) backscatter measurements. The retrieval model was based on a Water Cloud Model (WCM) and integrated two forest structural functions. The WCM trained with plot inventory GSV values or calibrated with the aid of auxiliary data products correctly reproduced the trend between SAR backscatter and GSV measurements across all sites. The WCM-predicted backscatter was within the range of measurements for a given GSV level with average model residuals being smaller than the range of the observations. The accuracy of the GSV estimated with the calibrated WCM was close to the accuracy obtained with the trained WCM. The difference in terms of root mean square error (RMSE) was less than 5% units. This study demonstrates that it is possible to predict biomass without providing reference measurements for model training provided that the modelling scheme is physically based and the calibration is well set and understood.

**Keywords:** forest; SAR backscatter; growing stock volume; Sentinel-1; ALOS-2 PALSAR-2; ICESat-2; Water Cloud Model



**Citation:** Santoro, M.; Cartus, O.; Antropov, O.; Miettinen, J. Estimation of Forest Growing Stock Volume with Synthetic Aperture Radar: A Comparison of Model-Fitting Methods. *Remote Sens.* **2024**, *16*, 4079. <https://doi.org/10.3390/rs16214079>

Academic Editors: Armando Marino and Michele Martone

Received: 21 September 2024

Revised: 29 October 2024

Accepted: 30 October 2024

Published: 31 October 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Knowledge of the spatial distribution and magnitude of forest resources is of paramount importance to understand pathways of climate, to support the economy of communities, and to provide societal benefits [1]. Forest resources are best quantified with a variable that expresses the biomass stored by vegetation, this being a direct proxy for its carbon content. Estimation of forest biomass defined either as the structural variable growing stock volume (GSV, unit  $\text{m}^3 \text{ ha}^{-1}$ ) or as organic mass of dry matter, i.e., above ground biomass (AGB, unit  $\text{Mg ha}^{-1}$ ), requires measurements of both the horizontal (i.e., basal area) and vertical (i.e., tree height) properties of a forest, as well as the number of trees, to obtain the biomass density. To estimate the volume, in addition, the tree form factor is needed whereas the organic mass can be estimated from the wood volume by accounting for the wood density and the total-to-stem biomass ratio [2]. Field measurements provide accurate information locally. Such measurements, however, are not easily scalable and require statistical tools to obtain spatially explicit estimates [3].

Satellite-based active microwave remote sensing enables frequent and systematic monitoring of the Earth's surface. Over vegetated land, the observations contain information

on both the structure and water content of plants and soil [4]. Identifying each component in the microwave signal is therefore essential for quantifying the quantification of the biomass stored in the vegetation. Typically, this is achieved by preferring observations where the contribution of the forest structure to the recorded signal is predominant. Research demonstrated that data acquired by Synthetic Aperture Radar (SAR) operating at longer wavelengths such as in the Very High Frequency (VHF) band [5] and P-band [6] are more strongly related to forest biomass than data acquired by SAR sensors operating at shorter wavelengths ([7,8]). In space, however, SAR sensors have operated only up to the L-band (23 cm) so far, so that the sensitivity to biomass of the SAR backscattered intensity in satellite images is mostly moderate or weak due to the limited capability of the microwave signal to penetrate the canopy and the influence of environmental conditions and terrain slope on the backscattered signal [9–14]. Stronger agreement with biomass was obtained with polarimetric features [15–18]. To overcome the limitations intrinsic in the data, a wide range of parametric and non-parametric model-based approaches that aimed at estimating biomass from SAR images have been developed and tested [9,10,19]. With the availability of SAR image datasets acquired at different frequencies and polarizations, these approaches have been advanced to improve the accuracy of the estimation by exploiting complementarity of the observations, for example, through multiple linear regression techniques [20–22], physical-based models [23–25], or machine learning [20,26–28].

While the choice of the estimation approach, i.e., whether parametric or non-parametric, has mostly shown only a marginal impact on the estimation accuracy [20,29], retrieval approaches based on machine learning techniques have received more attention in the recent literature. This is due to their capability to handle diverse datasets and provide a reliable estimate of the biomass even at levels when the SAR predictors are reported to saturate [27,30–38]. Although the variance of the estimates was mostly limited, they were often erroneous at the tails of the biomass distribution, with over- and underestimation affecting the low and high biomass range, respectively. These issues were overcome in a few studies in which either the predicting features were first optimized with a pre-selection based on seasonal conditions [33] or forest types [38] or supported by a genetic algorithm [39]. Increasing the feature space of predictors with additional SAR backscatter values simulated with a Water Cloud Model was instead proven to stabilize the retrieval based on Neural Networks across sites [26].

One common observation in the cited literature on the effective use of non-parametric estimation models is that they typically require a densely populated dataset of biomass values that act as a reference to establish the relationship between the predictors and the biomass variable of interest. This aspect is crucial, especially when satellite measurements have limited sensitivity to forest biomass (as currently available satellite sensors are). In addition, these studies were undertaken at local sites. While a model trained on field reference data can predict biomass within forests for which the reference data are representative, the extension of the same model to dissimilar areas might be unreliable [26]. The scarcity of ground reference data on forest biomass for most regions of the world means that wide-area model-based biomass mapping from satellite imagery, e.g., continents, requires strong generalizations during the model calibration phase to prevent unrealistic biomass retrievals. Eventually, this results in strongly biased estimates of biomass [40,41]. The high spatial sampling and spatial resolution of LiDAR-based estimates of AGB from the Global Ecosystem Dynamics Investigation (GEDI) mission have recently been exploited to train retrieval models [25,38] and to derive nation-wide maps of AGB [35]. Although the results were not as accurate as when the retrieval model was trained with field plot measurements due to the superior quality of the latter, spaceborne LiDAR-based datasets may have a prominent role in future studies as their accuracy will improve.

As an alternative to calibrating the retrieval model with a training dataset of GSV or AGB values, the parameters of parametric models can be derived from a statistical analysis of the SAR image used as a predictor and masked for given forest conditions [23,24]. Despite several approximations in how the retrieval model is calibrated, the backscatter

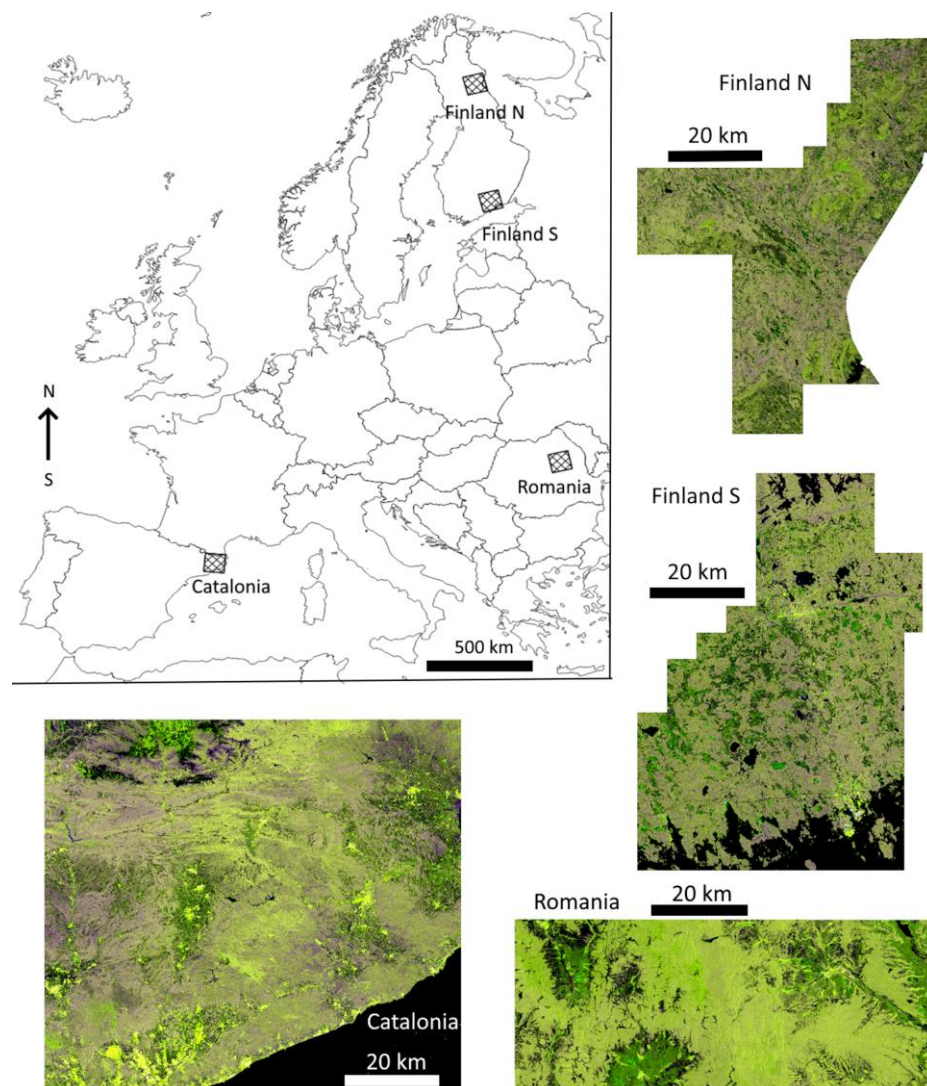
predicted by the model was within the range of the observations. These results suggested that the calibration method, referred to as BIOMASAR, could support large-scale mapping of biomass using spaceborne SAR backscatter observations at C- and L-bands [42] and thus overcoming the issues of limited availability of samples to train the retrieval model. Although BIOMASAR is used to create wall-to-wall estimates of GSV [43] and AGB [24,44], the retrieval performance with a calibrated model has been seldom explored elsewhere [45] and an assessment of factors that impact the calibration of the retrieval model is still missing.

In this study, we therefore compared the retrieval of forest GSV from SAR backscatter data acquired at C- and L-bands using the Water Cloud Model (WCM) trained on one side with field measurements and calibrated on the other side with the aid of auxiliary data products. The study was undertaken at multiple study sites representative of boreal, temperate, and Mediterranean forest landscapes in Europe to disentangle the effect of the training dataset, SAR backscatter measurements, forest structure, and modelling framework.

The focus is on GSV because it is a key parameter in the context of forest resource management in Europe [1] and major proxy of AGB and carbon stored in vegetation [46–50] to address global climate issues [51]. GSV represents the volume of the tree stems for all living species per unit area, including bark but excluding branches and stumps. The study sites and the data supporting our study are presented in Sections 2 and 3, respectively. For each site, plot-level GSV data from field inventories were used to evaluate the performance of the calibration and validate the GSV estimates (Section 3.1). The study in addition utilized sub-national averages of GSV values reported by National Forest Inventories (NFIs), which are easily accessible and provide information on the spatial distribution of GSV at a continental scale (Section 3.2). Sub-national averages were considered to simulate scenarios where GSV reference values, such as plot inventory data, are unavailable. The proposed framework for GSV retrieval foresaw the use of SAR backscatter acquired by Sentinel-1 (C-band) and Advanced Land Observing Satellite 2 (ALOS-2) Phased Array L-band SAR 2 (PALSAR-2) as predictors (Section 3.3) and canopy height metrics based on satellite LiDAR observations as an additional source to calibrate a set of forest structural functions that were integrated in the main retrieval model (Section 3.4). Our analysis focused on (i) understanding the relationship between SAR backscatter values and GSV, and (ii) discussing the training and the calibration of the WCM. By computing the accuracy of the GSV estimates from the trained and calibrated WCM with respect to the plot-level data, our aim was to assess the capability of a WCM to estimate GSV in a scenario when no reference data are available to train the model.

## 2. Study Sites

The study sites (Figure 1) included forests from major European ecological zones and different types of topographic conditions from gently undulating landscapes to mountains (Table 1). The Finland N and Finland S sites were located north of the Arctic Circle and by the Baltic Sea, respectively. At both sites, the forests consisted of semi-natural broadleaf and coniferous forest. The Finland N site was characterized by a hilly landscape with scattered open peatlands. Some of the hilltops had very sparse forests or no trees at all. The Finland S site was gently undulating and consisted of partly fragmented forest and agricultural landscapes. The Catalanian site was in the east of Catalonia, by the Mediterranean Sea. The forests consisted of semi-natural broadleaf and coniferous forest on hilly to mountainous landscape. The Romanian site was in the Carpathian Mountains. The forests consisted of semi-natural broadleaf (mainly beech and oak) and coniferous (mainly spruce and fir) species on hilly to mountainous landscape.



**Figure 1.** The location of the study sites. Each site is illustrated with a colour composite of Sentinel-1 imagery (Red: VV-polarized backscatter; Green: VH-polarized backscatter; Blue: difference in the VV- and VH-polarized backscatter).

**Table 1.** Summary of study sites and corresponding field measurements.

| Site<br>(Geographic Location) | Biome                                       | Topography           | Field Measurements |                   |                        |
|-------------------------------|---------------------------------------------|----------------------|--------------------|-------------------|------------------------|
|                               |                                             |                      | # Sample Plots     | Year of Inventory | Ownership              |
| Finland N                     | Boreal forests and taiga                    | Hilly                | 1004               | 2018              | Public                 |
| Finland S                     | Boreal forests and taiga                    | Gently undulating    | 1064               | 2018              | Public                 |
| Catalonia                     | Mediterranean woodlands, forests, and scrub | Hilly to mountainous | 663                | 2016              | Public                 |
| Romania                       | Temperate broadleaf and mixed forests       | Hilly to mountainous | 1306               | 2019              | Private (Tornator Oyj) |

### 3. Material

#### 3.1. Plot-Level GSV

Field measurements conducted by the Finnish Forest Centre (<https://www.metsakeskus.fi/en>, last accessed on 4 April 2024) at sample plots were used as a reference for the two Finnish sites. The plots were arranged in clusters, each containing six plots and chosen to cover the variation in forest types and site conditions. Three different plot radii were used in the

sampling: 9 m in young and advanced managed forests with relatively large tree densities; 12.62 m in forest with small stem densities but usually a large volume due to the mature development stage, and 5.64 m in seedling stands. For the Catalanian site, sample plot data measured by the Spanish NFI were used and consisted of circular sample plots with a radius varying from 5 to 25 m. For the Romanian site, field plot measurements were provided by a private company, which had contracted a field inventory on their forest area to support forest management decisions. The sample plots were circular and covered areas of 500 m<sup>2</sup> (radius of 12.6 m).

Among several features, the database of measurements from each site reported values for the basal area, tree height, diameter at breast height, growing stock volume, site type, and species proportions (pine, spruce, and broadleaf). The statistical distribution of the GSV from the field measurements is summarized in Table 2. The overall largest GSVs were measured in Romania. GSV at the two Finnish sites increased from north to south. The Catalanian site was mostly characterized by moderate-to-low GSV, with some patches of very high GSV.

**Table 2.** GSV statistics for each site.

| Site      | GSV (m <sup>3</sup> ha <sup>−1</sup> ) |                     |         |
|-----------|----------------------------------------|---------------------|---------|
|           | Average                                | Quartiles (1, 2, 3) | Min/Max |
| Finland N | 95                                     | 44/83/135           | 0/498   |
| Finland S | 157                                    | 47/129/233          | 0/751   |
| Catalonia | 108                                    | 55/92/144           | 1/480   |
| Romania   | 418                                    | 204/379/582         | 0/1677  |

### 3.2. Sub-National Average GSV Values

European NFIs regularly publish forest variable statistics at administrative unit levels, thus providing comprehensive information on the spatial distribution of biomass in Europe, which otherwise would not be attainable because the original measurements at the plot level are usually not accessible [52]. For this study, GSV average values were gathered from reports and table statistics published by the NFIs for 449 administrative or jurisdictional regions in 28 countries (Table S1). The reporting period is heterogeneous and spans the entire 2010 decade. An update with growth factors to the study years was not undertaken to unnecessarily bias the values since the effect of disturbances could not be accounted for.

### 3.3. SAR Dataset

The SAR dataset consisted of images acquired at the C-band (5.6 cm wavelength) by the European Space Agency (ESA) Sentinel-1 satellite and at the L-band (23 cm wavelength) by Phased Array L-band Synthetic Aperture Radar 2 (PALSAR-2) on Advanced Land Observing Satellite 2 (ALOS-2) operated by the Japan Aerospace Exploration Agency (JAXA, Tokyo, Japan). In correspondence with the area of each sample plot, we computed the area-weighted backscatter from the pre-processed Sentinel-1 and the ALOS-2 PALSAR-2 images.

#### 3.3.1. Sentinel-1 Dataset and Pre-Processing

The Sentinel-1 dataset consisted of all images acquired during the same year as the corresponding forest inventory (Table 1). Images were acquired in the Interferometric Wide Swath mode (IWS) in VV and VH polarization. For the sites of Catalonia, Finland N, and Romania, 59 dual-polarized pairs of images were available. For the site of Finland S, the dataset included 45 dual-polarized image pairs. The Sentinel-1 imagery consisted of detected amplitudes of the radar backscattered intensity and was obtained in ground-range geometry, with a pixel size of 10 m in the range and azimuth direction. Pre-processing of the SAR images followed a standard approach and consisted of (i) multi-looking to 20 m to be closer to the native spatial resolution of the data, (ii) terrain geocoding using the 30 m

Copernicus Digital Elevation Model (DEM) as a reference [53] with an automated matching procedure [54], and (iii) correction for the pixel area effect on the radar backscatter [55]. For each site, the Sentinel-1 images were clipped and geocoded to a common geometry and then filtered for speckle with a multi-channel approach [56].

### 3.3.2. ALOS-2 PALSAR-2 Dataset and Pre-Processing

The ALOS-2 PALSAR-2 dataset consisted of yearly mosaics of the radar backscatter at HH and HV polarization with a pixel size of 25 m × 25 m [57]. Mosaics have been generated by JAXA for each year since 2015.

The mosaics are based primarily on ALOS-2 Fine Beam Dual-polarization (FBD) images acquired between May and October of a given year. Each of the mosaics is provided in the form of 1° × 1° tiles and includes the (i) two backscatter images; (ii) an image of the local incidence angle with respect to the orientation of the pixel, derived from a DEM (3-arcsec Shuttle Radar Topography Mission (SRTM) or 1-arcsec ASTER DEM); (iii) a radar layover/shadow mask; (iv) an image of the date of acquisition; and (v) a mask detailing whether a pixel is labelled as land or water. Prior to mosaicking, the SAR images were geocoded, orthorectified, and calibrated by JAXA. The mosaics were compensated for variations in the pixel scattering area due to topography and for the dependence of backscatter on the local incidence angle [58]. Although the pre-processing of the ALOS-2 PALSAR-2 and the Sentinel-1 images followed the same sequence, the geocoding and terrain normalization procedures implemented different approaches. The visual inspection of the mosaics revealed partially uncompensated slope-induced effects on the backscatter. The SAR images were finally resampled to a pixel size of 1/4000th of a degree in both latitude and longitude, corresponding to roughly 25 m at the Equator.

For this work, we used version 2.1 of the mosaics and resampled them with bilinear interpolation to the geometry of the Sentinel-1 images. Besides the mosaic corresponding to the year of the field inventory data, we also considered the other mosaics to understand the impact of an L-band multi-temporal dataset on the estimation of GSV.

### 3.3.3. Uncertainty of the SAR Backscatter

The uncertainty of a SAR backscatter measurement was represented by a measure of its precision, i.e., its standard deviation. It was estimated empirically with the Equivalent Number of Looks (ENL) [59] where  $\mu$  represents the expected value of the SAR backscatter and  $\sigma$  its standard deviation.

$$ENL = \frac{\mu^2}{\sigma^2} \quad (1)$$

To compute the ENL, we superimposed a grid with a window size of 100 × 100 pixels on the stack of the SAR images and computed Equation (1) in each window and for each image. To avoid that heterogeneous landscapes or sloped terrain would impact the estimation, the ENL was defined as the 90th percentile of all values obtained for the grid cells. To account for the seasonal variability of C-band backscatter [42,60] and obtain an ENL representative of the Sentinel-1 dataset at a given site, we computed the median of the ENL estimates for the individual images.

Table 3 lists the estimates of ENL for each polarization and site for both sensors. The lower signal-to-noise ratio of the VH-polarized channel compared to the VV-polarized channel implied that the ENL of the Sentinel-1 VH-polarized data was higher than for VV polarization. The VV-polarized ENL of Sentinel-1 was rather constant between 4 and 6 except for the Finland N site where the estimate was slightly higher. The VH-polarized ENL for Sentinel-1 was also rather constant, between 11 and 14, except for Catalonia. For the ALOS-2 PALSAR-2 dataset, the ENL was mostly between 7 and 9, with the exception of Finland N where we estimated an ENL of 13 for the HH polarization. Compared to the Sentinel-1 dataset, the ENLs were lower. Although the ALOS-2 mosaics were obtained by multi-looking the original Single Look Complex images, the multi-channel filtering could not be applied.

**Table 3.** Estimates of the ENL from the Sentinel-1 and ALOS-2 PALSAR-2 datasets.

| Site      | ENL        |    |                 |    |
|-----------|------------|----|-----------------|----|
|           | Sentinel-1 |    | ALOS-2 PALSAR-2 |    |
|           | VV         | VH | HH              | HV |
| Finland N | 9          | 14 | 13              | 9  |
| Finland S | 6          | 13 | 8               | 7  |
| Catalonia | 4          | 7  | 8               | 9  |
| Romania   | 4          | 11 | 8               | 9  |

Finally, a single global value for each sensor was computed by taking the median of the site-specific estimates. For Sentinel-1, the ENL was set as equal to 6 for the VV polarization and 11 for the VH polarization, corresponding to a standard deviation of 1.5 dB and 1.1 dB, respectively. For the ALOS-2 PALSAR-2 dataset, the ENL was set as equal to 8 for both the HH and HV polarizations, corresponding to a standard deviation of 1.3 dB.

### 3.4. ICESat-2 Dataset and Pre-Processing

The Advanced Topographic Laser Altimeter System (ATLAS) onboard the Ice Cloud and Elevation Satellite (ICESat-2) is a photon-counting instrument that splits the laser into six beams. Beams are arranged in pairs approximately 3.3 km apart [61], each pair consisting of a strong and weak energy beam (4:1 ratio). Geophysical parameters related to vegetation and terrain heights are available in the ATL08 product as data segments with a 100 m step size along the flight direction [62]. From the original photon data, the original files were reformatted into segments of a 100 m length and 25 m width. Data are available starting at October 2018; for this study, observations from 2019 were used. To reduce the number of segments with potentially erroneous height metrics, we discarded segments that were (i) acquired by the weak beams because of the very small number of photons recorded from the forest floor, in particular in dense canopies; (ii) characterized by less than three photons reflected by the canopy; (iii) flagged as not belonging to natural vegetation in the metadata; and (iv) exhibited an elevation differing by more than 25 m from the reference DEM used in all ATL products. In addition, values larger than 50 m were filtered out because they frequently occurred in sparsely vegetated or unvegetated regions. As suggested by the product developers, the canopy height was represented by the 98th percentile of all height values within a segment (RH98).

From this dataset, an estimate of the maximum canopy height was computed for each site to constrain the retrieval of GSV (Table 4). This parameter was set as equal to the 99th percentile of all valid canopy height values to avoid that individual unrealistic values would bias the estimate if set as equal to the largest of all canopy height values. The difference between the ICESat-2-based values and the maximum canopy height derived from the field inventory dataset (Table 4) was small at Finland N, Catalonia, and Romania. At Finland S, the discrepancy was due to the different sampling in either the ICESat-2 dataset or the plot dataset, respectively, in correspondence with the tallest forests. The uncertainty of the maximum height was assessed by computing the same value with ICESat-2 data acquired in 2020 and 2021. The difference between the estimates of maximum canopy height for the three years was less than 1 m so that the uncertainty associated with this variable was considered negligible in this study.

**Table 4.** Maximum canopy height for each site from the field inventory data and the ICESat-2 metric.

| Site      | Field Inventory |         | ICESat-2   |            |
|-----------|-----------------|---------|------------|------------|
|           | Height [m]      | # Plots | Height [m] | # Segments |
| Finland N | 23              | 1005    | 21         | 26,049     |
| Finland S | 35              | 1065    | 28         | 18,867     |
| Catalonia | 33              | 663     | 30         | 26,751     |
| Romania   | 44              | 1306    | 42         | 10,900     |

## 4. Methods

### 4.1. Forest Backscatter Model

The relationship between the SAR backscatter and GSV was expressed with the Water Cloud Model with gaps [63] integrated with forest structural functions [64]. The original WCM described the backscattered intensity from a forest as a function of the backscatter from the forest floor through gaps in the canopy, the backscatter from the forest floor attenuated by the canopy, and the backscatter from the canopy.

$$\sigma_{for}^0 = (1 - \eta)\sigma_{gr}^0 + \eta\sigma_{gr}^0 T_{tree} + \eta\sigma_{veg}^0 (1 - T_{tree}) \quad (2)$$

The forest backscatter was expressed as a function of two structural parameters describing horizontal and vertical properties of the vegetation. The variable  $\eta$  represented the canopy density from a microwave perspective; as a first approximation, we considered it equal to the canopy density in the optical domain even though the gaps should decrease for an increasing wavelength. The two model parameters  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  represent the backscattering coefficient of the ground and vegetation layer, respectively.  $T_{tree}$  represents the two-way tree transmissivity and was expressed in terms of a two-way attenuation per metre through a tree canopy,  $\alpha$ , and the depth of the attenuating layer, which was assumed to correspond to the canopy height,  $h$ .

$$T_{tree} = e^{-\alpha h} \quad (3)$$

To express the backscatter from a forest as a function of GSV, we used two structural functions relating canopy density, canopy height, and GSV. Using spaceborne LiDAR metrics of canopy height,  $h$ , and density,  $CD$ , the exponential model in Equation (4) was demonstrated as a reliable fit to the data both at regional [64] and global [65] scales. The coefficient  $q$  was of an empirical nature.

$$CD = 1 - e^{-qh} \quad (4)$$

The power-law relationship between canopy height,  $h$ , and GSV,  $V$ , in Equation (5) with  $a$  and  $b$  being coefficients of an empirical nature was demonstrated when relating either field measurements or LiDAR-based metrics of canopy height to GSV [63,64].

$$V = a \cdot h^b \quad (5)$$

With  $\eta = CD$ , the integrated GSV retrieval model was therefore expressed as

$$\begin{aligned} \sigma_{for}^0 = & \sigma_{gr}^0 \left( e^{-q(aV)^b} + e^{-\alpha(aV)^b} - e^{-(q+\alpha)(aV)^b} \right) \\ & + \sigma_{veg}^0 \left( 1 - e^{-q(aV)^b} - e^{-\alpha(aV)^b} + e^{-(q+\alpha)(aV)^b} \right) \end{aligned} \quad (6)$$

### 4.2. Estimation of Model Parameters

The forest backscatter model in Equation (6) contained six parameters. The structural parameters  $q$ ,  $a$ , and  $b$  were estimated from a least squares regression between LiDAR-

based metrics of canopy density, canopy height, and GSV. The WCM parameters  $\sigma_{gr}^0$ ,  $\sigma_{veg}^0$ , and  $\alpha$  described the interaction of the microwave signal with the woody vegetation and the soil underneath and were estimated from the SAR backscatter data following different strategies.

The estimation of the coefficient  $q$  in Equation (4) required values of canopy density and canopy height as, for example, obtained from the ICESat LiDAR measurements [66]. Here, we used the spatially explicit dataset of the coefficient presented in [66] and associated with each site and the most frequent estimate of  $q$ . A site-specific estimate based on footprint-level data was discarded because of the sparse sampling by ICESat. The use of more recent LiDAR observations was not considered because it either would not cover all sites (Global Ecosystem Dynamics Investigation, GEDI) or would not provide a suitable canopy density metric (ICESat-2).

The coefficients of the structural function in Equation (5) were estimated by regressing to the sub-national averages of GSV and the corresponding averages of LiDAR-based canopy height. To account for the spatial variability of the association between height and volume, the equation was fitted for each of the four largest terrestrial biomes in Europe according to [67], i.e., “temperate broadleaf and mixed forests”, “temperate conifer forests”, “boreal forests and taiga”, and “Mediterranean forests, woodlands, and scrub”. The smaller biomes “temperate grasslands, savannas, and scrublands” and “tundra” were merged with temperate broadleaf forests and boreal forests, respectively. Estimates of the two coefficients for the respective biome were assigned to each site accordingly (Table 1). Although measurements of forest height and GSV were available for each plot and at each site, we did not create site-specific estimates because this study aimed at assessing the performance of the retrieval model in a scenario where training data in the form of plot measurements are not available.

Having set the two structural functions’ set, the three unknown parameters of the WCM are estimated. Traditionally, the WCM is regressed to a training set of GSV measurements and corresponding backscatter values. In practice, two major issues are associated with the regression.

1. The weak sensitivity of the C- and L-band backscatter to GSV implies that a model fit with three degrees of freedom may be characterized by large uncertainties. It is also likely that some of the physical constraints of the three model parameters are violated so that they may become pure regression parameters. This applies especially to the coefficient  $\alpha$  that governs the rate of change of the backscatter with GSV.
2. When the training data are not available or the area of interest is far away from the location of the training data, it is likely that the estimates of the three parameters are not representative of the area.

To overcome these issues, the estimation of the three parameters was performed in two steps. Firstly, the coefficient  $\alpha$  was estimated. Values and estimates of tree transmissivity reported in the literature are scarce and uncertain [68]. The only alternative was to rely on the generic values reported in the literature and corroborate them with local estimates by means of a least squares’ regression of the WCM with a training set of volume and backscatter measurements to capture variations dependent on the frequency, polarization, look angle, and seasonal conditions. Given an estimate of  $\alpha$ , the other two parameters,  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$ , were estimated either with a regression to a training dataset or with calibration using auxiliary data.

In absence of a training dataset, one approach to estimate the model parameters  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  was based on image statistics in areas where backscatter stems from either unvegetated or densely vegetated terrain, respectively [23]. Since  $\sigma_{gr}^0$  represented the backscatter from an unvegetated surface, it was reasonable to assume that an estimate might be derived from the backscatter measurements observed over surfaces that are unvegetated. A mean or median value from such measurements was then associated with  $\sigma_{gr}^0$ . The parameter  $\sigma_{veg}^0$  represented the backscatter of the fully opaque canopy. The backscatter for the densest forests within the area of interest was the type of measurement

closest to the definition of the model parameter. However, the measurements had to be corrected for an offset that removed the residual ground contribution since the measured backscatter contained a contribution from the soil underneath the canopy that is visible through the canopy gaps or is attenuated by the canopy.

The compensation of the backscatter for such an offset used the model in Equation (1) and expressed  $\sigma_{veg}^0$  as a function of the backscatter representative of dense forests,  $\sigma_{df}^0$ ; the canopy density of dense forests,  $\eta_{df}$ ; and the corresponding canopy height of dense forests,  $h_{df}$ , which is contained in the tree transmissivity term  $T_{tree,df}$ , together with the attenuation coefficient  $\alpha$  and the backscatter coefficient  $\sigma_{gr}^0$ .

$$\sigma_{veg}^0 = \frac{\sigma_{df}^0 - \left(1 - \eta_{df} + \eta_{df} T_{tree,df}\right) \sigma_{gr}^0}{\eta_{df} (1 - T_{tree,df})} \quad (7)$$

In previous developments,  $\sigma_{gr}^0$  and  $\sigma_{df}^0$  were estimated by selecting pixels in correspondence with sparse and dense forest cover according to a map of canopy density, e.g., [69], and calculating the mean observed backscatter in such areas, respectively. The estimation of  $\sigma_{veg}^0$  from  $\sigma_{df}^0$  with Equation (7) required an estimate of the canopy density and canopy height of dense forests. In [43], the values corresponded to the 90th percentile of the histogram of each variable derived from waveform data acquired by ICESat GLAS. This percentile minimized the difference in the estimate of  $\sigma_{veg}^0$  from Equation (7) and the estimate of  $\sigma_{veg}^0$  when fitting the backscatter model to field measurements of biomass. The empirically defined sensor- and region-specific thresholds to delineate sparse and dense forest cover and a definition of  $\eta_{df}$  and  $h_{df}$  based on LiDAR data, independent of the definition of  $\sigma_{df}^0$ , implied a risk of obtaining biased estimates of  $\sigma_{veg}^0$  and  $\sigma_{gr}^0$  locally.

To reduce the impact of uncertainties on the estimation of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$ , we investigated here an alternative calibration approach that regressed between backscatter measurements and corresponding canopy density values with Equation (1). To express the backscatter as a function of canopy density only, the structural function in Equation (4) was first inverted. By substituting canopy height with a function of canopy density in Equation (1), the WCM was rewritten as

$$\sigma_{for}^0 = (1 - \eta) \sigma_{gr}^0 + \eta \sigma_{gr}^0 e^{\alpha \frac{\log(1-\eta)}{q}} + \eta \sigma_{veg}^0 \left(1 - e^{\alpha \frac{\log(1-\eta)}{q}}\right) \quad (8)$$

This simplification was motivated by the availability of canopy density map products [69,70], a variable that is closely related to observations by optical remote sensing. The advantage of this calibration approach over the previous one was the independence from empirical thresholds used to select “ground” and “dense forest” pixels. The caveat was that uncertainties affecting the canopy density observations would have propagated to the estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$ .

According to its definition, the estimate of  $\sigma_{veg}^0$  from the regression of Equation (8) to observations of the SAR backscatter and canopy density expressed the backscatter of a canopy density of 100%. However, a canopy density of 100% corresponded to a large range of canopy heights and GSV values. Assuming that the backscatter for a canopy density of 100% was somewhat sensitive to forest structure, the initial estimate of  $\sigma_{veg}^0$  from the regression was offset with a term here represented by the standard deviation of the SAR backscatter for the largest canopy density.

$$\sigma_{veg}^0 = \hat{\sigma}_{veg}^0 + 2 * SD_{CD=100\%,meas} \quad (9)$$

In Equation (9),  $\hat{\sigma}_{veg}^0$  represents the initial estimate of  $\sigma_{veg}^0$  obtained from the regression and  $SD_{CD=100\%,meas}$  represents the standard deviation of the backscatter for a canopy density of 100%. The factor 2 was used to simulate a confidence interval from the stan-

dard deviation; herewith, it was assumed that the backscatter measurements for a given canopy density range were normally distributed. Since the observed standard deviation consisted of both the measurement noise and signal related to biomass variability, the term  $SD_{CD=100\%}$  was obtained by subtracting the measurement noise from the observed value  $SD_{CD=100\%, meas}$ :

$$SD_{CD=100\%} = \sqrt{(SD_{CD=100\%, meas})^2 - \frac{\hat{\sigma}_{veg}^0{}^2}{ENL}} \quad (10)$$

The measurement noise required the original measurement representative of the backscatter for a canopy density of 100%,  $\hat{\sigma}_{veg}^0$ , and the ENL.

#### 4.3. Retrieval of GSV

GSV was estimated from the SAR backscatter with the WCM in Equation (6). Since this form of the WCM could not be inverted mathematically, an estimate of GSV was obtained by minimizing the difference between the measured backscatter and the modelled backscatter.

Because of the loss of sensitivity of the backscatter for increasing GSV, the retrieval was constrained within a range of values between 0 and an estimate of the maximum GSV so that unnatural values were avoided [24]. An estimate larger than the maximum GSV was replaced with the maximum GSV value. Likely, a negative estimate of GSV was replaced with 0. In this study, we constrained the retrieval to a modelled value,  $V_{max}$ , using Equation (11). The first term represented the value obtained from the maximum canopy height,  $h_{max}$ , and Equation (5). As this value corresponded to an average value of several potential maximum GSVs, we introduced the second term that accounted for the range of maximum GSV estimates. This term was expressed with the standard deviation of the maximum GSV estimate corresponding to the maximum canopy height,  $\delta V(h_{max})$ .

$$V_{max} = a \cdot (h_{max})^b + 2 * \delta V(h_{max}) \quad (11)$$

With several backscatter measurements from one sensor, the corresponding estimates of GSV were then combined to form a new estimate that averaged out individual noises. The weighted average in Equation (12) combined the individual GSV estimates from SAR data of a given sensor and defined the weights as the difference between the backscatter parameters of the WCM,  $w_i = \sigma_{veg}^0 - \sigma_{gr}^0$  [23].

$$GSV_{mt} = \frac{\sum_{i=1}^N w_i GSV_i}{\sum_{i=1}^N w_i} \quad (12)$$

To alleviate systematic deficiencies in GSV estimates from multiple sensors, the sensor-specific estimates of GSV were eventually combined with an additional weighting scheme [43] where the definition of the weights accounted for (i) differences in the sensitivity of C- and L-band data to GSV and (ii) residual topographic effects in the ALOS-2 PALSAR-2 mosaics.

The first weighting scheme accounted for the sensitivity of the backscatter to GSV at a given frequency. It was modelled with the transmissivity term in Equation (1).

$$T = 1 - \eta + \eta \sigma_{gr}^0 \quad (13)$$

A simple weighting scheme that reflected the difference in sensitivity between C- and L-bands consisted of differencing in the derivatives of Equation (13):

$$w_s = \left| \frac{\partial T_L}{\partial V} \right| - \left| \frac{\partial T_C}{\partial V} \right| \quad (14)$$

where the derivatives of the transmissivity at L- and C-bands,  $T_L$  and  $T_C$ , were determined using the structural functions in Equations (4) and (5). Equation (14) was then constrained with the number of images available at L- and C-bands,  $N_L$  and  $N_C$ , respectively:

$$w_s = \left| \frac{\partial T_L}{\partial V} \sqrt{N_L} \right| - \left| \frac{\partial T_C}{\partial V} \sqrt{N_C} \right| \quad (15)$$

This weight was finally rescaled to the range 0 to 1 to obtain normalized weights:

$$w_{s,norm} = \frac{w_s - \min(w_s)}{\max(w_s - \min(w_s))} \quad (16)$$

The second weighting scheme was introduced because the effect of the pixel scattering area on the SAR backscatter was compensated for differently in the C- and L-band datasets. The additional weight was defined by calculating the percent difference in pixel scattering area between flat terrain (i.e., for the nominal incidence angle of  $38^\circ$  of ALOS-2 PALSAR-2) and the pixel area estimated from the local incidence angle,  $\theta_i$ .

$$\Delta_{area} = 100 \cdot \left| 1 - \frac{\sin(38^\circ)}{\sin \theta_i} \right| \quad (17)$$

The weight,  $w_{topo}$ , was then obtained by linearly scaling  $\Delta_{area}$  between 0 and 1 with 1 representing flat terrain (i.e., for a  $38^\circ$  incidence angle) and 0 reflecting  $\Delta_{area}$  values beyond 30, i.e., for steep slopes.

The final weight for the L-band was then the product of  $w_s$  and  $w_{topo}$ . Accordingly, the weight for the C-band was the complement of the L-band weight.

$$w(L) = w_{s,norm} \cdot w_{topo} \quad (18)$$

$$w(C) = 1 - w(L) \quad (19)$$

The final estimate of GSV,  $GSV_{fin}$ , was obtained from the C- and L-band GSV estimates,  $GSV_{mt,C}$  and  $GSV_{mt,L}$ , with Equation (20):

$$GSV_{fin} = w(L) \cdot GSV_{mt,L} + w(C) \cdot GSV_{mt,C} \quad (20)$$

#### 4.4. Implementation of GSV Retrieval and Validation

To train the WCM, the field measurements of GSV were sorted and every other sample was selected. The remaining samples formed the test set. With this procedure, we ensured that both sets had the same statistical distribution of GSV. The test set was used to validate GSV estimated from both the trained and the calibrated WCM.

Equation (12) combined the GSV estimates from the Sentinel-1 images and from several combinations of the six ALOS-2 PALSAR-2 mosaics separately. After that, a final estimate of GSV was obtained for each of the two approaches using Equation (20). All retrievals were assessed against the field measurements of GSV in the test set. The retrieval error was quantified with

- The root mean square error (RMSE);
- The RMSE relative to the mean value of the GSV measurements in the test set;
- The bias, i.e., the systematic difference between estimated and reference GSV values;
- The coefficient of determination ( $R^2$ ).

## 5. Results

### 5.1. Correlation Between Radar Backscatter and GSV

Pearson's coefficient of correlation between the plot-level GSV values and the backscatter observations was computed for each SAR image to understand the strength of the relationship between the two covariates (Figures S1 and S2). Statistics derived from the

correlation coefficients are reported in Table 5 for the Sentinel-1 dataset and in Table 6 for the ALOS-2 PALSAR-2 dataset. Although this was not a rigorous way to assess the strength of the relationship because Pearson's coefficient assumes collinearity between two variables while the backscatter at C- and L-bands loses sensitivity to GSV for increasing GSV, it provided indications on the theoretical performance of the GSV retrieval in space and time.

**Table 5.** Pearson's coefficient of correlation between plot inventory GSV and Sentinel-1 backscatter. For each site and polarization, the mean, minimum, and maximum correlation values across the Sentinel-1 dataset are reported.

| Site      | VV   |       |      | VH   |       |      |
|-----------|------|-------|------|------|-------|------|
|           | Mean | Min   | Max  | Mean | Min   | Max  |
| Finland N | 0.50 | 0.38  | 0.73 | 0.52 | 0.42  | 0.65 |
| Finland S | 0.15 | −0.13 | 0.40 | 0.14 | −0.04 | 0.38 |
| Catalonia | 0.37 | 0.26  | 0.48 | 0.42 | 0.25  | 0.54 |
| Romania   | 0.17 | 0.08  | 0.26 | 0.14 | 0.01  | 0.31 |

**Table 6.** Pearson's coefficient of correlation between plot inventory GSV and ALOS-2 PALSAR-2 backscatter. For each site and polarization, the mean, minimum, and maximum correlation values across the ALOS-2 PALSAR-2 dataset are reported.

| Site      | HH   |      |      | HV   |      |      |
|-----------|------|------|------|------|------|------|
|           | Mean | Min  | Max  | Mean | Min  | Max  |
| Finland N | 0.31 | 0.24 | 0.42 | 0.51 | 0.44 | 0.62 |
| Finland S | 0.42 | 0.29 | 0.54 | 0.48 | 0.36 | 0.60 |
| Catalonia | 0.21 | 0.17 | 0.25 | 0.35 | 0.34 | 0.39 |
| Romania   | 0.15 | 0.13 | 0.20 | 0.23 | 0.19 | 0.25 |

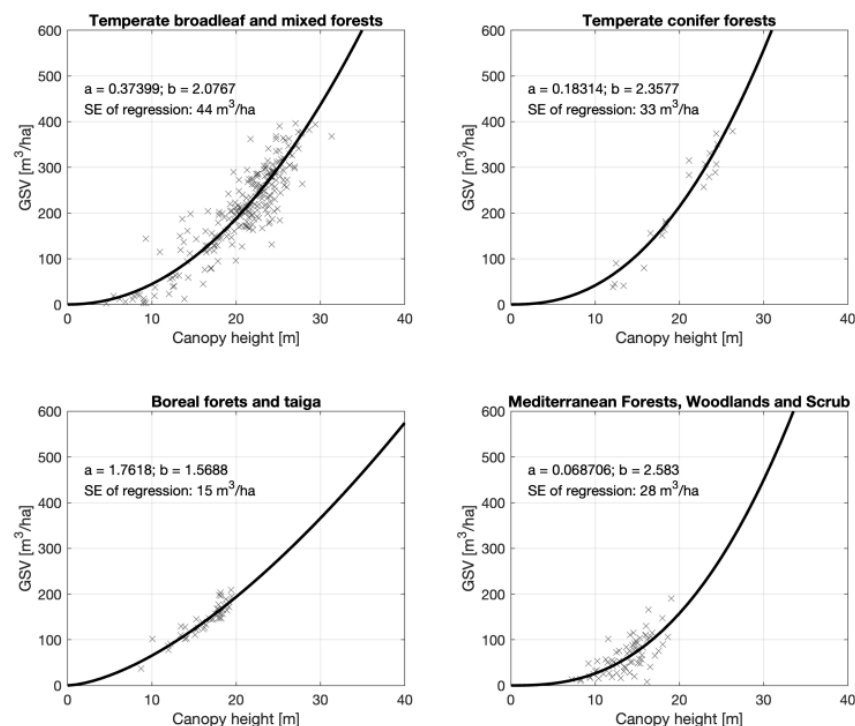
Except for images acquired in early spring over Finland N, the correlation between the Sentinel-1 backscatter and GSV was low (i.e.,  $<0.3$ ) to moderate (i.e., between 0.3 and 0.6) (Table 5 and Figure S1). While Finland N was characterized by the largest correlation, Finland S was characterized by the lowest correlation among all sites. This result was explained by the larger range of GSVs and the wet conditions prevalent at this site, as both factors negatively impact the correlation between biomass and SAR backscatter [19]. In addition, Finland S was the only site where different levels of correlation for ascending and descending passes (i.e., night-time and daytime) were observed (Figure S1). The night-time data were all characterized by correlation coefficients close to 0 whereas the correlation coefficients for the images acquired during daytime were higher. As for the two Finnish sites, the correlation coefficients in Catalonia were similar at VV and VH polarization (Table 5), with some seasonality occurring at VH polarization (Figure S1). On the contrary, in Romania, the correlation coefficient was low regardless of polarization because of the large proportion of plots with high GSV.

The correlation coefficients for the ALOS-2 PALSAR-2 dataset were either low or moderate (Table 6) with a few exceptions in Finland S (Figure S2). Compared to the C-band values, the correlation coefficients were higher at Finland S whereas they were lower in Finland N and Catalonia. The images used by JAXA to create the ALOS-2 PALSAR-2 mosaics over these sites were acquired in October. For the Finland N site, the weak correlation was a consequence of the air temperature being around 0 °C, this being consistent with previous studies [71]. In Catalonia, the weather was mostly dry and warm, thus not explaining the weaker correlation at the L-band compared to C-band observations. The visual analysis of the ALOS-2 PALSAR-2 mosaics revealed imperfections caused by

sloped terrain, leading to strong variation in backscatter observations in relation to the GSV measurements. The low correlation between GSV and ALOS-2 PALSAR-2 backscatter for the Romanian site was explained by the overall high GSV.

### 5.2. Estimation of Parameters of GSV Structural Function

Figure 2 shows the regression of Equation (5) to the measurements of average canopy height from the ICESat-2 dataset and the corresponding GSV value for 377 NFI units in 26 countries. Data from Switzerland, Austria, Ireland, and Great Britain were discarded because for each of these countries, the correlation between canopy height and GSV values was smaller than 0.3. Topography in Austria and Switzerland and the fragmented forest landscape in all four countries except Austria likely contributed to observed weak correlation between forest variables.



**Figure 2.** Canopy height from ICESat-2 data averaged at the level of sub-national units and corresponding average GSV values together with the fit of Equation (5) after stratifying by forest biome. Estimates of the coefficients  $a$  and  $b$  in Equation (5) and the standard error of the regression are visualized in the upper left corner of each panel.

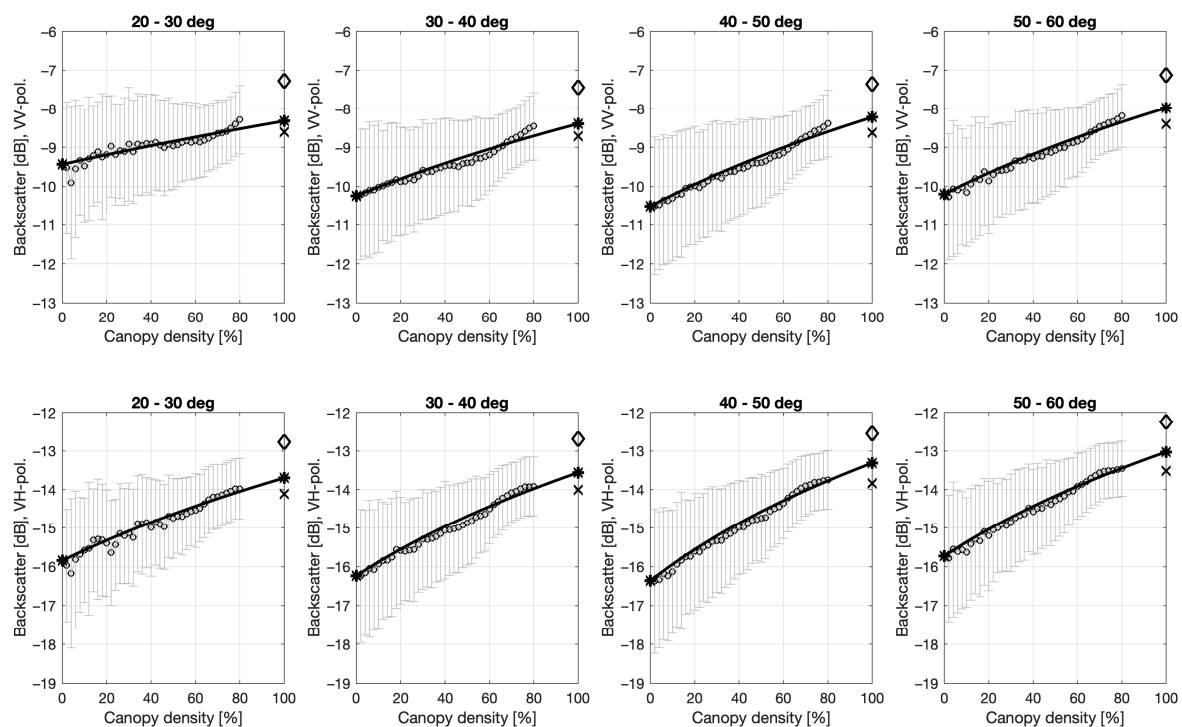
### 5.3. Estimation of the Water Cloud Model Parameters

The coefficient  $\alpha$  was obtained by allowing it to vary between a certain range of values and fitting each time the WCM to the observations of SAR backscatter and GSV. Based on experimental results reported in the literature, a plausible range of values between 0.1 dB and 2 dB was used [72,73]. For each value of  $\alpha$ , the sum of the model residuals was computed for each SAR image at a given site. The value corresponding to the smallest of the cumulative residuals was selected as an estimate of  $\alpha$ .

At the C-band, the procedure was strongly affected by the weak correlation between GSV and backscatter. The cumulative residuals showed no dependence on  $\alpha$  except for the Finland N site where the sum of the residuals decreased as  $\alpha$  increased. Since an estimate of 2 dB/m was considered plausible in previous studies [23], it was adopted regardless of the polarization, acquisition date, or site. At the L-band, we identified local minima in the cumulative residuals as a function of  $\alpha$ , which were site-specific, being 0.8 dB/m in Finland S and Catalonia, and 1 dB/m in Finland N. Because of the very weak correlation between GSV and backscatter in Romania, the estimate of 0.5 dB/m was considered unreliable. As a

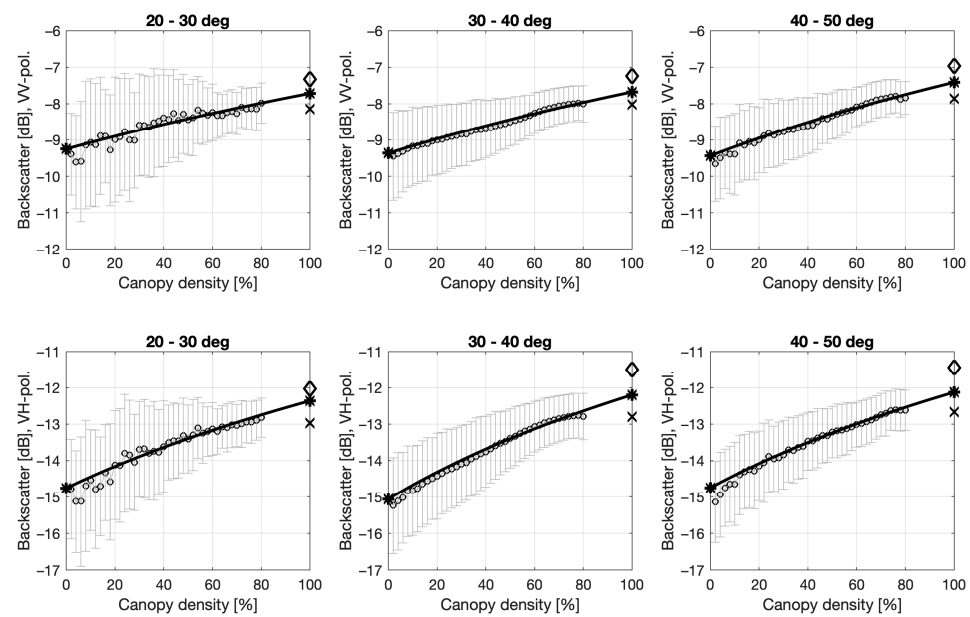
first-order approximation, we finally assigned  $\alpha$  the average of the site-specific values, i.e., 0.9 dB/m.

Figures 3 and 4 show examples of measured and modelled backscatter as a function of canopy density for a pair of Sentinel-1 images covering the sites with the highest correlation between backscatter and GSV, i.e., Catalonia and Finland N. The relationship between the backscatter and canopy density differed depending on the local incidence angle. The sensitivity increased from steep to shallow angles as the thickness of the vegetation layer sensed by Sentinel-1 grew and the backscatter from the ground beneath the canopy diminished. Although the variability in the backscatter observations was substantial, the mean value per canopy density level presented a clear trend, which was captured by the WCM in Equation (8). Compared to the value representative of the backscatter for dense forests, the value of  $\hat{\sigma}_{veg}^0$  obtained from the regression of the WCM to the data was slightly higher.

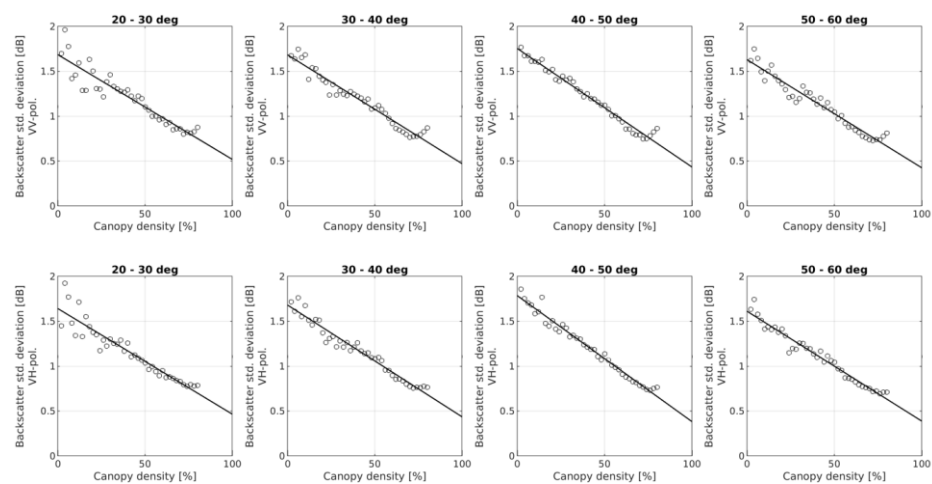


**Figure 3.** Measured and modelled Sentinel-1 VV- and VH-polarized backscatter over the Catalanian site stratified by the local incidence angle and illustrated as a function of the canopy density level (circles: average value; vertical bars: two-sided one standard deviation). The Sentinel-1 image was acquired on 17 July 2016. The asterisks at the canopy densities of 0% and 100% represent the estimates of  $\sigma_{gr}^0$  and  $\hat{\sigma}_{veg}^0$  obtained by regressing Equation (1) to the observations. The diamond and cross symbols at 100% canopy density represent the estimate of  $\sigma_{veg}^0$  obtained by compensating for  $\hat{\sigma}_{veg}^0$  for the standard deviation of the observations and the backscatter of dense forests, respectively (see also Figure 5).

To estimate the offset term  $SD_{CD=100\%}$  in Equation (9), a model was fitted to canopy density and backscatter standard deviation per canopy density level. Figures 5 and 6 show the relationship between canopy density and backscatter standard deviation for the same Sentinel-1 images used in the displays of Figures 3 and 4, respectively. The standard deviation decreased monotonically with canopy density, and the use of a linear function was deemed reasonable.

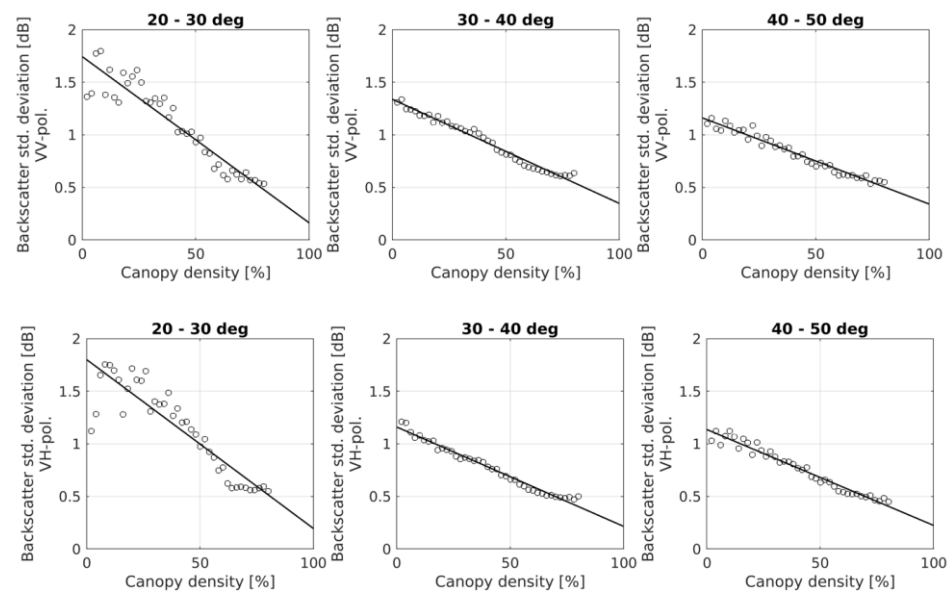


**Figure 4.** Measured and modelled Sentinel-1 VV- and VH-polarized backscatter over the Finland N site as a function of canopy density. The Sentinel-1 image was acquired on 12 July 2018. Plot notations are the same as in Figure 3.

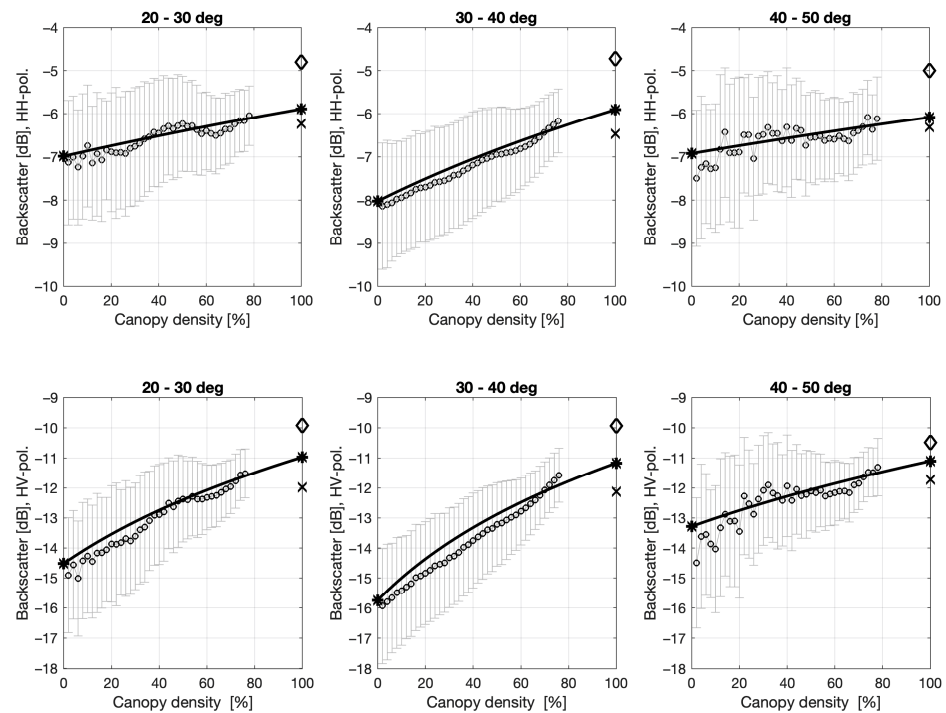


**Figure 5.** The standard deviation of the VV- and VH-polarized backscatter observations per canopy density level (circles) and linear regression (solid line) for the Sentinel-1 image acquired over the Catalanian site on 17 July 2016.

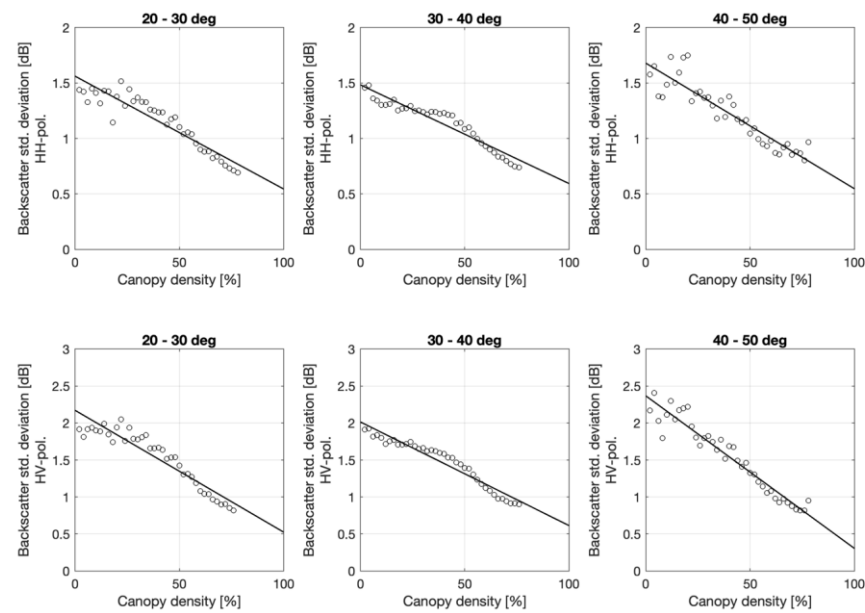
For the ALOS-2 PALSAR-2 data, Figure 7 shows measured and modelled backscatter as a function of canopy density for the Finland N site where the correlation between GSV and SAR backscatter was the largest among all study sites. And Figure 8 shows the standard deviation of the backscatter modelled as a function of canopy density for the same site. The sensitivity of the backscatter to canopy density was strongest for the 30–40 deg incidence angle interval, regardless of polarization (2 dB at HH and 4 dB at HV polarization). The WCM reproduced the trend of the measurements. The correction applied to the estimate of  $\hat{\sigma}_{veg}^0$  was approximately 1 dB. This is a result of compensating for the standard deviation of the observations for a canopy density of 100%, which was indeed slightly larger than 0.5 dB (Figure 8), with the ENL-related term.



**Figure 6.** The standard deviation of the VV- and VH-polarized backscatter observations for the Sentinel-1 image acquired over the Finland N site on 12 July 2018 at VV and VH polarization. Plot notations follow Figure 5.

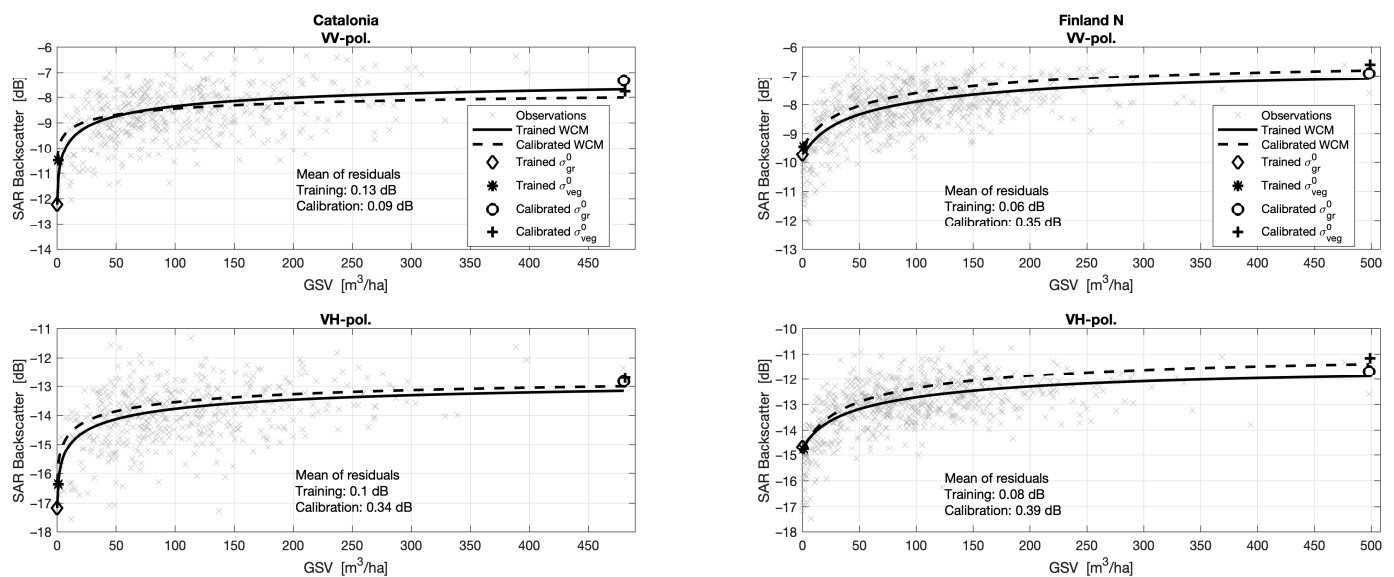


**Figure 7.** Measured and modelled ALOS-2 PALSAR-2 HH- and HV-polarized backscatter over the Finland N site stratified by the local incidence angle and illustrated as a function of the canopy density level (circles: average value; vertical bars: two-sided one standard deviation). The asterisks at canopy densities of 0% and 100% represent the estimates of  $\sigma_{gr}^0$  and  $\hat{\sigma}_{veg}^0$  obtained by fitting Equation (1) to the observations. The diamond and cross symbols at 100% canopy density represent the estimate of  $\sigma_{veg}^0$  obtained by compensating for  $\hat{\sigma}_{veg}^0$  for the standard deviation of the observations and the backscatter of dense forests, respectively.



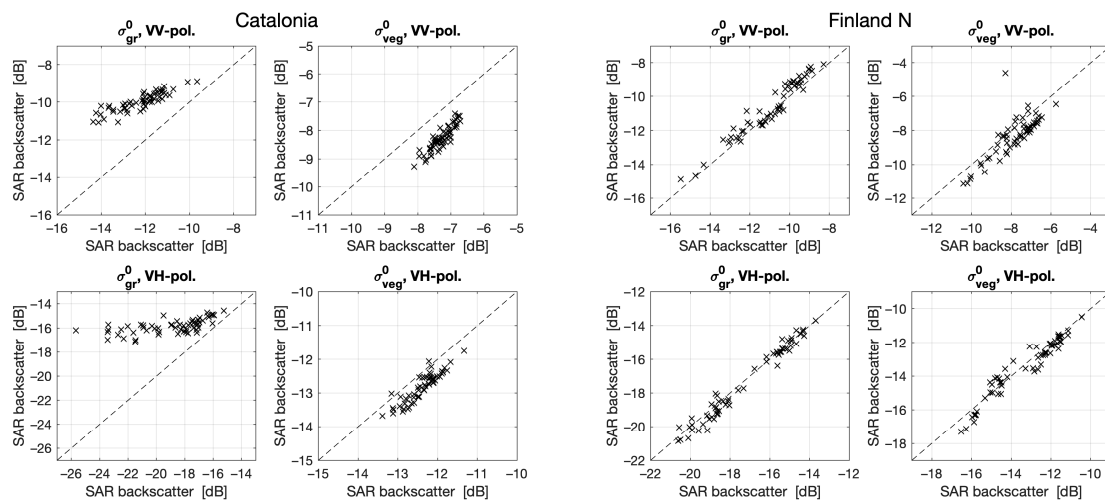
**Figure 8.** The standard deviation of the backscatter observations per canopy density level (circles) and linear regression (solid line) for the ALOS-2 PALSAR-2 HH- and HV-polarized backscatter acquired over the Finland N site.

Figure 9 shows an example of trained and calibrated WCM for the same Sentinel-1 images over Catalonia and Finland N used in Figures 3 and 4. Each panel also reports the mean value of the model residuals obtained with training and calibration. A comparison of all Sentinel-1 estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  from Catalonia and Finland N obtained with the two model-fitting procedures is illustrated in Figure 10.

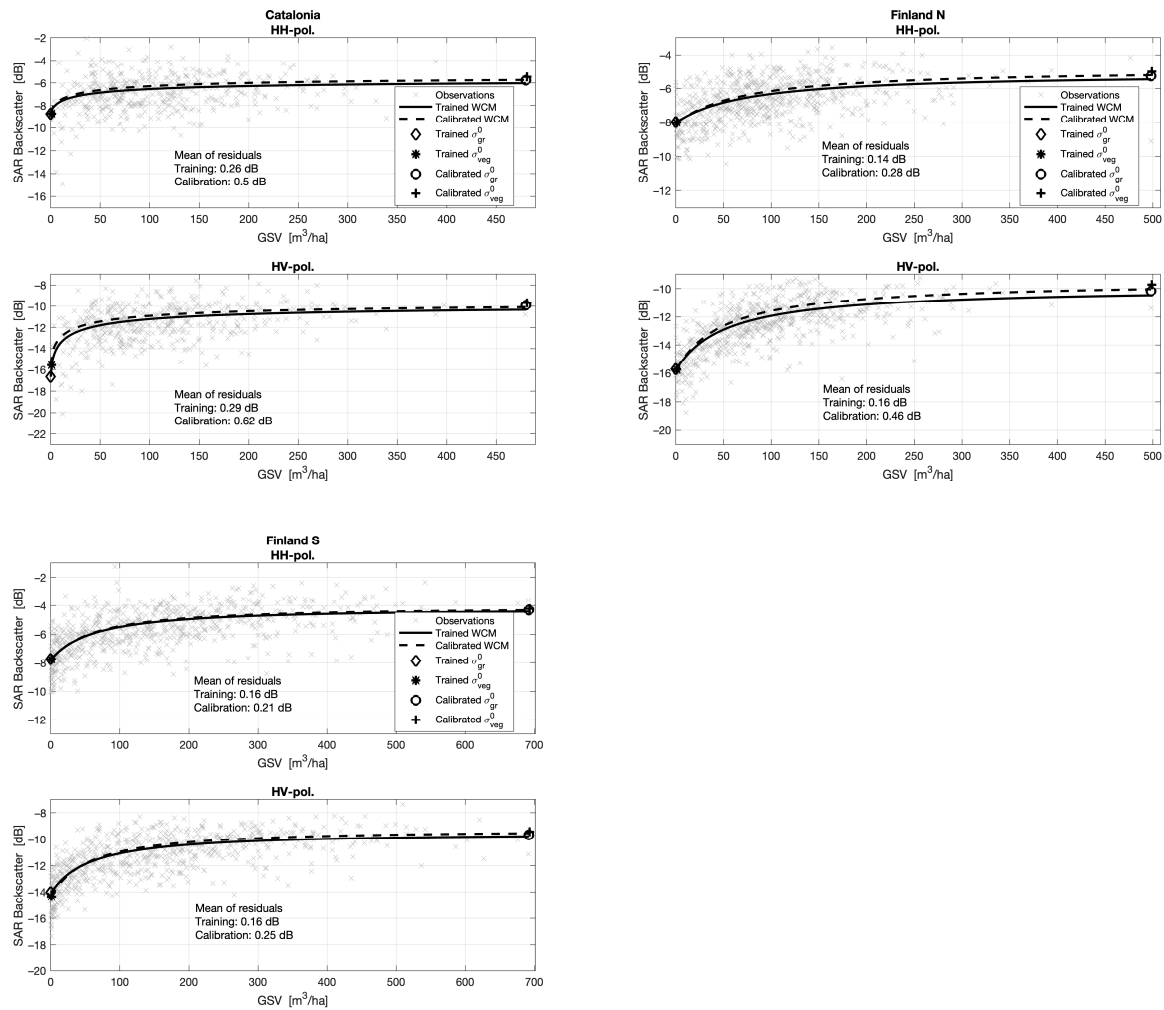


**Figure 9.** Measured and modelled VV- and VH-pol. backscatter as a function of GSV for the sites of Catalonia (left panels) and Finland N (right panels) for the Sentinel-1 dataset used in Figures 3 and 4.

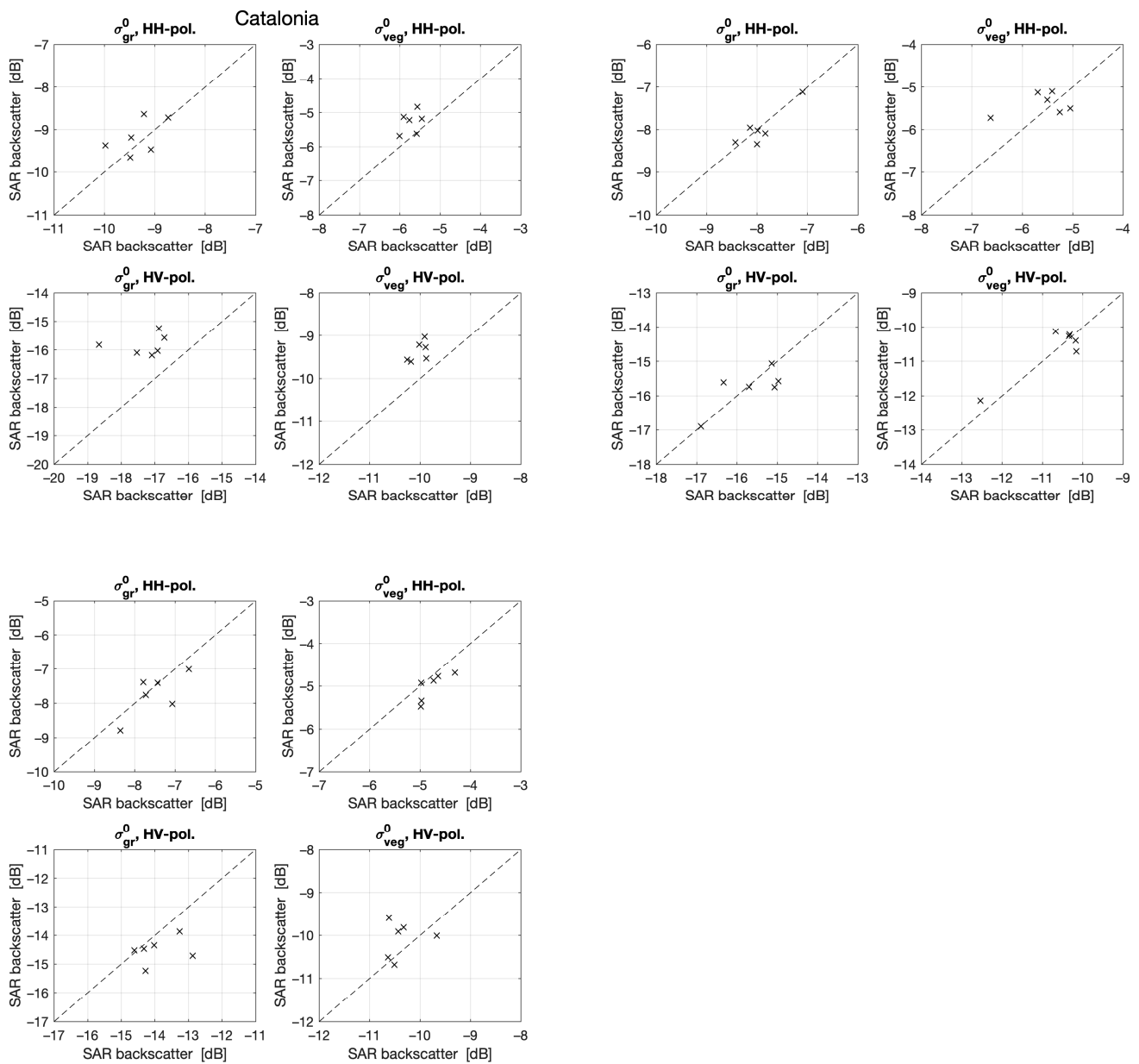
The same analysis was undertaken for the ALOS-2 PALSAR-2 dataset over the study sites of Catalonia, Finland N, and Finland S. Figure 11 shows measured and modelled backscatter as a function of GSV for HH and HV polarizations for Catalonia, Finland N, and Finland S. The training and calibration of the WCM were tested on all ALOS-2 mosaic data and the estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  are displayed in Figure 12.



**Figure 10.** Scatter plots illustrating the estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  from training (x axis) and calibration (y axis) for VV- and VH-polarized Sentinel-1 images over the sites of Catalonia and Finland N. The dashed line represents the identity line.



**Figure 11.** Measured and modelled ALOS-2 PALSAR-2 HH- and HV-pol. backscatter as a function of GSV grouped for the sites of Catalonia, Finland N, and Finland S.



**Figure 12.** Scatter plots illustrating the estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  from training (x axis) and calibration (y axis) for HH- and HV-polarized ALOS-2 PALSAR-2 mosaics acquired between 2015 and 2020 over the sites of Catalonia, Finland N, and Finland S. The dashed line represents the identity line.

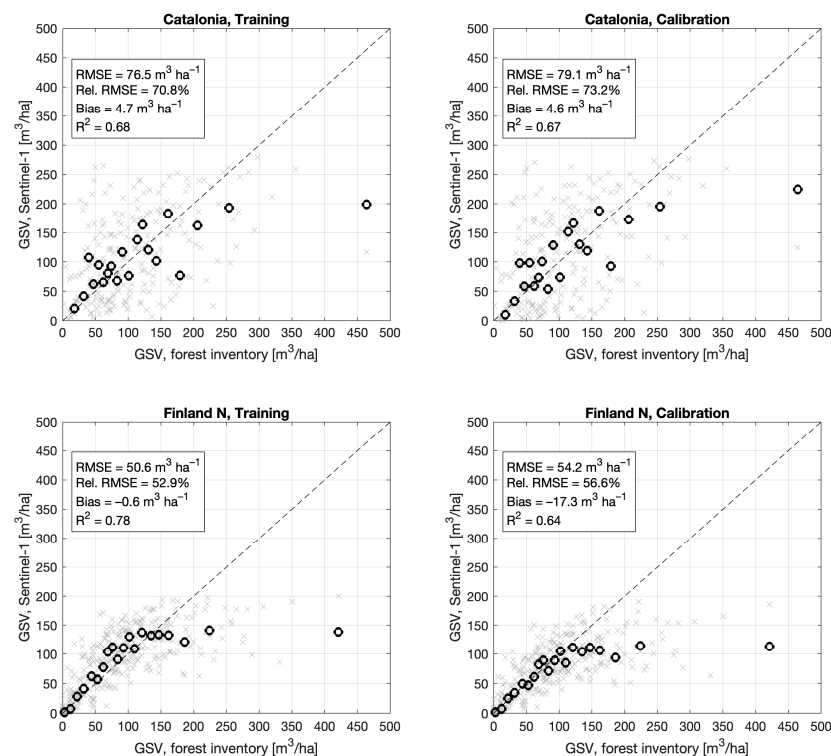
#### 5.4. Retrieval of GSV

GSV retrieval was performed using the trained and the calibrated WCM in Equation (6). The retrievals were constrained by the maximum GSV values obtained with Equation (11) in which we applied site-specific values of (i) ICESat-2 maximum canopy height (Table 4), (ii) the structural model coefficients  $a$  and  $b$  in Figure 2, and (iii) the standard deviation of GSV values corresponding to the ICESat-2  $h_{max}$  value. Table 7 compares our estimates of maximum GSV with the largest value from the dataset of plot-based measurements.

**Table 7.** Site-specific maximum GSV derived from the plot-based measurements and modelled with Equation (11) using the ICESat-2 maximum canopy height in Table 4. Values in brackets represent the percentile at which the plot-based GSV measurements are equal to the modelled value of maximum GSV.

| Site      | Maximum GSV [ $\text{m}^3/\text{ha}$ ] |             |
|-----------|----------------------------------------|-------------|
|           | Plot-Based                             | Modelled    |
| Finland N | 498                                    | 238 (96th)  |
| Finland S | 751                                    | 395 (94th)  |
| Catalonia | 480                                    | 504 (100th) |
| Romania   | 1677                                   | 1295 (99th) |

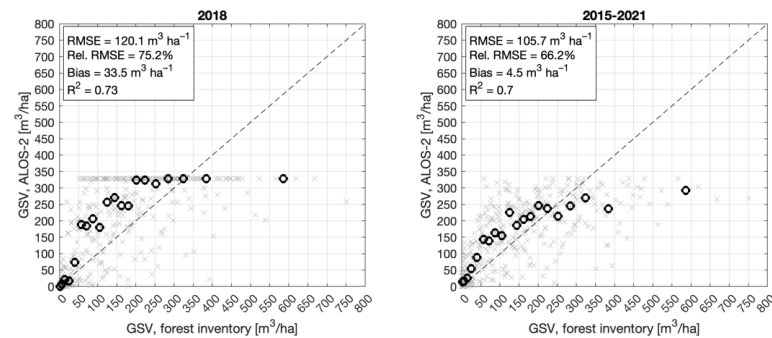
The performance of the retrieval using the Sentinel-1 dataset is illustrated with scatter plots comparing reference and estimated GSV and with retrieval statistics for the two sites of Catalonia and Finland N in Figure 13. For the remaining sites where the correlation between the Sentinel-1 backscatter and GSV was low, the scatter plots are included in the (Supplementary Material Figure S3).



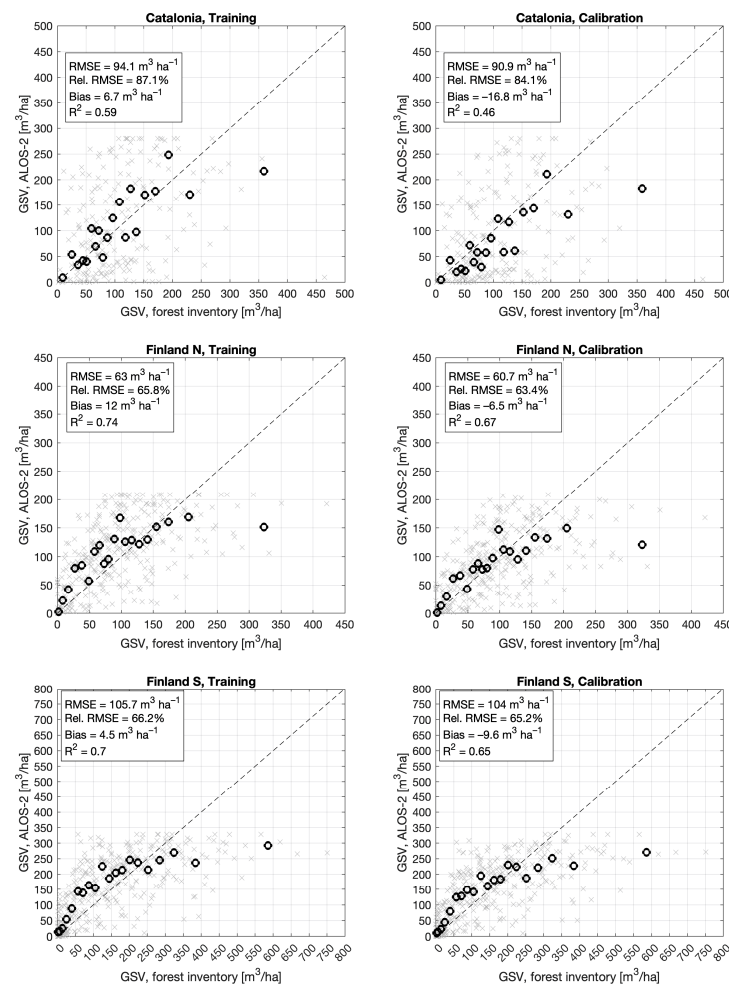
**Figure 13.** The comparison of GSV values estimated from the Sentinel-1 dataset and from field inventory for the sites of Catalonia and Finland N. Crosses refer to individual field plots. Circles represent the median value of the estimated GSV for  $10 \text{ m}^3/\text{ha}$  large bins of reference GSV. The dashed line represents the identity line.

The benefit of using multiple observations at the L-band is illustrated in Figure 14, which shows the retrieval results for the ALOS-2 PALSAR-2 mosaics of the same year of the field inventory data (2 SAR images) and for the 2015–2021 interval (12 SAR images). Results are shown for the Finland S site characterized by the highest correlation coefficients between SAR backscatter observations and GSV. The accuracy of the retrieval increased with the number of SAR observations; nonetheless, combining data acquired over several years might have been detrimental because the risk of not accounting for changes on the ground due to growth and losses would have biased the GSV estimates. In the remainder

of this section, we therefore assessed the L-band retrieval with ALOS-2 PALSAR-2 mosaics based on images acquired between the year prior to and after the field inventory. Results are displayed in Figure 15 for the sites of Catalonia, Finland N, and Finland S. The results for the Romanian site where correlation between SAR backscatter and GSV was low are illustrated in the (Supplementary Material Figure S4).

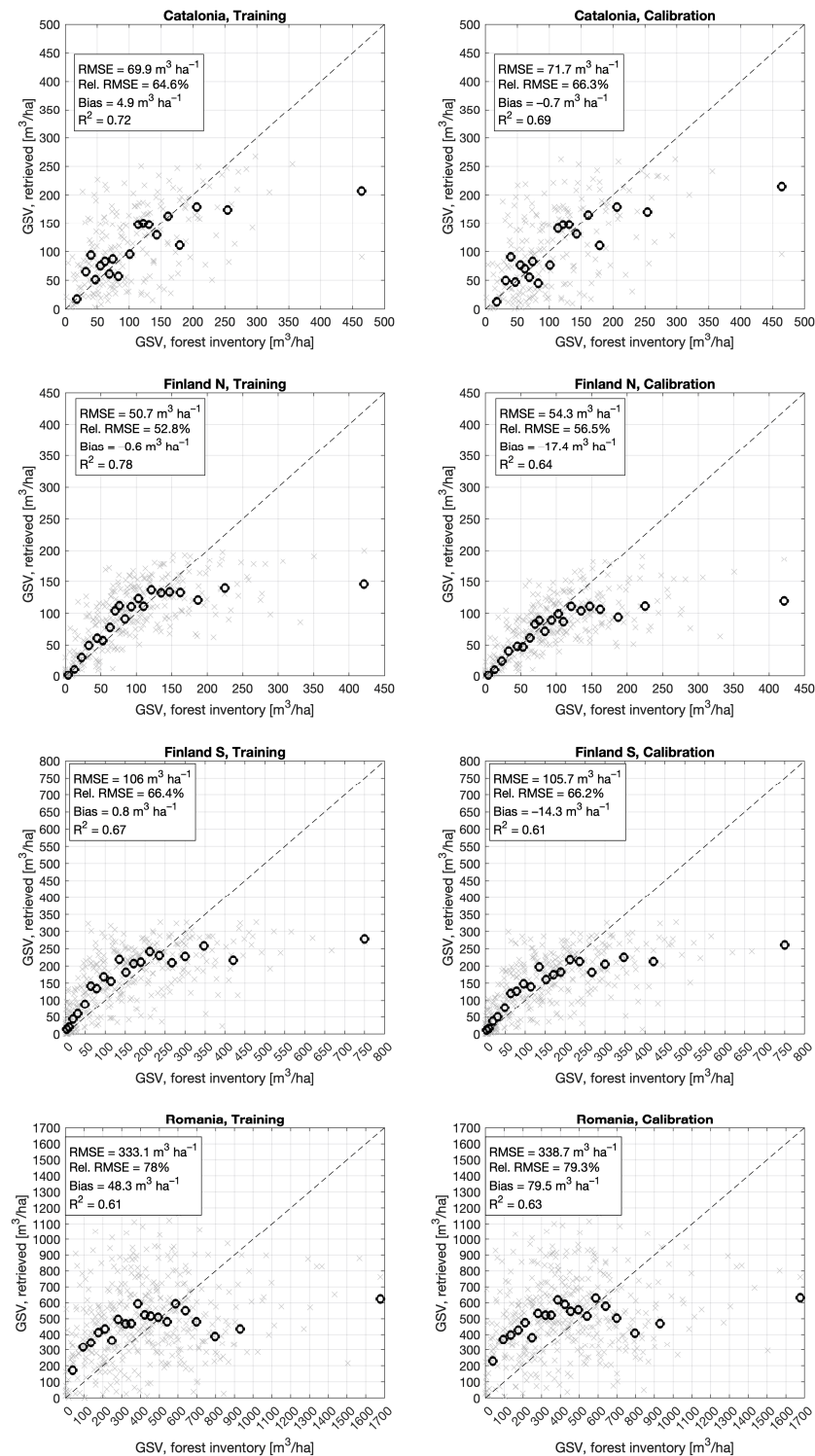


**Figure 14.** GSV estimates from the ALOS-2 PALSAR-2 mosaic of 2018 and the mosaics of 2015–2021 compared to the field inventory values for the site of Finland S. Plot arrangement and notations follow Figure 13.



**Figure 15.** The comparison of GSV values estimated from three years of ALOS-2 PALSAR-2 mosaics and from field inventory for the sites of Catalonia, Finland N, and Finland S. Plot arrangement and notations follow Figure 13.

The weights used for merging the Sentinel-1 and PALSAR-2 GSV estimates differed depending on the site. The weighting factor for the L-band GSV estimates ranged between nearly 0 (Finland N), 0.23 (Catalonia), 0.54 (Romania), and 0.9 (Finland S). The scatter plots illustrating reference and estimated GSV are illustrated for each site in Figure 16.



**Figure 16.** Scatter plots comparing the SAR-based and the field-measured GSV for all study sites. Plot arrangement and notations follow Figure 13.

## 6. Discussion

Estimating forest GSV using satellite data that correlate moderately or poorly with biomass suggests that any imperfection in the setup of the retrieval approach would cause substantial estimation errors. In the retrieval scheme proposed here, errors could be due to (i) the structural functions relating the forest variables in the WCM with GSV, (ii) the data used for training the structural models, (iii) the forest backscatter model, (iv) the procedure for estimating the parameters of this model, and (v) the procedure for combining GSV estimates from individual observations of the SAR backscatter. Assessment of any of these aspects requires a dataset of reference GSV measurements. In this study, such data consisted of plot-level measurements collected by field inventories at four sites. At the spatial scale of individual pixels, our datasets of SAR backscatter and GSV measurements were moderately or poorly correlated (Tables 5 and 6). Plot-level data and pixel-level backscatter are known to correlate weakly. This effect was reinforced by the limited penetration of the C- and L-band signal into the canopy (see, e.g., Figures 9 and 11) and, in part, by the pre-processing of the SAR data, which especially over sloped terrain accounted for the pixel area effect on the backscatter but not for scatterer-specific effects [74]. Under these circumstances, we acknowledge that the setup for assessing the performance of the method proposed here is sub-optimal. Nonetheless, our interest was to explain errors associated with a model calibration approach independent of GSV reference data. Hence, the plot-based datasets of GSV were assumed to be the most accurate data source that could be used for model calibration; i.e., the errors associated with forest inventories were not further addressed.

The training and calibration of the WCM relied on the same set of structural functions. Although the plot data could have been used to derive site-specific functions, our focus was on the WCM calibration so that the same structural functions had to be used to maintain consistency in the retrieval phase. While the relationship between canopy density and canopy height in Equation (4) was extensively documented, the experimental evidence for the power-law function relating canopy height and GSV was relatively limited. To overcome the lack of GSV measurements for large areas, we trained Equation (5) with sub-national average values that provided an almost gap-free coverage. The structural function in Equation (5) reproduced the relationship between canopy height and GSV in each of the four major European biomes. The two types of temperate forest biomes and the Mediterranean forest biome were characterized by a power-law relationship resembling a quadratic function (i.e., exponent close to 2). Boreal forests exhibited a more linear relationship with an exponent close to 1. The standard error (SE) of the regression indicated a low error across all biomes (between 9% and 20% of the average GSV) except for Mediterranean forest. There, SE of 45% relative to the average GSV was caused by the large spread of the observations, which was due to the variety of vegetation types in this environment.

The validity of the curves in Figure 2 can be questioned when considering that the structural function was trained on large-scale averages while our aim was to apply it at the pixel level of a SAR image. Our understanding was that the region of validity of the curves corresponded to the range of GSV values present in the training dataset (e.g., up to 400 m<sup>3</sup>/ha in temperate broadleaf and mixed forests). The structural function might instead produce inaccurate predictions of GSV for forest patches with GSV exceeding the largest average GSV in the dataset. As shown by the GSV statistics in Table 2, the largest of the GSV values in the dataset of sub-national statistics was considerably smaller than the maximum value of GSV at the plot level and in some cases comparable with the third quartile (Finland S) or with the second quartile (Romania). The impact of a presumably imperfect calibration of the structural function was further evidenced by the different values for the maximum GSV (Table 7). The estimate of the maximum GSV with the structural function for Catalonia was in line with the largest values reported in the plot-level GSV dataset. For both Finnish sites and the Romanian site, the maximum GSV predicted by the structural function instead corresponded only to a very high percentile in the reference plot dataset, being much smaller than the maximum measured GSV. For the two sites in the boreal zone, the discrepancy was attributed to the rather linear trend between height and

GSV (Figure 2), which was caused by the lack of data points with a GSV above 200 m<sup>3</sup>/ha so that the high end of the function, where typically a non-linear correspondence occurs, could not be accurately constrained. Indeed, when fitting Equation (5) to the plot data, we obtained an estimate for the exponent close to 2 (not shown here), leading to a much steeper increase in GSV with height. For the Romanian site, the large difference between maximum GSV values (>400 m<sup>3</sup>/ha) was apparent since only 12 of the total 1306 plots had GSV larger than our model's maximum estimate (Figure 2).

The SAR backscatter steadily increased with canopy density, regardless of the site, frequency, polarization, or incidence angle. The strongest sensitivity was observed at the L-band for HV-polarized data with an increase of 2.5 dB up to 4 dB (Figure 7). C-band VH-polarized data showed an increase of 2–3 dB (Figures 3 and 4). The co-polarized data were characterized by weak sensitivity to canopy density, being mostly between 1 dB and 2 dB (Figures 3, 4 and 7). The overall poor-to-moderate association between SAR backscatter and canopy density did not affect the performance of the calibration. Indeed, the WCM in Equation (2) fitted to the observations of canopy density and backscatter was able to reproduce their trend (Figures 3 and 4 at C-band and Figure 7 at L-band) and generated plausible estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  despite the large range of backscatter values for a given canopy density. The low ENL of the Sentinel-1 and ALOS-2 PALSAR-2 datasets (Table 3) partly explaining the large spread of the observations caused the strong correction to be applied to the initial estimate of  $\sigma_{veg}^0$  to obtain its final value with Equation (9).

The relationship between canopy density and variability of the SAR backscatter measurements after correcting for their uncertainty (Figures 5 and 6 at C-band and Figure 8 at L-band) was clearly linear, so that the correction factor to be applied to the initial estimate of  $\sigma_{veg}^0$  was easily determined. This correction factor was between 0.2 and 0.5 dB at the C-band (Figures 5 and 6) and around 0.5 dB at the L-band (Figure 8). For any SAR image, the resulting estimate of  $\sigma_{veg}^0$  was among the largest measured SAR backscatter values within the study site, this being consistent with our assumption that  $\sigma_{veg}^0$  is an ideal value of the backscatter for the largest possible stocks of wood volume in the area of interest.

The calibrated WCM was close to the trained WCM and both curves represented the trend in the observations (Figure 9 for C-band and Figure 11 for L-band). The mean of the model residuals confirmed the visual impression that the model fit obtained with the calibration tended to slightly overestimate the backscatter, regardless of frequency and polarization. The bias of the calibrated models was, however, substantially smaller than the range of measured backscatter values and less than the uncertainty characterizing the SAR backscatter data. In addition, the moderate correlation between backscatter and GSV did not allow us to conclude to which extent one curve is more representative of the true association between the two variables. Indeed, the trained WCM in Catalonia was characterized by somewhat unrealistic values for  $\sigma_{gr}^0$  both at C- (Figure 9) and L-bands (Figure 11), because they were lower than any measurement, and by a steep increase in the modelled backscatter in correspondence with a moderate increment of GSV.

The comparison of all estimates of  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  from the training and the calibration showed strong agreement at both C- (Figure 10) and L-bands (Figure 12) except for the Catalanian site where a systematic offset for  $\sigma_{gr}^0$  at the C-band for both polarizations and at the L-band at cross-polarization occurred. The calibration generated larger values than those obtained with the training of the WCM (Figures 11 and 12), the latter one being unrealistically low.

With Sentinel-1 C-band data, the retrieval based on the calibrated WCM was comparable to the retrieval based on the trained WCM for the two sites of Finland N and Catalonia characterized by the highest correlation between backscatter and GSV (Figure 13). The retrieval error of 56% and 75% relative to the mean GSV was only minimally affected by the choice of the model-fitting procedure. The discrepancy between the retrievals for the two other sites was instead substantial and none of the model-fitting approaches could reproduce the plot-based GSVs in the test set. The overall large retrieval error was explained by the considerable variance of model predictions (Figure 13), a consequence of the

weak sensitivity of the C-band backscatter to GSV (Figure 9). The accuracy was similar to previous studies comparing the estimates of forest biomass (GSV or AGB) at pixels with plot-level values. Nonetheless, the level of GSV was estimated correctly, with a tendency to underestimate the largest GSVs. This was a consequence of (i) the rather large estimate of  $\sigma_{veg}^0$  and (ii) the inversion rule that set a maximum retrievable value, which often was smaller than the site-level value derived from the plot data.

With ALOS-2 PALSAR-2 data, the retrieval based on the trained WCM produced estimates of GSV that were only slightly overestimated in the central range of GSVs and underestimated for the highest GSVs (Figure 15). As in the C-band case, these were underestimated because of the limit imposed to the range of retrievable GSVs and because of the high value of  $\sigma_{veg}^0$ . The performance of the retrieval from the calibrated WCM was comparable to the best-case scenario at the two Finnish sites (Figure 15). For Catalonia, slight underestimation occurred because of the slightly higher estimates of  $\sigma_{veg}^0$  (Figure 12). This could be due to imperfect terrain flattening of the mosaic leading to uncompensated backscatter in sloped terrain. Translating variance of backscatter into variance of GSV estimations explained the large retrieval error between 63% and 87%. The error was marginally affected by the procedure used to estimate the model parameters, except for Catalonia. At the Romanian site, the retrievals were characterized by poor accuracy (Figure S4) due to the low correlation between forest backscatter and GSV (Table 6).

The weights used to merge the C- and L-band GSV estimates differed depending on the study site. For Catalonia, the weight was geared towards the C-band estimates, which was reasonable given the poorer performance of the retrieval with the L-band (Figure 15). The merged GSV estimates (Figure 16) were slightly more accurate than in the C-band-only case (Figure 13) with the RMSE decreasing from 79.1 m<sup>3</sup>/ha to 71.7 m<sup>3</sup>/ha and the bias decreasing from 4.6 m<sup>3</sup>/ha to 0.7 m<sup>3</sup>/ha for the calibrated WCM. For Finland N, the weighting excluded the L-band estimates although their accuracy was comparable to the C-band retrieval (Figures 13 and 15). This was a shortcoming of how the normalization procedure in Equation (16) was implemented for this analysis. For Finland S, the L-band estimates were preferred to the C-band estimates, being consistent with the superior performance of the L-band retrieval and a consequence of the higher correlation between backscatter and GSV (Table 6). The error statistics of the merged estimates (Figure 16) were comparable to those obtained for the L-band estimates and were only minimally affected by the C-band retrieval. For Romania, the weights were similar, as a consequence of the poor correlation between GSV and backscatter at both C- and L-bands (Tables 5 and 6) so that none of the GSV estimates prevailed.

The accuracy of the merged estimates of GSV from the trained and from the calibrated WCM was comparable (Figure 16). The closest agreement was obtained in Catalonia with a relative RMSE close to 65%,  $R^2$  of approximately 0.7, and negligible bias for both model-fitting approaches. At Finland N, the retrieval based on the calibrated WCM was more accurate below 100 m<sup>3</sup>/ha, whereas the opposite occurred above 100 m<sup>3</sup>/ha with increasing underestimation regardless of the model-fitting approach. At Finland S, both sets of GSV estimates were characterized by both over- and underestimations in an equal manner. We explained this result as partial incapability of the WCM to reproduce the relationship between SAR backscatter and GSV. In Romania, the retrieval led to strong overestimation, which was expected given the lack of correlation between SAR backscatter and GSV. An investigation to explain the origin of the errors was beyond the scope of this study.

When putting these overall encouraging results in a broader context of biomass estimation with SAR backscatter data, our results show that the levels of GSV are well captured with the inversion of a Water Cloud Model and the combination of estimates from C- and L-bands. Except for the Romanian site, the estimates of GSV clustered along the identity line with systematic underprediction occurring for the highest GSVs only. Nonetheless, the estimates of GSV are characterized by substantial variance as shown by relative RMSE values between 53% and 79% of the average GSV and  $R^2$  values between

0.61 and 0.78 (Figure 16). The weak sensitivity of the C- and L-band backscatter to GSV explains the large dispersion of the estimates.

Unlike studies that explained the systematic underprediction of high biomass as a consequence of backscatter saturation [75], we instead associate biases to either an incorrect value of the maximum retrievable GSV (see also [43,44]) or an overestimate of the parameter  $\sigma_{veg}^0$  (see also [44]). As C- and L-band backscatter lose sensitivity for increasing AGB or GSV, measurements that are outside of the range of values predicted by a retrieval model might either not be inverted or associated with a biomass value within the range of the modelled backscatter. As a consequence, the high range of biomass is underestimated. By instead applying a constraint such as a maximum retrievable biomass, such backscatter measurements are here associated with a plausible biomass value (see, e.g., Figure 14). Because of the mostly weak sensitivity of the C- and L-band backscatter to biomass, the biomass estimates from a multi-temporal stack of observation are noisy and the combination of values suggested in this study acts as a filter so that (i) the impact of erroneous values and “hard coded” biomass on the final estimates is limited, and (ii) the statistical distribution of biomass is well reproduced.

## 7. Conclusions

This study compared the performance of forest GSV retrieval using two approaches to fit C- and L-band SAR backscatter data to forest GSV with a physically based WCM. Our key finding was that the performance of the WCM calibrated on auxiliary image data and independent from reference GSV data was comparable to a WCM trained on such reference data. To corroborate the results, the analysis was undertaken at four sites representing various forest conditions in Europe. Although the GSV estimates from the SAR data and from the plot inventory measurements were affected by large spread, the proposed procedure was able to correctly reproduce the spatial distribution of GSV except for very high GSV over sloped terrain. A calibrated WCM is therefore a viable approach when GSV reference data are poor or unavailable. This scenario applies to wide-area mapping, e.g., at the continental level, when trans-national reference measurements may be cumbersome to access due to data policy restrictions or difficult to interpret due to different measuring protocols and reporting.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs16214079/s1>, Figure S1. Pearson’s correlation coefficient between the Sentinel-1 backscatter measurements and the plot measurements of GSV for each acquisition date, study site and polarization (DOY: day of year). Figure S2. Pearson’s correlation coefficient between the ALOS-2 PALSAR-2 backscatter measurements and the plot measurements of GSV for each year of the mosaics, study site and polarization. Figure S3. Comparison of GSV values estimated from the Sentinel-1 dataset and from field inventory for the sites of Finland S (top row) and Romania (bottom row). The SAR-based estimates were obtained with the WCM trained with field measurements (left panels) and with the calibrated WCM (right panels). Crosses refer to individual field plots. Circles represent the median value of the estimated GSV for 10 m<sup>3</sup>/ha large bins of reference GSV. The dashed line represents the identity line. Figure S4. Same as in Figure S3 for PALSAR-2-based GSV estimates for the site of Romania. Table S1. Details on NFI-derived GSV datasets for European countries. References [76,77] are cited in the Supplementary Materials.

**Author Contributions:** Conceptualization, M.S. and O.C.; methodology, M.S. and O.C.; validation, M.S.; formal analysis, M.S., O.C., O.A. and J.M.; investigation, M.S.; resources, J.M.; data curation, J.M.; writing—original draft preparation, M.S.; writing—review and editing, O.C. and O.A.; visualization, M.S. and J.M.; supervision, M.S.; project administration, J.M.; funding acquisition, J.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was carried out under the EOEP5 programme and funded by the European Space Agency (ESA), contract 4000135015/21/I-NB—Forest Carbon Monitoring.

**Data Availability Statement:** Data generated in this study may be available on request.

**Acknowledgments:** The authors want to thank Tornator Oyj for providing the field plot data for the Romanian study site. The field measurements from the Finnish Forestry Center were publicly available (<https://www.metsakeskus.fi/fi/avoin-metsa-ja-luontotieto/aineistot-paikkatieto-ohjelmille/paikkatietoaineistot>, last accessed on 4 April 2024). The Spanish NFI plot measurements for Catalonia are openly available ([https://laboratoriforestal.creaf.cat/nfi\\_app/](https://laboratoriforestal.creaf.cat/nfi_app/), accessed on 18 September 2024). We are thankful to the two anonymous reviewers for their comments and suggestions.

**Conflicts of Interest:** Authors Oleg Antropov and Jukka Miettinen were employed by the company VTT Technical Research Centre of Finland. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. FAO. *Global Forest Resources Assessment 2020*; FAO: Rome, Italy, 2020.
2. Picard, N.; Saint-André, L.; Henry, M. *Manual for Building Tree Volume and Biomass Allometric Equations: From Filed Measurement to Prediction*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 2012; ISBN 978-92-5-107347-6.
3. Mitchard, E.T.A.; Feldpausch, T.R.; Brien, 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. *Glob. Ecol. Biogeogr.* **2014**, *23*, 935–946. [CrossRef] [PubMed]
4. Ulaby, F.T.; Moore, R.K.; Fung, A.K. *Microwave Remote Sensing, Active and Passive*; Artech House: Dedham, MA, USA, 1986; Volume 3.
5. Israelsson, H.; Ulander, L.M.H.; Askne, J.L.H.; Fransson, J.E.S.; Frolind, P.-O.; Gustavsson, A.; Hellsten, H. Retrieval of Forest Stem Volume Using VHF SAR. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 36–40. [CrossRef]
6. Le Toan, T.; Quegan, S.; Davidson, M.W.J.; Balzter, H.; Paillou, P.; Papathanassiou, K.; Plummer, S.; Rocca, F.; Saatchi, S.; Shugart, H.; et al. The BIOMASS Mission: Mapping Global Forest Biomass to Better Understand the Terrestrial Carbon Cycle. *Remote Sens. Environ.* **2011**, *115*, 2850–2860. [CrossRef]
7. Ranson, K.J.; Sun, G.; Lang, R.H.; Chauhan, N.S.; Cacciola, R.J.; Kilic, O. Mapping of Boreal Forest Biomass from Spaceborne Synthetic Aperture Radar. *J. Geophys. Res. Atmos.* **1997**, *102*, 29599–29610.
8. Cartus, O.; Santoro, M.; Wegmüller, U.; Rommen, B. Benchmarking the Retrieval of Biomass in Boreal Forests Using P-Band SAR Backscatter with Multi-Temporal C- and L-Band Observations. *Remote Sens.* **2019**, *11*, 1695. [CrossRef]
9. Santoro, M.; Cartus, O.; Fransson, J.E.S.; Wegmüller, U. Complementarity of X-, C-, and L-Band SAR Backscatter Observations to Retrieve Forest Stem Volume in Boreal Forest. *Remote Sens.* **2019**, *11*, 1563. [CrossRef]
10. Englhart, S.; Keuck, V.; Siegert, F. Aboveground Biomass Retrieval in Tropical Forests—The Potential of Combined X- and L-Band SAR Data Use. *Remote Sens. Environ.* **2011**, *115*, 1260–1271. [CrossRef]
11. Joshi, N.P.; Mitchard, E.T.A.; Schumacher, J.; Johannsen, V.K.; Saatchi, S.; Fensholt, R. L-Band SAR Backscatter Related to Forest Cover, Height and Aboveground Biomass at Multiple Spatial Scales across Denmark. *Remote Sens.* **2015**, *7*, 4442–4472. [CrossRef]
12. Atwood, D.K.; Andersen, H.-E.; Matthijs, B.; Holec, F. Impact of Topographic Correction on Estimation of Aboveground Boreal Biomass Using Multi-Temporal, L-Band Backscatter. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3262–3273.
13. Lone, J.M.; Sivasankar, T.; Sarma, K.K.; Qadir, A.; Raju, P.L.N. Influence of Slope Aspect on above Ground Biomass Estimation Using ALOS-2 Data. *Int. J. Sci. Res. IJSR* **2017**, *6*, 1422–1428. [CrossRef]
14. Kumar, A.; Kishore, B.S.P.C.; Saikia, P.; Deka, J.; Bharali, S.; Singha, L.B.; Tripathi, O.P.; Khan, M.L. Tree Diversity Assessment and above Ground Forests Biomass Estimation Using SAR Remote Sensing: A Case Study of Higher Altitude Vegetation of North-East Himalayas, India. *Phys. Chem. Earth Parts ABC* **2019**, *111*, 53–64. [CrossRef]
15. Huang, W.; Sun, G.; Ni, W.; Zhang, Z.; Dubayah, R. Sensitivity of Multi-Source SAR Backscatter to Changes in Forest Aboveground Biomass. *Remote Sens.* **2015**, *7*, 9587–9609. [CrossRef]
16. Antropov, O.; Rauste, Y.; Häme, T.; Praks, J. Polarimetric ALOS PALSAR Time Series in Mapping Biomass of Boreal Forests. *Remote Sens.* **2017**, *9*, 999. [CrossRef]
17. Chowdhury, T.A.; Thiel, C.; Schumilius, C.; Stelmaszczuk-Górska, M. Polarimetric Parameters for Growing Stock Volume Estimation Using ALOS PALSAR L-Band Data over Siberian Forests. *Remote Sens.* **2013**, *5*, 5725–5756. [CrossRef]
18. Tanase, M.A.; Panciera, R.; Lowell, K.; Tian, S.; Hacker, J.M.; Walker, J.P. Airborne Multi-Temporal L-Band Polarimetric SAR Data for Biomass Estimation in Semi-Arid Forests. *Remote Sens. Environ.* **2014**, *145*, 93–104.
19. Pulliainen, J.T.; Mikkilä, P.J.; Hallikainen, M.T.; Ikonen, J.-P. Seasonal Dynamics of C-Band Backscatter of Boreal Forests with Applications to Biomass and Soil Moisture Estimation. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 758–770.
20. Englhart, S.; Keuck, V.; Siegert, F. Modeling Aboveground Biomass in Tropical Forests Using Multi-Frequency SAR Data—A Comparison of Methods. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 298–306. [CrossRef]
21. Berninger, A.; Lohberger, S.; Stängel, M.; Siegert, F. SAR-Based Estimation of Above-Ground Biomass and Its Changes in Tropical Forests of Kalimantan Using L- and C-Band. *Remote Sens.* **2018**, *10*, 831. [CrossRef]
22. Omar, H.; Misman, M.; Kassim, A. Synergetic of PALSAR-2 and Sentinel-1A SAR Polarimetry for Retrieving Aboveground Biomass in Dipterocarp Forest of Malaysia. *Appl. Sci.* **2017**, *7*, 675. [CrossRef]

23. Santoro, M.; Beer, C.; Cartus, O.; Schmullius, C.; Shvidenko, A.; McCallum, I.; Wegmüller, U.; Wiesmann, A. Retrieval of Growing Stock Volume in Boreal Forest Using Hyper-Temporal Series of Envisat ASAR ScanSAR Backscatter Measurements. *Remote Sens. Environ.* **2011**, *115*, 490–507. [\[CrossRef\]](#)
24. Cartus, O.; Santoro, M.; Kellndorfer, J. Mapping Forest Aboveground Biomass in the Northeastern United States with ALOS PALSAR Dual-Polarization L-Band. *Remote Sens. Environ.* **2012**, *124*, 466–478. [\[CrossRef\]](#)
25. Khati, U.; Laval, M.; Singh, G. The Role of Time-Series L-Band SAR and GEDI in Mapping Sub-Tropical Above-Ground Biomass. *Front. Earth Sci.* **2021**, *9*, 752254. [\[CrossRef\]](#)
26. Santi, E.; Paloscia, S.; Pettinato, S.; Fontanelli, G.; Mura, M.; Zolli, C.; Maselli, F.; Chiesi, M.; Bottai, L.; Chirici, G. The Potential of Multifrequency SAR Images for Estimating Forest Biomass in Mediterranean Areas. *Remote Sens. Environ.* **2017**, *200*, 63–73. [\[CrossRef\]](#)
27. Shao, Z.; Zhang, L.; Wang, L. Stacked Sparse Autoencoder Modeling Using the Synergy of Airborne LiDAR and Satellite Optical and SAR Data to Map Forest Above-Ground Biomass. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5569–5582. [\[CrossRef\]](#)
28. Stelmaszczuk-Górska, M.A.; Urbazhev, M.; Schmullius, C.; Thiel, C. Estimation of Above-Ground Biomass over Boreal Forests in Siberia Using Updated In Situ, ALOS-2 PALSAR-2, and RADARSAT-2 Data. *Remote Sens.* **2018**, *10*, 1550. [\[CrossRef\]](#)
29. Tanase, M.A.; Panciera, R.; Lowell, K.; Aponte, C.; Hacker, J.M.; Walker, J.P. Forest Biomass Estimation at High Spatial Resolution: Radar Versus Lidar Sensors. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 711–715. [\[CrossRef\]](#)
30. Han, H.; Wan, R.; Li, B. Estimating Forest Aboveground Biomass Using Gaofen-1 Images, Sentinel-1 Images, and Machine Learning Algorithms: A Case Study of the Dabie Mountain Region, China. *Remote Sens.* **2021**, *14*, 176. [\[CrossRef\]](#)
31. Hu, Y.; Nie, Y.; Liu, Z.; Wu, G.; Fan, W. Improving the Potential of Coniferous Forest Aboveground Biomass Estimation by Integrating C- and L-Band SAR Data with Feature Selection and Non-Parametric Model. *Remote Sens.* **2023**, *15*, 4194. [\[CrossRef\]](#)
32. Ge, S.; Tomppo, E.; Rauste, Y.; McRoberts, R.E.; Praks, J.; Gu, H.; Su, W.; Antropov, O. Sentinel-1 Time Series for Predicting Growing Stock Volume of Boreal Forest: Multitemporal Analysis and Feature Selection. *Remote Sens.* **2023**, *15*, 3489. [\[CrossRef\]](#)
33. Long, J.; Zheng, H.; Ye, Z.; Zhang, T.; Li, X. The Impacts of Phenological Stages within the Annual Cycle on Mapping Forest Stock Volume Using Multi-Band Dual-Polarization SAR Images in Boreal Forests. *Forests* **2024**, *15*, 1660. [\[CrossRef\]](#)
34. Salazar Villegas, M.H.; Qasim, M.; Csaplovics, E.; González-Martínez, R.; Rodríguez-Buritica, S.; Ramos Abril, L.N.; Salazar Villegas, B. Examining the Potential of Sentinel Imagery and Ensemble Algorithms for Estimating Aboveground Biomass in a Tropical Dry Forest. *Remote Sens.* **2023**, *15*, 5086. [\[CrossRef\]](#)
35. Shendryk, Y. Fusing GEDI with Earth Observation Data for Large Area Aboveground Biomass Mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *115*, 103108. [\[CrossRef\]](#)
36. Stelmaszczuk-Górska, M.; Rodríguez-Veiga, P.; Ackermann, N.; Thiel, C.; Balzer, H.; Schmullius, C. Non-Parametric Retrieval of Aboveground Biomass in Siberian Boreal Forests with ALOS PALSAR Interferometric Coherence and Backscatter Intensity. *J. Imaging* **2015**, *2*, 1. [\[CrossRef\]](#)
37. Vafaei, S.; Soosani, J.; Adeli, K.; Fadaei, H.; Naghavi, H.; Pham, T.; 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 Sens.* **2018**, *10*, 172. [\[CrossRef\]](#)
38. Wang, C.; Zhang, W.; Ji, Y.; Marino, A.; Li, C.; Wang, L.; Zhao, H.; Wang, M. Estimation of Aboveground Biomass for Different Forest Types Using Data from Sentinel-1, Sentinel-2, ALOS PALSAR-2, and GEDI. *Forests* **2024**, *15*, 215. [\[CrossRef\]](#)
39. Ji, Y.; Xu, K.; Zeng, P.; Zhang, W. GA-SVR Algorithm for Improving Forest Above Ground Biomass Estimation Using SAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6585–6595. [\[CrossRef\]](#)
40. Mitchard, E.T.A.; Saatchi, S.S.; Baccini, A.; Asner, G.P.; Goetz, S.J.; Harris, N.L.; Brown, S. Uncertainty in the Spatial Distribution of Tropical Forest Biomass: A Comparison of Pan-Tropical Maps. *Carbon Balance Manag.* **2013**, *8*, 10. [\[CrossRef\]](#)
41. Avitabile, V.; Herold, M.; Heuvelink, G.B.M.; Lewis, S.L.; Phillips, O.L.; Asner, G.P.; Armston, J.; Ashton, P.S.; Banin, L.; Bayol, N.; et al. An Integrated Pan-Tropical Biomass Map Using Multiple Reference Datasets. *Glob. Change Biol.* **2016**, *22*, 1406–1420. [\[CrossRef\]](#)
42. Santoro, M.; Beaudoin, A.; Beer, C.; Cartus, O.; Fransson, J.E.S.; Hall, R.J.; Pathe, C.; Schepaschenko, D.; Schmullius, C.; Shvidenko, A.; et al. Forest Growing Stock Volume of the Northern Hemisphere: Spatially Explicit Estimates for 2010 Derived from Envisat ASAR Data. *Remote Sens. Environ.* **2015**, *168*, 316–334. [\[CrossRef\]](#)
43. Santoro, M.; Cartus, O.; Carvalhais, N.; Rozendaal, D.; Avitabile, V.; Araza, A.; de Bruin, S.; Herold, M.; Quegan, S.; Rodríguez Veiga, P.; et al. The Global Forest Above-Ground Biomass Pool for 2010 Estimated from High-Resolution Satellite Observations. *Earth Syst. Sci. Data* **2021**, *13*, 3927–3950. [\[CrossRef\]](#)
44. Santoro, M.; Cartus, O.; Quegan, S.; Kay, H.; Lucas, R.M.; Araza, A.; Herold, M.; Labrière, N.; Chave, J.; Rosenqvist, Å.; et al. Design and Performance of the Climate Change Initiative Biomass Global Retrieval Algorithm. *Sci. Remote Sens.* **2024**, *10*, 100169. [\[CrossRef\]](#)
45. Askne, J.I.H.; Santoro, M. On the Estimation of Boreal Forest Biomass From TanDEM-X Data Without Training Samples. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 771–775. [\[CrossRef\]](#)
46. Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. National-Scale Biomass Estimators for United States Tree Species. *For. Sci.* **2003**, *49*, 12–35.

47. Häme, T.; Salli, A.; Lahti, K. Estimation of Carbon Storage in Boreal Forests Using Remote Sensing Data. In *The Finnish Research Program on Climate Change, Progress Report*; Kanninen, M., Anttila, P., Eds.; Pilot Study; Academy of Finland: Helsinki, Finland, 1992; Volume 3, pp. 250–255.
48. Shvidenko, A.; Schepaschenko, D.; Nilsson, S.; Bouloui, Y. Semi-Empirical Models for Assessing Biological Productivity of Northern Eurasian Forests. *Ecol. Model.* **2007**, *204*, 163–179. [\[CrossRef\]](#)
49. Zianis, D.; Mencuccini, M. On Simplifying Allometric Analyses of Forest Biomass. *For. Ecol. Manag.* **2004**, *187*, 311–332. [\[CrossRef\]](#)
50. Somogyi, Z.; Teobaldelli, M.; Federici, S.; Matteucci, G.; Pagliari, V.; Grassi, G.; Seufert, G. Allometric Biomass and Carbon Factors Database. *iForest* **2008**, *1*, 107–113.
51. Houghton, R.A. Aboveground Forest Biomass and the Global Carbon Balance. *Glob. Change Biol.* **2005**, *11*, 945–958. [\[CrossRef\]](#)
52. Avitabile, V.; Pilli, R.; Migliavacca, M.; Duveiller, G.; Camia, A.; Blujdea, V.; Adolt, R.; Alberdi, I.; Barreiro, S.; Bender, S.; et al. Harmonised Statistics and Maps of Forest Biomass and Increment in Europe. *Sci. Data* **2024**, *11*, 274. [\[CrossRef\]](#)
53. Fahrland, E. *Copernicus DEM Product Handbook*; AIRBUS: Blagnac, France, 2022; Version 4.0.
54. Wegmüller, U. Automated Terrain Corrected SAR Geocoding. In Proceedings of the IGARSS'99, Hamburg, Germany, 28 June–2 July 1999; IEEE Publications: Piscataway, NJ, USA, 1999; pp. 1712–1714.
55. Frey, O.; Santoro, M.; Werner, C.L.; Wegmüller, U. DEM-Based SAR Pixel-Area Estimation for Enhanced Geocoding Refinement and Radiometric Normalization. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 48–52. [\[CrossRef\]](#)
56. Quegan, S.; Le Toan, T.; Yu, J.J.; Ribbes, F.; Floury, N. Multitemporal ERS SAR Analysis Applied to Forest Mapping. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 741–753.
57. Shimada, M.; Itoh, T.; Motooka, T.; Watanabe, M.; Shiraishi, T.; Thapa, R.; Lucas, R. New Global Forest/Non-Forest Maps from ALOS PALSAR Data (2007–2010). *Remote Sens. Environ.* **2014**, *155*, 13–31. [\[CrossRef\]](#)
58. Shimada, M.; Ohtaki, T. Generating Large-Scale High-Quality SAR Mosaic Datasets: Application to PALSAR Data for Global Monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 637–656.
59. Oliver, C.; Quegan, S. *Understanding Synthetic Aperture Radar Images*; Artech House: Boston, MA, USA, 1998.
60. Proisy, C.; Mougin, E.; Dufrêne, E.; Le Dantec, V. Monitoring Seasonal Changes of a Mixed Temperate Forest Using ERS SAR Observations. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 540–552. [\[CrossRef\]](#)
61. Markus, T.; Neumann, T.; Martino, A.; Abdalati, W.; Brunt, K.; Csatho, B.; Farrell, S.; Fricker, H.; Gardner, A.; Harding, D.; et al. The Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2): Science Requirements, Concept, and Implementation. *Remote Sens. Environ.* **2017**, *190*, 260–273. [\[CrossRef\]](#)
62. Neuenschwander, A.; Pitts, K. The ATL08 Land and Vegetation Product for the ICESat-2 Mission. *Remote Sens. Environ.* **2019**, *221*, 247–259. [\[CrossRef\]](#)
63. Askne, J.; Dammert, P.B.G.; Ulander, L.M.H.; Smith, G. C-Band Repeat-Pass Interferometric SAR Observations of the Forest. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 25–35. [\[CrossRef\]](#)
64. Santoro, M.; Cartus, O.; Fransson, J.E.S. Integration of Allometric Equations in the Water Cloud Model towards an Improved Retrieval of Forest Stem Volume with L-Band SAR Data in Sweden. *Remote Sens. Environ.* **2021**, *253*, 112235. [\[CrossRef\]](#)
65. Kay, H.; Santoro, M.; Cartus, O.; Bunting, P.; Lucas, R. Exploring the Relationship between Forest Canopy Height and Canopy Density from Spaceborne LiDAR Observations. *Remote Sens.* **2021**, *13*, 4961. [\[CrossRef\]](#)
66. Santoro, M.; Cartus, O.; Wegmüller, U.; Besnard, S.; Carvalhais, N.; Araza, A.; Herold, M.; Liang, J.; Cavlovic, J.; Engdahl, M.E. Global Estimation of Above-Ground Biomass from Spaceborne C-Band Scatterometer Observations Aided by LiDAR Metrics of Vegetation Structure. *Remote Sens. Environ.* **2022**, *279*, 113114. [\[CrossRef\]](#)
67. Dinerstein, E.; Olson, D.; Joshi, A.; Vynne, C.; Burgess, N.D.; Wikramanayake, E.; Hahn, N.; Palminteri, S.; Hedao, P.; Noss, R.; et al. An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience* **2017**, *67*, 534–545. [\[CrossRef\]](#)
68. Ballester-Berman, J.D. Reviewing the Role of the Extinction Coefficient in Radar Remote Sensing. *arXiv* **2020**. [\[CrossRef\]](#)
69. DiMiceli, C.M.; Carroll, M.L.; Sohlberg, R.A.; Huang, C.; Hansen, M.C.; Townshend, J.R.G. *Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000–2010, Collection 5 Percent Tree Cover 2011*; University of Maryland: College Park, MD, USA, 2011.
70. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21-St Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [\[CrossRef\]](#)
71. Santoro, M.; Fransson, J.E.S.; Eriksson, L.E.B.; Magnusson, M.; Ulander, L.M.H.; Olsson, H. Signatures of ALOS PALSAR L-Band Backscatter in Swedish Forest. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 4001–4019. [\[CrossRef\]](#)
72. Praks, J.; Antropov, O.; Hallikainen, M. LiDAR-Aided SAR Interferometry Studies in Boreal Forest: Scattering Phase Center and Extinction Coefficient at X- and L-Band. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3831–3843. [\[CrossRef\]](#)
73. Askne, J.; Santoro, M.; Smith, G.; Fransson, J.E.S. Multitemporal Repeat-Pass SAR Interferometry of Boreal Forests. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1540–1550. [\[CrossRef\]](#)
74. Hoekman, D.H.; Reiche, J. Multi-Model Radiometric Slope Correction of SAR Image of Complex Terrain Using a Two-Stage Semi-Empirical Approach. *Remote Sens. Environ.* **2015**, *156*, 1–10. [\[CrossRef\]](#)
75. Rodríguez-Veiga, P.; Quegan, S.; Carreiras, J.; Persson, H.J.; Fransson, J.E.S.; Hoscilo, A.; Ziolkowski, D.; Stereńczak, K.; Lohberger, S.; Stängel, M.; et al. Forest Biomass Retrieval Approaches from Earth Observation in Different Biomes. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *77*, 53–68. [\[CrossRef\]](#)

76. Čavlović, J. *Prva Nacionalna Inventura Šuma u Republici Hrvatskoj. Šumarski Fakultet SVEUČILIŠTA u Zagrebu i MRRŠVG*; First National Forest Inventory in Republic of Croatia (in Croatian); Ministry of Regional Development and Forestry & Faculty of Forestry: Zagreb, Croatia, 2010; p. 300.
77. Čavlović, J.; Božić, M.; Teslak, K.; Vedriš, M. Chapter 15—Croatia. In *National Forest Inventories—Assessment of Wood Availability and Use*; Vidal, C., Alberdi, I., Hernandez, L., Redmond, J., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 281–305, ISBN 978-3-319-44015-6.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.