
This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Ge, Shaojia; Tomppo, Erkki; Rauste, Yrjö; McRoberts, Ronald E.; Praks, Jaan; Gu, Hong; Su, Weimin; Antropov, Oleg

Sentinel-1 Time Series for Predicting Growing Stock Volume of Boreal Forest: Multitemporal Analysis and Feature Selection

Published in:
Remote Sensing

DOI:
[10.3390/rs15143489](https://doi.org/10.3390/rs15143489)

Published: 01/07/2023

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Ge, S., Tomppo, E., Rauste, Y., McRoberts, R. E., Praks, J., Gu, H., Su, W., & Antropov, O. (2023). Sentinel-1 Time Series for Predicting Growing Stock Volume of Boreal Forest: Multitemporal Analysis and Feature Selection. *Remote Sensing*, 15(14), Article 3489. <https://doi.org/10.3390/rs15143489>



Article

Sentinel-1 Time Series for Predicting Growing Stock Volume of Boreal Forest: Multitemporal Analysis and Feature Selection

Shaojia Ge ¹, Erkki Tomppo ², Yrjö Rauste ³, Ronald E. McRoberts ⁴, Jaan Praks ⁵, Hong Gu ¹, Weimin Su ¹ and Oleg Antropov ^{3,*}

¹ School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; geshaojia@njjust.edu.cn (S.G.); guhong666@njjust.edu.cn (H.G.); suweimin@njjust.edu.cn (W.S.)

² Department of Forest Sciences, University of Helsinki, 00014 Helsinki, Finland; erkki.tomppo@helsinki.fi

³ VTT Technical Research Centre of Finland, 00076 Espoo, Finland; yrjo.rauste@gmail.com

⁴ Department of Forest Resources, University of Minnesota, Saint Paul, MN 55108, USA; mcrob001@umn.edu

⁵ Department of Electronics and Nanoengineering, Aalto University, 02150 Espoo, Finland; jaan.praks@aalto.fi

* Correspondence: oleg.antropov@vtt.fi

Abstract: Copernicus Sentinel-1 images are widely used for forest mapping and predicting forest growing stock volume (GSV) due to their accessibility. However, certain important aspects related to the use of Sentinel-1 time series have not been thoroughly explored in the literature. These include the impact of image time series length on prediction accuracy, the optimal feature selection approaches, and the best prediction methods. In this study, we conduct an in-depth exploration of the potential of long time series of Sentinel-1 SAR data to predict forest GSV and evaluate the temporal dynamics of the predictions using extensive reference data. Our boreal coniferous forests study site is located near the Hyytiälä forest station in central Finland and covers an area of 2500 km² with nearly 17,000 stands. We considered several prediction approaches and fine-tuned them to predict GSV in various evaluation scenarios. Our analyses used 96 Sentinel-1 images acquired over three years. Different approaches for aggregating SAR images and choosing feature (predictor) variables were evaluated. Our results demonstrate a considerable decrease in the root mean squared errors (RMSEs) of GSV predictions as the number of images increases. While prediction accuracy using individual Sentinel-1 images varied from 85 to 91 m³/ha RMSE, prediction accuracy with combined images decreased to 75.6 m³/ha. Feature extraction and dimension reduction techniques facilitated the achievement of near-optimal prediction accuracy using only 8–10 images. Examined methods included radiometric contrast, mutual information, improved k-Nearest Neighbors, random forests selection, Lasso, and Wrapper approaches. Lasso was the most optimal, with RMSE reaching 77.1 m³/ha. Finally, we found that using assemblages of eight consecutive images resulted in the greatest accuracy in predicting GSV when initial acquisitions started between September and January.

Keywords: synthetic aperture radar; growing stock volume; boreal forests; Sentinel-1; support vector regression; random forests regression



Citation: 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. <https://doi.org/10.3390/rs15143489>

Academic Editors: Vahid Akbari, Nizar Bouhlel, Alireza Tabatabaeenejad and Esra Erten

Received: 21 May 2023

Revised: 1 July 2023

Accepted: 7 July 2023

Published: 11 July 2023



Copyright: © 2023 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

Space-borne synthetic aperture radar (SAR) is a highly versatile information source for Earth observation that can provide terrain-related data with fine resolution regardless of weather conditions. This makes it particularly useful in areas with persistent cloud cover or near polar regions. SAR images have been extensively used in forest remote sensing, particularly for wide-area mapping [1–3], where models of relationships between forest structural parameters and measured backscatter signatures are essential [4]. SAR signals have been demonstrated to be highly sensitive to various forest structure variables, including above-ground tree biomass (AGB) and growing stock volume (GSV) (m³/ha) [5]. Depending on the wavelengths of the sensor, the relationship between SAR backscattering

data and GSV can quickly become saturated. In addition, environmental and ionospheric conditions can affect this relationship. When signal saturation is reached, the radiometric data become less useful for predicting volume or biomass, especially when only one image is used. This means that data from commonly used space-borne SAR systems operating in the C-band and X-band are often considered suboptimal for wide-area forest inventory [5].

Cross-polarised (cross-pol) backscatter demonstrates greater sensitivity to forest biomass than co-polarised (co-pol) backscatter, although multiple polarizations are recommended for use in forest biomass mapping algorithms. C-band SAR is mostly useful for forests with very small biomass levels (30–50 t/ha) when a single SAR intensity image is used. The shorter wavelengths have larger extinction, and limited penetration within the forest layer compared to the L- and P-bands, although multi-temporal and texture analysis of fine resolution C-band data may provide useful inputs [6–10]. On the other hand, C-band radar data acquired by multiple space-borne sensors, historically including ERS-1, ERS-2, and Envisat ASAR, presently RADARSAT-2 and ESA Sentinel-1, and only recently RADARSAT Constellation, represent attractive sources of information for timely assessment of forest cover due to their global coverage (in the cases of ASAR and Sentinel-1) and frequent revisits. This is particularly important for Sentinel-1 satellites [11] with data offered free of charge to users.

To date, methodological studies using long time series of C-band SAR intensity data for GSV mapping over boreal forests have been relatively rare [10,12–14] compared to studies on multitemporal approaches for GSV estimation using L-band SAR imagery in boreal forests, particularly with JERS [12,15,16] and ALOS PALSAR [17–19]. The common understanding is that C-band data are suboptimal compared to L- and P-band data [5]; however, due to the availability of C-band data (Envisat ASAR in Global Monitoring mode), wide-area maps of GSV have been produced in boreal forests [13]. Parametric semi-empirical and physics-based models have been used with historical C-band SAR data, such as ERS-1 and Envisat ASAR, to predict forest stem volume and GSV [13,14,20]. Detailed historical studies of the temporal dynamics of C-band SAR backscatter in boreal forests are scarce, whereas in notable examples [14,20] the SAR data were represented by ERS-1 single-pol VV imagery at a 23° nominal incidence angle. In [20], the auxiliary parameters of a semi-empirical model were estimated separately from each processed SAR image, then a vector of backscatter measurements for each areal unit was iteratively inverted to predict the forest stem volume. Correlation analysis with reference plot-level stem volume was used to select suitable images. In [14], the parameters of a semi-empirical model and the stem volume were estimated simultaneously for each SAR image. Separate stem volume estimates were later combined using a linear regression model calibrated using the training data set. In addition, use of multiple linear regression (MLR) for optimally combining predictions from individual SAR images has been reported in [12]. A popular approach in forest biomass mapping is to initially predict the GSV using the Water Cloud Model (WCM) [21] for individual SAR images and then combining these predictions via optimal weighting of the predictions [13]. Weights can be estimated using image-wise accuracy statistics in the presence of reference data or with auxiliary datasets [13,22].

Recently, nonparametric approaches such as random forests (RF) and support vector regression (SVR) have been tested to predict GSV using features calculated from multiple C-band SAR images, often in combination with L-band and optical images [23]. Using features calculated from several SAR images as independent variables in GSV prediction has rarely been studied [10,24], and has typically included both SAR intensity and textural image features. Use of feature selection approaches has often been limited to RF-based rankings and simple correlation to forest biomass [24–28]. In [29,30], Wrapper methods such as stepwise regression analysis, recursive feature elimination, and Boruta were applied to select feature variables from multi-source satellite data in forest height mapping. In [31], the performance of importance-based selection methods, including RF and XGBoost, was analyzed in the context of forest biomass estimation. A binary genetic algorithm was tested

for selecting the features for forest GSV prediction to reduce the computational complexity of a SVR kernel [32].

The availability of freely available very long time series data from Sentinel-1, the presence of two polarizations (co- and cross-polarizations), and access to better reference data motivates more in-depth exploration of seasonal dynamics and their influence on GSV predictions compared to earlier studies with ERS-1 and ASAR. Studying various ways of combining C-band images via feature selection to predict GSV, as well as studying the temporal dynamics of GSV predicted using selected features (including analysis of long sets of consecutive SAR images) has now become possible.

Several recent studies using multiple Sentinel-1 acquisitions in the context of forest monitoring or classification [33–36] have concentrated on analysis of the temporal patterns of Sentinel-1 data. However, these did not elaborate on the influence of the length of the time series on classification accuracy, perform explicit feature selection trials, or study ways of combining consecutive observations for predicting continuous forest variables.

In the context of forest biomass prediction, research has been conducted on the fusion of Sentinel-1 data with SAR data from instruments operating at different wavelength, as well as with optical data. In particular, the combination of Sentinel-1, ALOS-2, and Sentinel-2 data has produced greater accuracy, although the role of Sentinel-1 data in this case was relatively limited [23]. Another study compared single-date and multi-temporal Sentinel-1 data and demonstrated that multi-temporal Sentinel-1 substantially improved AGB estimation, though with a much smaller adjusted R^2 because of the limitations of the short wavelengths [37]. When using WCM-based approaches for AGB prediction from ASAR hypertemporal data, a relative root mean square error (rRMSE) in the range of 34.2–48.1% was achieved for 1 km pixel predictions [13]. Use of multiple stand-level features was instrumental in achieving an rRMSE of 30% for a Finnish test site [10]. There, stand-level backscatter intensity and other stand-level features such as the standard deviation of the intensities and the averages of the ratios of the intensities of the images from different dates were used as predictor variables, resulting in good GSV prediction accuracies. Several sets of RADARSAT-2 images were studied using machine learning approaches in combination with other SAR and optical data using various features and machine learning approaches, achieving rRMSE values of 42–47% [24].

A general conclusion that can be reached is that the use of a single Sentinel-1 image does not lead to accurate forest biomass mapping. Therefore, the present research concentrates on the fusion of Sentinel-1 and L-band SAR or optical satellite data, while investigating the potential of multitemporal Sentinel-1 datasets. The performance of multi-temporal Sentinel-1 approaches is not yet satisfactory on either the forest plot or stand level. Most critically, in-depth analysis of the potential of very long time series Sentinel-1 data, their seasonal fluctuations (dynamics), and the effects of these fluctuation on the accuracy of GSV prediction is lacking, although the data for multiple years are readily available. Furthermore, a statistically sound method for optimising image and season selections for purpose of increasing prediction accuracy when using dual-pol Sentinel-1 data has not been demonstrated over boreal forests. Reported studies of feature selection are relatively scarce, as well as disparate and non-exhaustive. This motivates us to address these knowledge gaps. In the present study, we use representative reference forest stand-level data to study various GSV prediction models and feature selection approaches, including several that have not been previously studied with SAR data.

Here, we should acknowledge the long history of the operational use of optical area space-borne data, as well as recent studies for forest and agricultural applications, e.g., [38–43] and the potential of upcoming sensors [44]. However, it needs to be taken into account that optical area data are applicable in the boreal zone only during the growing season. Furthermore, the weather conditions often prevent data acquisition during that season. One purpose of the current article is to study and develop methods of utilizing remote sensing data that are applicable independently of weather and lightning conditions

and to take advantage of the seasonal variation of the backscattering mechanism of SAR in forestry applications.

Goals and Contributions of This Study

In this study, we expand on a lineage of forest biomass mapping using multitemporal radar imagery by evaluating the use of large numbers of Sentinel-1 images for a lengthy time period for Finnish boreal forests. We focus on one key variable, GSV (m^3/ha), defined here as the above-ground volume of living stems above stump to the stem top over a specified area. Included are the stem volumes of living trees standing or lying with a diameter at breast height (dbh) of more than 10 cm and height of more than 1.30 m, at which height dbh is measured. GSV is strongly correlated with other forest variables, such as tree basal area, AGB, and total tree carbon content. The latter is approximately $(0.3\text{--}0.4 \text{ t/m}^3) \times \text{GSV}$ in boreal forests depending on the tree species [45,46]. Thus, our results can be compared to the results from other relevant studies that use SAR to predict tree biomass for boreal forests.

The primary and overarching objective of this study was to develop a method for using Sentinel-1 imagery to maximize GSV prediction accuracy. Three supporting and subordinate objectives were addressed: (1) to analyse the effect of the acquisition date, weather conditions, and number of Sentinel-1 images on the GSV prediction accuracy; (2) to test and develop feature selection and transformation methods for backscattering coefficients and multitemporal features such as principal component analysis (PCA), radiometric contrast (RC), mutual information (MI), improved k-Nearest Neighbors (ik-NN), random forests selection (RF-Sel), Lasso and Wrapper approaches, and to evaluate the effects of the features and transformations on the prediction accuracies; and (3) to test a few commonly used optional forest variable prediction methods and their performance in GSV prediction.

These objectives were addressed using extensive analyses of a 3.5-year stack of Sentinel-1 images acquired over a large test site representative of boreal forestland. Both established and advanced statistical and machine learning approaches were used for feature selection and forest biomass prediction, in particular MLR, RF, and SVR. With management inventory needs in mind, we focused on stand-level GSV predictions and their errors, not necessarily on large area level estimates. The criterion was the mean standard error at the stand level when using separate validation data. Note that estimators with large standard errors at the stand level may produce consistent large area estimators, while precise stand-level estimators may be biased when used as large area estimators.

2. Materials

2.1. Study Site

The $50 \text{ km} \times 50 \text{ km}$ study area in Southern Finland was chosen around the Hyytiälä Forestry Field station of the University of Helsinki, with the following centre coordinates: 61.8°N , 24.3°E , as shown in Figure 1. The study site includes forest, agricultural land, lakes, and several population centres. The forests are dominated by Norway spruce (*Picea abies* (L.) H. Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula pendula* Ehrh. and *Betula pubescens* Roth). The soils are mainly glacial drift, with sandy and clayey soils common as well, especially in agricultural areas. Mires and open bogs are frequent. The area is relatively hilly, with terrain variation ranging between 95 and 230 m above sea level.

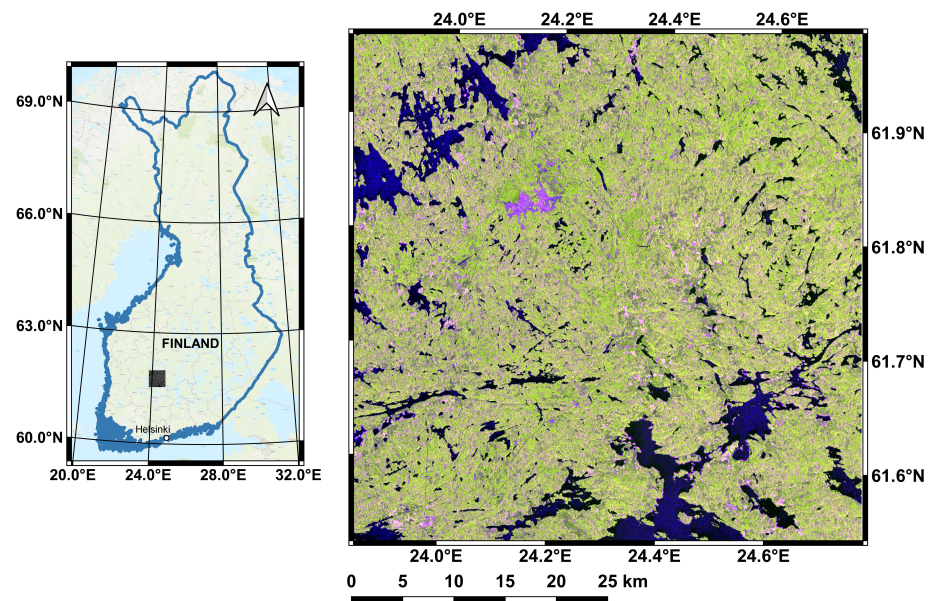


Figure 1. Location of the study site in Southern Finland (left) along with representative Sentinel-1 image as RGB composite (Red: VV, Green: VH, Blue: VV/VH; Date: 9 October 2014).

2.2. Sentinel-1 Data

Ninety-six Sentinel-1 images acquired between 9 October 2014 and 21 May 2018 were used in the study as Level-1 (Ground-Range Detected High-resolution, GRDH) products with VV and VH polarizations. The images were acquired in Interferometric Wide-swath (IW) mode on descending orbits.

The Sentinel-1 images were multilooked with factor 2×2 (range \times azimuth) to obtain images with pixel dimensions corresponding to the 20 m grid spacing. The bilinear interpolation method was used for resampling in connection with orthorectification. A digital elevation model (DEM) from the National Land Survey of Finland was used and averaged to $20 \text{ m} \times 20 \text{ m}$ pixels. Radiometric normalization with respect to the projected area of the scattering element was used to eliminate the topography-induced radiometric variation. In this way, a time series of co-registered images in the terrain-corrected “gamma-naught” coefficient was constructed [47] with a pixel size of $20 \text{ m} \times 20 \text{ m}$. A 3×3 median window was used to filter the speckle noise.

2.3. Ground Reference Data

Stand-level forest resource data collected, prepared, and intended for forest management purposes were used for both the training and validation data. Figure 2 presents the stand-level reference and weather conditions. The data for the test site corresponded to the year 2016 and were obtained from the Finnish Forest Centre [48,49]. These forest resource data consist of predictions from statistical models using airborne laser scanning (ALS) data and data from field measurements. A single remote sensing site has a total area of about 300,000 hectares. From one inventory site, 700–800 circular sample plots or 150–200 new types of larger tree map sample plots are measured. Different plot types can be combined [48]. Field checks on each stand were used to correct noticeable deviations in predictions, e.g., in tree species level predictions, following the practical procedure recommended by the Forest Centre. It is important to note that despite this procedure stand-level predictions always include some errors. The estimated average relative stand level root mean square error (RMSE) of the GSV predictions is around 9–11% [49], with the error depending on the actual GSV. Earlier, new data were produced on every tenth year, and are now measured on every fifth year. The data are updated annually based on growth models and the reported forestry regimes, such as harvests; thus, they are in principle up to date. Because the stand edges may influence prediction, one-pixel morphological erosion was

applied to the stand mask, that is, only pixels separated from the closest stand boundary by at least one pixel were included in the analyses. Further, only stands with areas of at least 1.0 ha after the erosion and with a GSV of $2.0 \text{ m}^3/\text{ha}$ or more were used. The GSV threshold was used to remove stands with recent regeneration cuts from the lengthy time series datasets. In total, 17,762 stands were included in our approach, of which 8881 were randomly selected for training and 8881 remained for validation (Table 1).

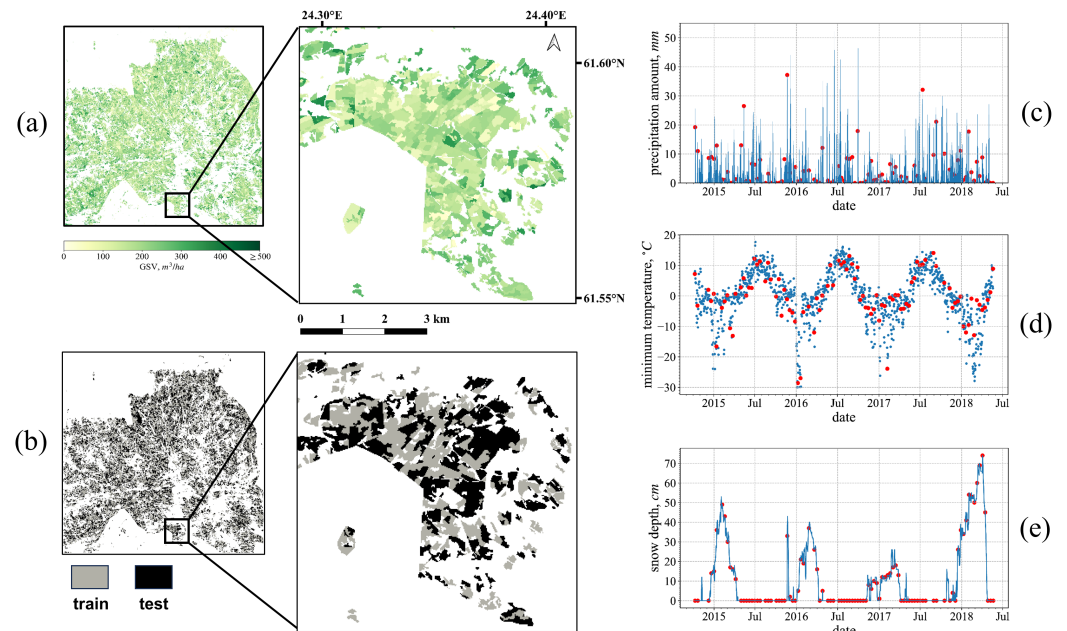


Figure 2. Forest reference data and weather conditions: (a) masks of the forest stands; (b) locations of training and testing stands; (c) three-day accumulated precipitation amount; (d) minimum temperature; and (e) snow depth. Sentinel-1 acquisition times are highlighted with red dot marks, while blue dots indicate the weather history throughout the entire observation period.

Table 1. Ground reference data, including the stand area and the growing stock volume, divided into training and validation sets.

Data Set	Number of Stands	Stand Area, ha			Volume, m^3/ha	
		Median	Mean	Std	Mean	Std
Training	8881	2.16	3.05	3.26	170.39	93.03
Validation	8881	2.20	3.03	2.77	170.70	93.83
Both	17,762	2.20	3.04	3.03	170.55	93.43

3. Methodology

Our approach focused on predicting GSV at the stand level using optional regression techniques, including MLR (Section 3.1.1), SVR (Section 3.1.2), and RF regression (Section 3.1.3).

The general approach is shown schematically in Figure 3. First, a subset of images and optional features was selected, then the backscattering was averaged within each stand for each feature, spatially guided by the forest stand masks. Later, using the stand-level averaged features as the predictors, the response variable GSV was predicted utilizing the different regression approaches.

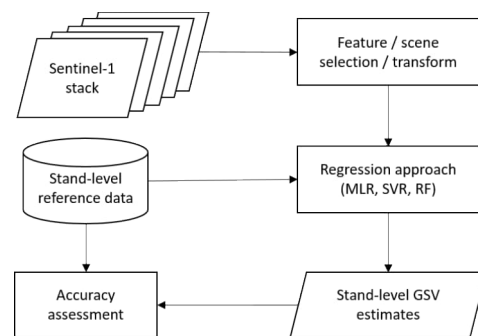


Figure 3. General processing flowchart for growing stock prediction using selected Sentinel-1 images.

We used RMSE as the criterion for evaluating regression performance. To ensure that the evaluation was not affected by the size of the forest stands, stand size after a morphological erosion was used as a weighting factor to normalize the RMSE, which was calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{v}_i - v_i)^2 s_i}{\sum_{i=1}^N s_i}}, \quad (1)$$

where v_i denotes the GSV of stand i , \hat{v}_i is its prediction, s_i the area of the stand i , and N is the total number of stands.

Three supervised GSV prediction methods are described in Section 3.1. The approaches for selecting the optimal subsets of features are described in Section 3.2. The multi-temporal analysis method is described in Section 3.3.

3.1. Methods for Predicting Growing Stock Volume

3.1.1. Multiple Linear Regression

MLR was used as one of the parametric methods for predicting GSV. This is a basic regression approach often used for modeling the relationship between GSV or forest biomass and SAR backscattering. In certain cases, transformations can be applied to the predictor and response variables to further linearize the modeled relationship [15,18,26,27,50–56]. In this study, SAR image features were used as predictor variables and multiple optional transformations of the response variable GSV were tested, including exponential (e.g. with a power of 0.5) and logarithmic, in addition to a non-transformed GSV. The final model was of the form

$$\log(v_i + a) = w_0 + \sum_{j=1}^{N_f} w_j f_{i,j} + \epsilon_i, \quad (2)$$

where v_i is the GSV for stand i , $f_{i,j}$ denotes the SAR feature j for stand i , N_f is the number of the SAR features, w_j with $j \in 0, \dots, N_f$ are the coefficients to be estimated, a is a small positive number, and ϵ_i is random error that is assumed to be independently distributed. In this study, the SAR feature represents the logarithmic scaled stand-level intensity at the VV or VH polarizations, or both. A bias correction was made to GSV predictions due to the use of a logarithmic transformation in model estimation [57]. The correction was based on Taylor series expansion, assuming that ϵ_i is normally distributed and is calculated as half of the variance estimate of ϵ_i . Note that the correction is additive, is made to the prediction on the log scale before back-transforming, and is half the residual variance on the log scale.

To evaluate the separate contributions to GSV prediction, the values of f_n were selected from different scenes (timestamps) and polarizations (VV-only, VH-only, or VV and VH), as was the case for the other regression approaches as well.

3.1.2. Support Vector Regression

Support vector machine (SVM) classification was first proposed by [58] to solve traditional classification problems. It is based on statistical learning theory. Later, [59] extended SVM to regression problems, namely, SVR. This converts a regression problem

to a convex quadratic programming problem. The introduction of the ϵ -insensitive loss, which ignores the error near true values, makes SVR less sensitive to overfitting. Using a nonlinear kernel function, SVR can be transformed from a nonlinear problem into a linear one in a higher-dimensional feature space [60,61].

The Gaussian Radial Basis Function (RBF) kernel was used in our study. This kernel has previously shown good performance in forest biomass estimation tasks [26,62]. As SVR is sensitive to hyperparameters, it is necessary to find the optimal hyperparameters by cross-validation on the training data.

3.1.3. Random Forests

RF is a supervised ensemble learning algorithm that uses a bagging strategy to fit a large collection of classification and regression trees (CART), leading to improved overall modelling prediction ability [63]. First proposed in 1995, it has been widely used for forest classification, change detection, and biomass estimation [64–67].

Compared to SVR and linear regression, RF has multiple advantages: (i) it can avoid overfitting by introducing two layers of randomisation, i.e., bootstrapping when sampling and a random feature subset when splitting nodes [63]; (ii) parallel decision trees accelerate computation while making nonlinear modelling possible; (iii) high-dimensional data can benefit from its automated feature selection, which is especially needed in SAR long time series image processing; and (iv) by using out-of-bag (OOB) data in internal examination of the model, RF can substantially decrease the systematic prediction error without any need for more training data [68].

In the training stage, data subsets for each decision tree are randomly sampled from the original data using bootstrapping, then each decision tree is trained independently until each leaf belongs to one specific label. When splitting nodes, for regression tasks impurity criterion such as mean squared error (MSE) or mean absolute error (MAE) are used to select the splitting variable (feature) and splitting point. In the end, the predictions are averaged over trees to provide the final results.

In addition, the impact of each feature on the impurity of the decision trees can reflect the importance of that feature. This principle is used for feature ranking, namely, RF selection (RF-Sel), along with the other methods elaborated in Section 3.2.

After five-fold cross-validation on a small subset of training data, the number of trees within the model was set to 400 and MSE was chosen as the impurity criterion (variance reduction). To minimise the computational intensity and achieve greater accuracy, the maximum depth of trees was varied from 5 to 35 according to the number of input features. The minimum number of sample units required to split an internal node was set to 5.

3.2. Feature Transformation and Selection

Because the Sentinel-1 time series has 96 images (timestamps) covering more than three years and each image has two bands (VV/VH), considerable redundancy can be expected in the feature space. To circumvent the effects of this redundancy and to reduce the computational burden, feature extraction or reduction can be useful. Here, we describe several approaches for transforming or selecting an optimal subset of SAR features. We discuss both the optimal subset and insights into the choice of the optimal observation dates. By carefully choosing a few specific dates, we expect to achieve a prediction accuracy close to what can be achieved using the entire original time series.

3.2.1. Principal Component Analysis

Principal Component Analysis (PCA) is a common approach for reducing the dimensionality of the feature space [69]. PCA can be used to convert the original features into a set of linearly uncorrelated components using an eigenvalue decomposition. The importance of each component is evaluated by its eigenvalue. The components corresponding to the largest eigenvalues are expected to retain most information about variation in the data. In this way, a subset of principal components can be used as a set of new features,

thereby effectively reducing the original data features while preserving the majority of information. However, because PCA is essentially a feature transformation rather than a feature selection approach, it is not suitable for certain tasks, such as determining the optimal season for acquiring SAR imagery. Therefore, PCA analysis is primarily used as a standard for comparing feature selection approaches, as described below.

3.2.2. Feature Selection Using Radiometric Contrast

Radiometric contrast (RC) is used for choosing a set of images with feature variables that have the greatest radiometric variation between forested and non-forested terrain. This is similar to the approaches used for feature selection in earlier studies, and is preferred because of its simplicity and lack of need for reference data [13]. The essential idea is that SAR image features with greater radiometric contrast between forested and non-forested areas are better able to discriminate different values of GSV due to a greater dynamic range of forest backscattering measurements.

Forested and non-forested areas are each represented by a specific number of stands that have the least or greatest GSV, respectively. First, we sorted the stands in descending order according to their GSV. Denoting the 10 stands with the greatest GSV as the subset \mathcal{P}_{max} and the 10 stands with the least GSV as \mathcal{P}_{min} for a specific polarization of a certain image, the radiometric contrast γ is evaluated as the ratio of the mean amplitude by using the formula below:

$$\gamma = \frac{\bar{A}_{\mathcal{P}_{max}}}{\bar{A}_{\mathcal{P}_{min}}} = \frac{\sum_{i \in \mathcal{P}_{max}} A_i}{\sum_{j \in \mathcal{P}_{min}} A_j} \quad (3)$$

where A_i denotes the stand-averaged feature/backscattering amplitude of stand $i \in \mathcal{P}_{max}$.

Further, all the images were ranked according to the calculated radiometric contrast. The selected images and their relative performance are analysed further in Section 4.2.

3.2.3. Feature Selection Using Mutual Information

Mutual information (MI) is a Shannon entropy-based measure of dependence between random variables. It uses the naive idea that more mutual information means more contributions to prediction. Unlike Pearson linear correlation, mutual information can capture nonlinear dependency [70,71]. The mutual information is calculated as

$$\phi = \sum_{v \in \mathcal{V}} \sum_{f \in \mathcal{F}} p(f, v) \log \left(\frac{p(f, v)}{p(f)p(v)} \right), \quad (4)$$

where f is the SAR feature variable, v is the GSV, p denotes the probability distribution, and ϕ is always non-negative and equals zero when and only when $p(\mathcal{F}, \mathcal{V}) = p(\mathcal{F})p(\mathcal{V})$.

Accurate estimation of the probability distribution is not trivial, particularly from finite samples. Here, we use a k-Nearest Neighbors (k-NN) based approach to estimate the joint distribution $p(\mathcal{F}, \mathcal{V})$ [72]. The number of neighbours k was set to $k = 7$ for this task [70].

3.2.4. Feature Selection Using Lasso

Least Absolute Shrinkage and Selection Operator (Lasso) is another linear regression method [73]. For totals of N_s stands and N_f features, the Lasso loss function is

$$\hat{w} = \operatorname{argmin}_w \left\{ \sum_{i=1}^{N_s} (v_i - w_0 - \sum_{j=1}^{N_f} f_{i,j} w_j)^2 + \lambda \sum_{j=1}^{N_f} |w_j| \right\} \quad (5)$$

where $f_{i,j}$ denotes the SAR feature j for stand i , v_i denotes the corresponding reference GSV, and w_j is the coefficient of feature j . The loss is similar to least squares estimation (LSE) for conventional linear regression; the difference is that in Lasso an $L1$ regularization $\lambda \sum_{j=1}^{N_f} |w_j|$ is applied on the coefficients w_j . The regularization makes some of the weak feature weights equal to 0, thereby achieving selection of important features in high dimensional cases. λ is

the weight to control the strength of regularization. Lasso has known shortcomings; when there is a set of highly relevant variables, Lasso tends to choose one of them while ignoring others, leading to instability of the regression/feature selection process.

3.2.5. Feature Selection Using ik-NN

The well-known k-NN prediction method is well-suited for cases in which multiple target parameters or variables are estimated simultaneously. In addition, during the estimation process k-NN can obtain the weights of different features, and as such can be an optional method for feature selection [10,38–40].

In this study, we utilize an improved k-nearest neighbors (ik-NN) with a genetic algorithm in the context of feature selection. Consider a stand p and denote its k nearest feasible field stands as p_1, \dots, p_k , where the distance is measured in the feature space. The weight $\beta_{i,p}$ of field stand i with respect to stand p is defined as

$$\beta_{i,p} = \frac{1}{d_{p_i,p}^t} \bigg/ \sum_{r \in \{1, \dots, k\}} \frac{1}{d_{p_r,p}^t} \quad (6)$$

$$i \in \{1, \dots, k\}$$

$$= 0 \quad \text{otherwise,}$$

where p_r is a field stand with its index $r \in \{1, \dots, k\}$. We fixed the value of k to be 15 after tests on the validation subset, for which we used the RMSE and mean deviation as criteria. The distance weighting power t is a real number, usually $t \in [0, 2]$; here, we set $t = 1$. A small quantity greater than zero is added to d when $d = 0$.

The distance metric d employed was

$$d_{p_r,p}^2 = \sum_{j=1}^{N_f} w_j^2 (f_{p,j} - f_{p_r,j})^2, \quad (7)$$

where $f_{p,j}$ is the j -th SAR feature variable for stand p , N_f is the number of SAR feature variables, and w_j is the weight of the j -th SAR feature variable.

The values of the elements w_j of the weight vector w were selected using a genetic algorithm [38,39]. The fitness function to be minimized when calculating the w vector for GSV estimation was

$$f(w, \gamma_\delta, \gamma_e, \hat{\delta}, \hat{e}) = \sum_{r=1}^{N_e} \gamma_{\delta,r} \hat{\delta}_r(w) + \sum_{r=1}^{N_e} \gamma_{e,r} \hat{e}_r(w), \quad (8)$$

where:

$\hat{\delta}_r(w)$ is the standard error of the prediction of the forest variable r at the stand level,

$\hat{e}_r(w)$ is the mean deviation of the forest variable r at the stand level,

N_e the number of the forest variables used in the algorithm (here, one variable),

γ_δ and γ_e are fixed constant vectors.

After trial and error tests as well as heuristic trade-offs between the standard error and mean deviation, the values of $\gamma_\delta = 10$ and $\gamma_e = 1$ were selected, resulting in RMSE having greater weight. The other parameters of the genetic algorithm itself were the same as in [38].

As an iterative procedure, genetic algorithms are computationally intensive and may produce a local rather than global optimum. To circumvent this possibility, multiple runs are usually needed to find a near-global optimum.

3.2.6. Feature Selection Using Wrapper Methods

Based on ranking mechanisms, feature selection approaches can be classified into three categories: Filter, Embedded, and Wrapper methods. Filter methods select features by statistically measuring their correlation with the target variable. These methods typically assume the independence of the features. Furthermore, the selection process is not directly

related to subsequent tasks. Although filter methods are fast, they may not yield the best features. In the current study, RC and MI are examples of filter methods. On the other hand, embedded methods incorporate the feature selection process into model training. For instance, RF has a built-in feature selection mechanism. With a fully trained model, the importance of different features can be extracted in parallel. In our study, RF-Sel, Lasso, and ik-NN are typical representatives. In contrast to filter methods, embedded methods can take into account information redundancy among the features.

In addition to filter and embedded methods, we employed a typical wrapper method called sequential feature selection (SFS) in our study. SFS utilizes a forward search direction to select key features from a large number of features. It follows a greedy search approach, where at each iteration the SFS-based wrapper selects one promising feature from the remaining feature pool based on the performance of its proxy task model. If the performance criteria are met, the feature is incorporated into the locally-optimal subset of features for the next iteration. After a certain number of iterations, a collection of key features is obtained. The flowchart illustrating the method is shown in Figure 4. Compared to filter and embedded methods, wrapper methods are generally considered to be more precise; however, they require substantially more computational resources due to their iterative nature and the need to evaluate the performance of multiple subsets of features.

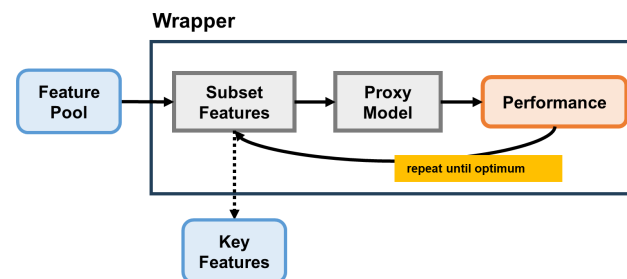


Figure 4. Flowchart of the wrapper method.

For the choice of the proxy task model, we opted for RF regression in our study. The reason for this choice was its nonlinear modelling capability and relatively modest computation workload compared to other methods. To estimate the performance of the proxy model, we employed five-fold cross validation. In this case, the RMSE of GSV prediction was the evaluation criteria. We additionally tested MLR as the proxy model, with the selection performance being slightly worse. SVR was another option; however, the computational requirements appeared to be too heavy to accomplish within a reasonable time.

3.3. Multitemporal Analysis Using “Sliding Window”

The aim of multitemporal analysis of SAR imagery in this study was to assess the prediction accuracy gain achieved on multiple SAR images. These images were used as either sequences of consecutively collected images or as specific images chosen using the feature selection approaches (Section 3.2). For the latter, optimal SAR images could be spread over the three-year observation period. However, using all the data seems impractical, as strong saturation can be observed; indeed, any additional value when using more than 10 images is questionable. The use of such a long time series can be computationally demanding as well. Therefore, we attempted to use RMSE as the criterion in order to determine the best time to acquire SAR images, i.e., the optimal observation season, and to analyze the seasonal characteristics of the SAR image time series.

The size of the sliding window corresponds to the number of images that are aggregated and used to derive the feature variables. Several aggregation approaches are possible in the context of GSV estimation. The first, multitemporal filtering, facilitates reduction of speckle by calculating the combined estimate from several images via averaging [74]. The filtered image is further used to calculate image features to predict GSV. Another approach

performs image-by-image GSV estimation, then merges individual image-based estimates using a multiple regression approach [12]. Individual GSV estimates can be obtained using methods such as inversion of physics-based models [18,75]. However, for this study a set of feature variables derived from SAR images was used directly as input to a multiple regression approach, thereby constituting a possible third approach that seems to be the most practical.

An illustration of the sliding window is shown in Figure 5. The choice of the size of sliding window, i.e., the number of images used to select the feature variables for regression, required separate experimentation. On the one hand, it is expected that including more images will improve GSV prediction performance. On the other hand, the marginal value of additional images can be limited as the image quantity increases beyond a certain number, and there is increased computational intensity associated with larger sliding windows. Thus, several window sizes from 1 to 32 were tested for this study.

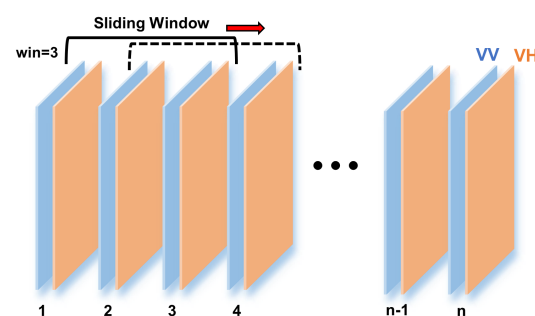


Figure 5. Illustration of the sliding window approach.

4. Results

We started our analysis by assessing the improvement in GSV prediction as the number of Sentinel-1 images increased when using the three GSV prediction approaches. The effect of the image acquisition time on the prediction accuracy and relative performance of the studied prediction methods is reported in Section 4.1. Images were aggregated in the order of their acquisition. Furthermore, the potential of the feature selection and transformation approaches and their effects on the accuracy of GSV prediction are analysed in Section 4.2. The effects of combining a small number of images within the “sliding window” for GSV prediction are reported in detail in Section 4.3.

4.1. Accuracy of GSV Prediction

4.1.1. Single SAR Images and Seasonal Variation of the Predictions

Figure 6 shows the accuracy of the GSV predictions when using feature variables derived from single SAR images acquired from November 2014 to July 2018. The accuracy metrics are shown in the order of image acquisition. RMSE varies within the range 85–94 m³/ha (rRMSE 50–55%) for various polarizations. Overall, the dynamics are quite variable, as expected for C-band imagery, which is strongly affected by environmental conditions. For MLR, the greatest accuracies were obtained using feature variables selected for images for the month of June for 2015, 2016, and 2017. For SVR and RF a similar observation can be made, with an optimum in early 2016. For all image and polarization combinations, the combined VV and VH polarizations provide smaller RMSEs than any single polarization alone, and provide VV-pol somewhat smaller than VH-pol. Scatterplots of the different GSV prediction methods for several selected sets of Sentinel-1 image features are shown in Figure 7.

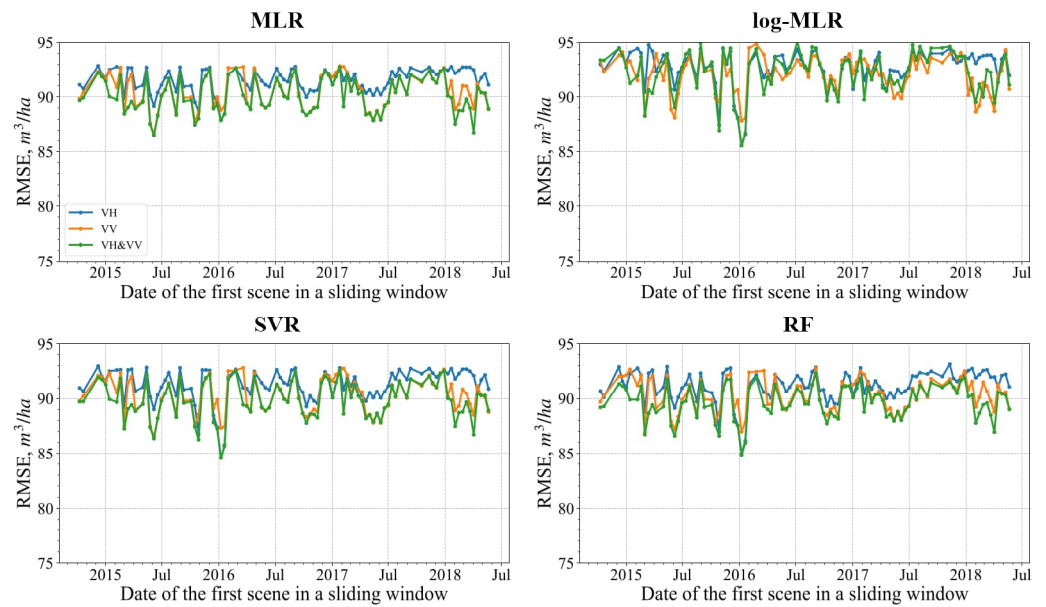


Figure 6. RMSEs of the predicted GSV using individual images (both polarizations) with MLR, log-MLR, SVR, and RF.

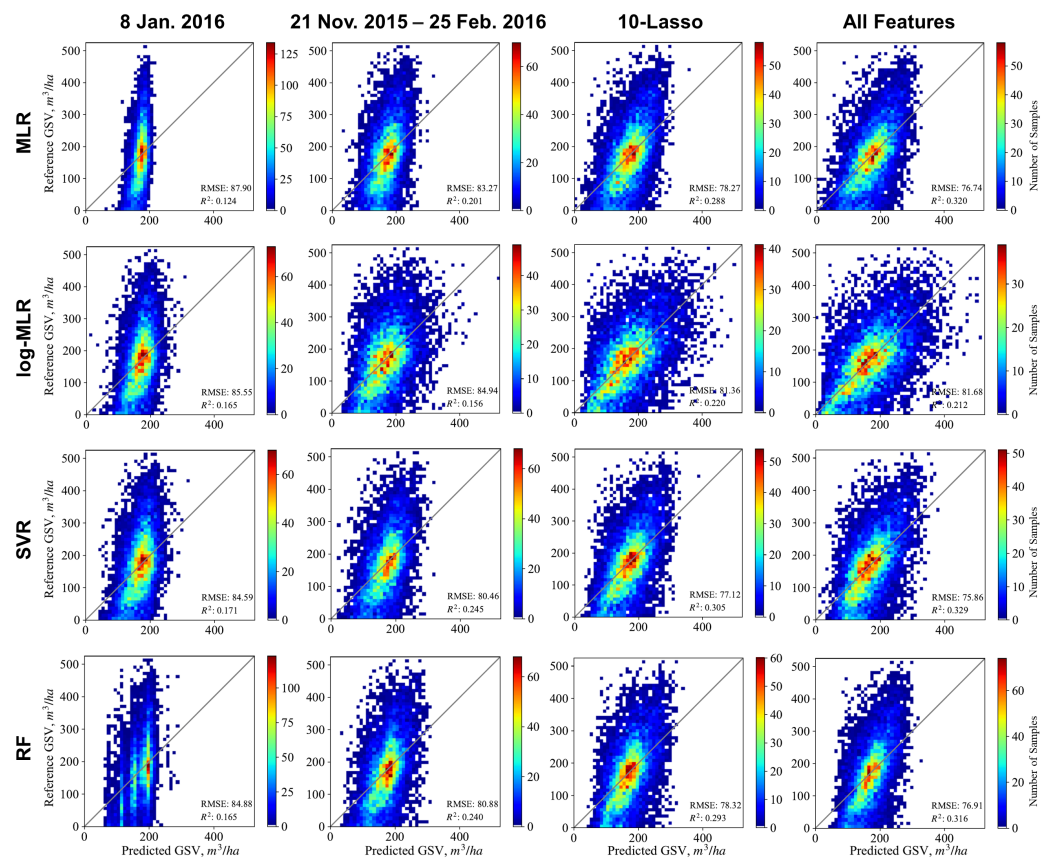


Figure 7. Scatterplots illustrating prediction performance for various methods and features. Rows correspond to the MLR, log-MLR, SVR, and RF methods. The first column shows predictions using only one image (acquired 8 January 2016), the second column shows predictions using eight consecutive scenes (21 November 2015–25 February 2016), the third column shows predictions using the 10 best Lasso features, and the last column shows predictions using all 192 features.

4.1.2. Accumulated SAR Time Series

We further tested the accuracy of the GSV predictions by using feature variables for multiple images (SAR time series) in modelling with MLR, SVR, and RF, and with three polarization options, VH, VV, and VV plus VH. The variations in RMSE for the GSV predictions are illustrated in Figure 8. Here, the images were accumulated in the order of their acquisition. GSV prediction for any given date was performed using all images acquired until the date on x-axis. Altogether, twelve RMSE curves obtained as a function of the number of accumulated images used in GSV prediction are presented in Figure 8.

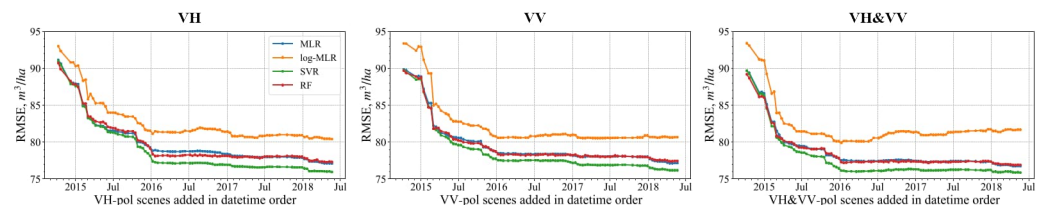


Figure 8. Accuracy of predicted GSV with image features in the order of image acquisition.

A nearly monotonic decrease in RMSE for the GSV predictions with all the polarization combinations and all three prediction methods can be seen. The RMSE rapidly decreases initially, then levels off shortly after July 2015, by which time more than 15 images have accumulated in the SAR time series. RMSE decreases somewhat faster for VV-pol than for VH-pol, with an RMSE of approximately 80 m³/ha for VV and approximately 82 m³/ha for VH for July 2015. Combined VH and VV had a smaller RMSE for GSV prediction than either VV or VH alone, less than 79 m³/ha (rRMSE 47%) for the same date.

For the majority of our experiments, combinations of image sets, polarizations, and prediction methods SVR produced somewhat smaller RMSEs than MLR or RF, about 1.0 m³/ha smaller, whereas there was practically no difference between the RMSEs for MLR and RF except for a short period from 2016 to 2017 for VH-pol.

4.2. Feature Selection

Here, we evaluated the potential of Sentinel-1 for GSV prediction using selected sets of images, in contrast to using images in the order of their acquisition dates as in Section 4.1.2. The idea was to achieve the greatest GSV prediction accuracy using a minimum number of images. In further analyses, we evaluated several feature selection approaches. When particular images were selected, their features were used in stand-level GSV prediction.

The results of GSV prediction using different feature selections and transformation approaches to form multitemporal SAR stacks are presented in Figure 9. Both polarizations were used together; thus, the original feature pool consists of 192 features. The PCA results are shown in Figure 9 as a benchmark. In the figures, only the first 50 components are shown, as there is no accuracy improvement beyond this number.

The SVR method produced slightly smaller RMSEs than MLR or RF, similar to the results in Section 4.1.2 (Figure 8). RF produced somewhat smaller RMSEs than MLR. In particular, MLR had larger RMSEs when using the initial 20–25 features and the features selected using radiometric contrast and mutual information.

All three studied regression methods demonstrated smaller RMSEs when using Lasso, RF-Sel, and Wrapper compared to RC, MI, or ik-NN.

Interestingly, GSV prediction accuracy with RF started to decrease when more than 20 PCA features were included in the model. A likely reason for this is that the PCA tail components consist largely of noise.

Other approaches, particularly MLR and SVR, appeared to be more robust in this respect, with consistent decreases even when incorporating noise-like PCA features. The achieved “plateau” of RMSE was approximately 75.1 m³/ha (rRMSE 44%).

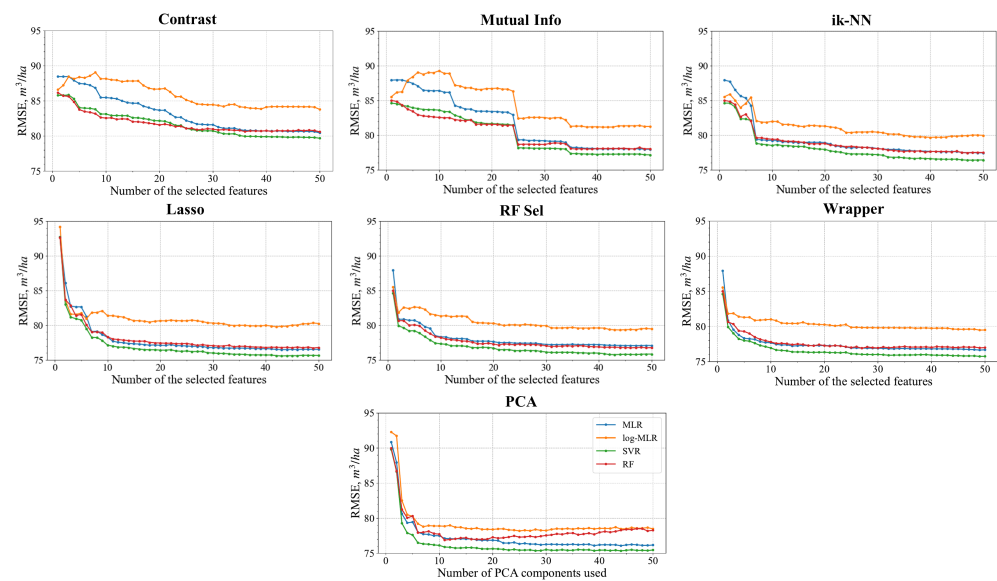


Figure 9. RMSE dependence for various feature sets. All 192 features (VH+VV) are ranked, with the top 50 shown in the figure. Features are ranked using RC, MI, ik-NN, RF-Sel, Lasso, and PCA, respectively. All four regression methods are used in the evaluation.

Among all examined feature selection methods, Lasso, RF-Sel, and Wrapper had similar levels of performance. The RMSEs for all regression methods reached less than $78 \text{ m}^3/\text{ha}$ when only the top 20 features were used. With only 20 features, we were able to obtain a prediction accuracy similar to the PCA upper bound benchmark. The performance of ik-NN was slightly worse, achieving $79 \text{ m}^3/\text{ha}$ RMSE for GSV prediction. The accuracy measure of MI exhibited an inconsistent “stepped” trend of decrease, suggesting that this feature selection approach is not ideal, with slow convergence to smaller RMSE values. Radiometric contrast appeared to be the least effective in selecting features, with RMSE decreasing at the slowest pace and not reaching the best possible values within the 50 top features.

This indicates that a blind feature selection approach with no reference data, such as RC, is generally suboptimal. However, in the absence of reference data it was the only approach with RMSEs that were better than the prediction results in Figure 8, where no feature selection was done at all.

To facilitate better comparison of the Lasso, RF-Sel, and Wrapper approaches, we show the respective RMSE dependencies in Figure 10. It is evident from the plot that wrapper has a smaller RMSE, particularly when fewer than 10 features are selected. This suggests that the SFS mechanism helps to improve the selection performance of the wrapper method even though it uses the same proxy model as RF-Sel. However, as the number of selected features increases, the RMSE curves of all studied selection approaches tend to converge. This implies two points: (1) wrapper can be considered the best feature selection in the context of our study, particularly when a small number of features is required; (2) when more than the top 10 features are allowed in a prediction, the superiority of wrapper is less apparent. Additionally, considering that the heavy computational workload associated with wrapper grows exponentially with the number of features selected, we recommend using Lasso instead in such cases. Lasso is significantly faster and demonstrates good feature selection performance. In Figure 7, we present the scatter plots using the top 10 Lasso features, which provides GSV prediction performance similar to the scenario in which all features are used. This further supports our conclusion on the efficiency of the Lasso approach for SAR feature selection.

Upon analyzing the feature selection performance using various polarizations, we found that the effect of polarizations was relatively small, especially for Lasso, RF-Sel, and Wrapper. These algorithms consistently selected appropriate features for GSV prediction

regardless of whether VV-pol or VH-pol was utilized. Higher accuracy was achieved by incorporating both VH and VV polarizations. MI had the strongest sensitivity in feature selection with respect to different polarizations. Feature selected using VV-pol was much better than with VH-pol in terms of overall prediction performance. This resulted in the deteriorated selection performance of VH and VV in combination. This phenomenon can be clearly seen in Figure 11a. Regarding RC-based feature selection, it was difficult to achieve competitive performance regardless of the polarization used.

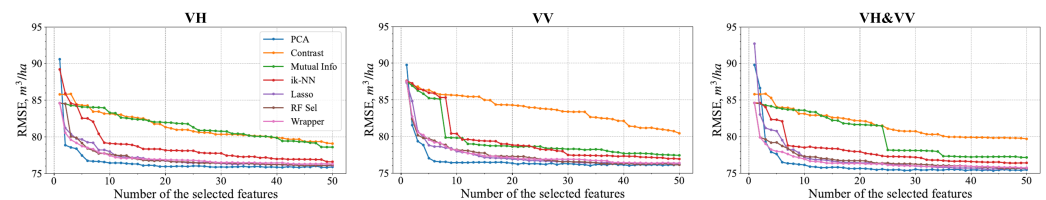


Figure 10. RMSE dependence on the number of input features at various polarizations with SVR as the regression method.

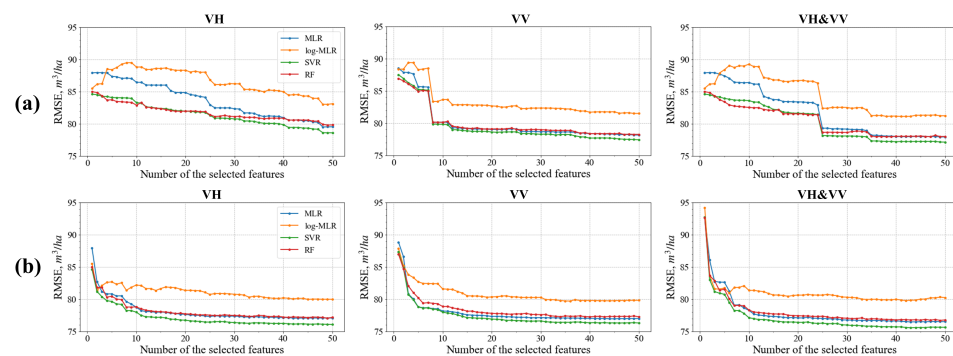


Figure 11. RMSE dependence on the number of input features at various polarizations with MLR/RF/SVR as regression methods: (a) Mutual Information feature selection and (b) Lasso feature selection.

In summary, the wrapper method demonstrates the most accurate feature selection, consistently yielding the best GSV accuracy for all regression models. Lasso proved a viable option when computational resources are limited. Features selected using VH polarization are slightly better than those selected using VV polarization. Lastly, SVR is confirmed to be the most accurate regression method in our study.

Table 2 shows the feature selection results for the three polarization combinations. For both VV and VH, the key features belonged mostly to images acquired under frozen conditions. However, several “dry summer” images acquired in May and June were included by the most accurate Lasso and ik-NN approaches. We performed additional experiments using Monte Carlo testing of random combinations of 10 images representing frozen ground conditions (VV and VH features combined) in GSV prediction using the SVR approach. After fifty trials, the RMSE was $79.3 \pm 1.1 \text{ m}^3/\text{ha}$ and never exceeded the accuracy of the top 10 Lasso approach of $77.1 \text{ m}^3/\text{ha}$ when all images were used (Table 2). The corresponding accuracy for 10 ik-NN selected features using all possible images reached $78.5 \text{ m}^3/\text{ha}$, exceeding the absolute majority of trials with ten exclusively frozen condition images. Similar dynamics were observed for other regression approaches, indicating that the idea of using dry summer images in addition to frozen condition images should be examined in further studies. Cross-pol performed slightly more accurately than co-pol. Several images, such as those acquired on March 2nd 2015 and January 8th 2016, had both polarizations among the selected features, indicating that near-optimal GSV prediction can be achieved while using only a few carefully selected images instead of the whole SAR image stack. This observation is in line with PCA feature selection, as discussed above.

Table 2. Feature selection results using several feature selection approaches.

Contrast			Mutual Info			ik-NN			Lasso			RF-Sel			Wrapper		
Image #	Timestamp	Pol.	Image #	Timestamp	Pol.	Image #	Timestamp	Pol.	Image #	Timestamp	Pol.	Image #	Timestamp	Pol.	Image #	Timestamp	Pol.
VH-pol only, 96 images, 96 features																	
33	2016-01-20	VH	32	2016-01-08	VH	9	2015-03-02	VH	32	2016-01-08	VH	32	2016-01-08	VH	32	2016-01-08	VH
51	2016-10-10	VH	33	2016-01-20	VH	33	2016-01-20	VH	7	2015-02-06	VH	33	2016-01-20	VH	87	2018-02-02	VH
9	2015-03-02	VH	26	2015-10-28	VH	32	2016-01-08	VH	92	2018-04-03	VH	87	2018-02-02	VH	26	2015-10-28	VH
32	2016-01-08	VH	9	2015-03-02	VH	38	2016-04-13	VH	33	2016-01-20	VH	92	2018-04-03	VH	8	2015-02-18	VH
38	2016-04-13	VH	30	2015-12-15	VH	11	2015-03-26	VH	8	2015-02-18	VH	26	2015-10-28	VH	66	2017-04-08	VH
25	2015-10-16	VH	31	2015-12-27	VH	94	2018-04-27	VH	87	2018-02-02	VH	7	2015-02-06	VH	92	2018-04-03	VH
63	2017-03-03	VH	25	2015-10-16	VH	64	2017-03-15	VH	29	2015-12-03	VH	31	2015-12-27	VH	33	2016-01-20	VH
93	2018-04-15	VH	1	2014-10-09	VH	88	2018-02-14	VH	26	2015-10-28	VH	67	2017-04-20	VH	7	2015-02-06	VH
37	2016-04-01	VH	16	2015-06-06	VH	87	2018-02-02	VH	11	2015-03-26	VH	9	2015-03-02	VH	51	2016-10-10	VH
68	2017-05-02	VH	37	2016-04-01	VH	61	2017-02-07	VH	71	2017-06-07	VH	85	2018-01-09	VH	29	2015-12-03	VH
VV-pol only, 96 images, 96 features																	
9	2015-03-02	VV	9	2015-03-02	VV	25	2015-10-16	VV	32	2016-01-08	VV	9	2015-03-02	VV	9	2015-03-02	VV
51	2016-10-10	VV	51	2016-10-10	VV	67	2017-04-20	VV	11	2015-03-26	VV	92	2018-04-03	VV	87	2018-02-02	VV
96	2018-05-21	VV	33	2016-01-20	VV	9	2015-03-02	VV	92	2018-04-03	VV	32	2016-01-08	VV	32	2016-01-08	VV
33	2016-01-20	VV	32	2016-01-08	VV	23	2015-09-10	VV	7	2015-02-06	VV	87	2018-02-02	VV	91	2018-03-22	VV
50	2016-09-28	VV	16	2015-06-06	VV	15	2015-05-25	VV	38	2016-04-13	VV	91	2018-03-22	VV	38	2016-04-13	VV
49	2016-09-16	VV	54	2016-11-15	VV	69	2017-05-14	VV	31	2015-12-27	VV	51	2016-10-10	VV	8	2015-02-18	VV
21	2015-08-17	VV	52	2016-10-22	VV	32	2016-01-08	VV	89	2018-02-26	VV	38	2016-04-13	VV	89	2018-02-26	VV
52	2016-10-22	VV	87	2018-02-02	VV	51	2016-10-10	VV	91	2018-03-22	VV	33	2016-01-20	VV	21	2015-08-17	VV
53	2016-11-03	VV	41	2016-05-31	VV	90	2018-03-10	VV	8	2015-02-18	VV	7	2015-02-06	VV	29	2015-12-03	VV
41	2016-05-31	VV	42	2016-06-12	VV	72	2017-06-19	VV	9	2015-03-02	VV	88	2018-02-14	VV	85	2018-01-09	VV
VH-pol and VV-pol combined, 96 images, 192 features																	
33	2016-01-20	VH	32	2016-01-08	VH	32	2016-01-08	VH	11	2015-03-26	VH	32	2016-01-08	VH	32	2016-01-08	VH
51	2016-10-10	VH	33	2016-01-20	VH	93	2018-04-15	VH	32	2016-01-08	VH	92	2018-04-03	VV	92	2018-04-03	VV
9	2015-03-02	VH	26	2015-10-28	VH	64	2017-03-15	VH	7	2015-02-06	VH	33	2016-01-20	VH	38	2016-04-13	VV
9	2015-03-02	VV	9	2015-03-02	VV	10	2015-03-14	VH	32	2016-01-08	VV	87	2018-02-02	VV	7	2015-02-06	VV
32	2016-01-08	VH	9	2015-03-02	VH	11	2015-03-26	VH	29	2015-12-03	VH	91	2018-03-22	VV	70	2017-05-26	VV
38	2016-04-13	VH	51	2016-10-10	VV	18	2015-06-30	VV	7	2015-02-06	VV	7	2015-02-06	VV	91	2018-03-22	VV
25	2015-10-16	VH	30	2015-12-15	VH	92	2018-04-03	VV	92	2018-04-03	VV	32	2016-01-08	VV	26	2015-10-28	VH
51	2016-10-10	VV	31	2015-12-27	VH	59	2017-01-14	VH	89	2018-02-26	VV	26	2015-10-28	VH	7	2015-02-06	VH
63	2017-03-03	VH	25	2015-10-16	VH	62	2017-02-19	VH	71	2017-06-07	VH	38	2016-04-13	VV	32	2016-01-08	VV
93	2018-04-15	VH	1	2014-10-09	VH	89	2018-02-26	VV	26	2015-10-28	VH	87	2018-02-02	VH	67	2017-04-20	VH

4.3. Sliding Window Analysis

Our sliding window analyses started by selecting a suitable window size for further tests (Section 3.3). For this purpose, we used SVR as the prediction method and tested several window sizes from one (corresponding to only one image) to as many as 15 images.

The results of GSV prediction using sliding windows of different sizes are shown in Figure 12. The date of the first image within a sliding window is shown on the x-axis. The RMSE varied from 85.3 m³/ha to 93.5 m³/ha, with an average RMSE of ~90 m³/ha for a sliding window size N_w equal to one, that is, when only one image at a time was used to predict GSV. The greatest accuracy was achieved using the image acquired on 8 January 2016. No pronounced RMSE seasonal dynamics can be observed. Furthermore, as N_w grew larger, the RMSE gradually decreased. The largest RMSE with $N_w = 4$ was about 92 m³/ha. The largest RMSE was smaller than 88 m³/ha with $N_w = 15$.

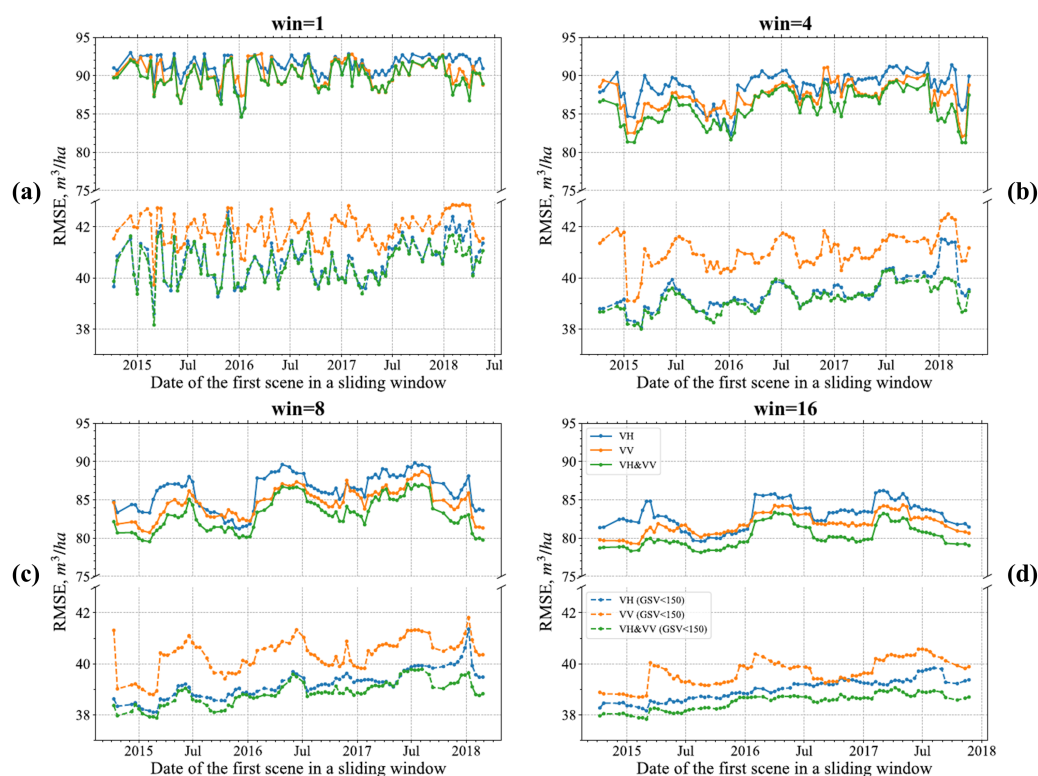


Figure 12. RMSE of GSV prediction as a function of image acquisition date for various sliding windows when using the SVR approach. The window sizes are 1, 4, 8, and 16 (shown in (a–d), respectively). Solid curves show the RMSE for experiments with all stands, while dashed curves show the RMSE results when using stands with biomass of less than 150 m³/ha.

The seasonal patterns indicated by the RMSE curves first become apparent and then more pronounced when expanding the sliding window size.

For example, in Figure 12b the seasonal pattern is not very clear, whereas in Figure 12c annual variations can be readily observed; the local minima of the RMSE curves are clearly visible in every winter, while the local maxima can be found every summer. When the window size is very large, approaching six months, the second peaks are almost eliminated and the three main annual peaks become clearly visible (Figure 12d).

Note that the x-axis tick marks denote acquisition dates for the first Sentinel-1 image within each sliding window, whereas all images within the window contribute to GSV prediction. The temporal resolution decreases when the window size increases, with the RMSE temporal variation becoming smooth. The width of the extreme is particularly affected. Consider the situation around January 2016 as an example. When the window

size is small, i.e., 1 to 4, there is a strong temporal fluctuation in RMSE, and the “valleys” (e.g., 8 January 2016) are sharp, as shown in Figure 12a. When the window size increases to half a year in Figure 12d, the fluctuation becomes smoother, as do the major “peaks”. The reason for this is that the image acquired on 8 January 2016 contributed strongly to GSV prediction with all sliding windows that include this image, exhibiting large prediction accuracies. This complicated the selection of an optimal season for GSV prediction. Thus, a balance between preserving seasonal patterns without sacrificing temporal resolution needs to be made. Based on our trade-off analysis, a window size of eight is most suitable. Using this window size, we proceeded to analyse GSV prediction using different polarizations and estimation methods.

The actual date of the RMSE minimum can be somewhat delayed compared to the date shown on x-axis due to larger window size. A window size of eight corresponds to approximately three months in the Sentinel-1 time series. The RMSE minimum around December means that the optimal timing for SAR images acquisition is in the winter around January and February. The peaks around May indicate that the acquisition times in the summer around June and July are not optimal for GSV prediction in boreal forests when using C-band SAR time series imagery.

Different Regressions over an Eight-Image Sliding Window

In Figure 12, a window size of eight is suggested as the best sliding window size. To study the influence of different regression methods on the sliding window size, we used MLR and RF in addition to SVR in this subsection. As shown in Figure 13, the seasonal pattern is obvious with all regression methods. Seasonal RMSE extrema occurred every summer and winter. The overall RMSE levels and dynamic range were nearly the same for all regression approaches, with MLR underperforming in the autumn of 2016. The differences among polarizations were more apparent, favoring VV over VH and overall suggesting combined use of both polarization channels.

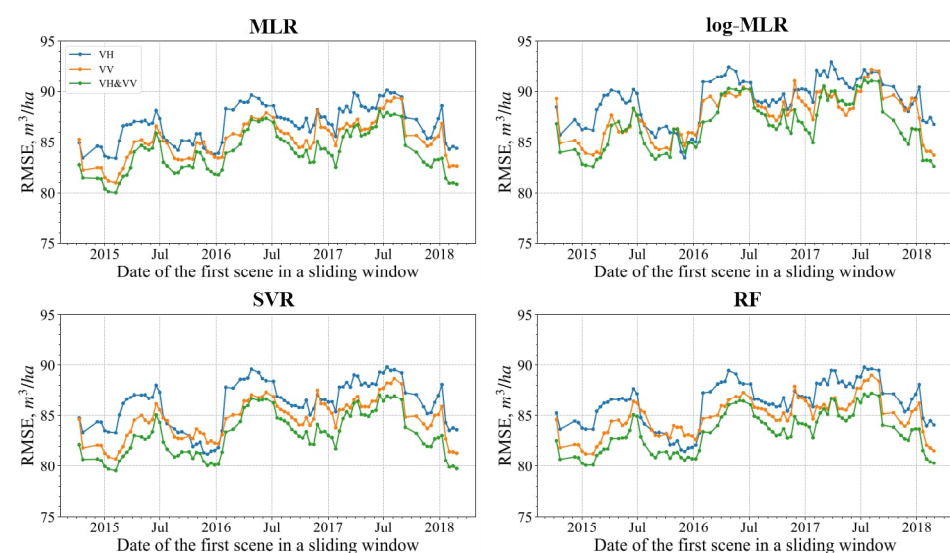


Figure 13. RMSEs of various prediction methods using an eight-image sliding window.

In summary, our sliding window analysis provided a more detailed perspective on the temporal dynamics and facilitated our determination of the optimal timing for image acquisition.

5. Discussion

5.1. On GSV Prediction Methods

Among the different prediction methods, SVM, log-MLR, and RF showed similar temporal dynamics, with no strong differences in GSV prediction. MLR with no log-

transformation was less accurate, and seems to be the least suitable. RF and SVR produced somewhat greater accuracy than log-MLR in the majority of our experiments. MLR with log-transformation of the target variable had somewhat greater variance particularly for larger GSV values; however, it provided the most accurate explanatory model for greater GSV values (Figure 7). No earlier benchmarking studies with C-band data in boreal forests could be found in the literature, while studies with C-band data have been reported for tropical forests [26].

When using only one polarization as the feature variable, no large differences in accuracy were found, with VV alone providing slightly more accurate GSV predictions than VH alone in non-frozen conditions and VH being more optimal in frozen conditions. When limiting analysis to stands with only smaller values of GSV, VH polarization starts to produce consistently more accurate results. In the small-biomass scenario VH is more sensitive to biomass, as cross-pol mostly contains volumetric scattering in forest environments. This discrepancy can be explained by somewhat smaller extinction at VV, contributing to its greater sensitivity to large-biomass forest when the ground is not visible. Under frozen conditions, when the ground is more visible, VV becomes suboptimal compared to VH even when all forest stands are considered, as VV-pol contains both scattering from the ground and volumetric scattering. For both scenarios of GSV values, the use of both polarizations in combination provided the most accurate predictions. The observation that VH is more suitable is supported by earlier studies; however, no analysis of different biomass classes was reported in those studies.

5.2. The Effects of Seasonality on GSV Prediction

The observation that C-band backscatter images acquired in frozen conditions are more suitable for forest prediction in boreal forests has been reported in earlier studies with ERS by [14,20] and ASAR [13]. Our study confirms this observation using Sentinel-1 images, though the difference was relatively weak when single images were used. Moreover, our feature selection experiments (when the 10 most accurate features were selected using the Lasso and ik-NN approaches) indicate that adding dry summer images in addition to frozen condition images provides more accurate GSV predictions compared to using exclusively frozen condition images. Naturally, combining dry summer SAR images with frozen SAR images provides a more comprehensive dataset that captures seasonal variations and various levels of canopy penetration, facilitating data fusion. In our opinion, important contributing factors include deeper penetration into frozen canopy in winter (sensitivity to somewhat larger GSV values) and better sensitivity to relatively biomass within dry summer images. On the other hand, when only dry summer images are used the SAR signal becomes saturated very quickly, and separation of moderate and large GSV becomes impossible. Autumn and spring images can be most problematic in this context, as rainy conditions reduce radiometric contrast between tall vegetation and sparsely vegetated areas, a finding that is in line with previous studies [13].

Using the sliding window approach suggested in this study for combining features calculated from consecutive Sentinel-1 images as inputs to the regression approach was able to facilitate detection of strong seasonality in GSV predictions compared to using single images. This difference is clearly visible in Figure 12. It is important to note that seasonality is not studied here at the level of backscatter (input variable), but rather at the level of GSV prediction (output variable). Using assemblages of features calculated from consecutive images with varying widths of sliding windows suggests there are optimal timing periods when Sentinel-1 images should be acquired to produce optimal GSV predictions. The use of a longer sliding window permits compensation for the presence of features from suboptimal images and taking advantage of “multitemporal variability” within multi-variable regression models, as features calculated from several consecutive images can be used. Our results suggest that even autumn and spring images can be used in principle. This is in contrast with single image retrieval followed by weighting estimates according to accuracy of predictions from individual images [13]. In the latter approach, the most

accurate strategy is to completely discard autumn and spring images, when the forest biomass to C-band SAR relationship can be most unpredictable. Our results with multi-variable regression models indicate that late autumn and early spring datatakes can be included. For the window size of eight, which corresponds to three months of observations, the optimal start can be early as December and as late as February based on three years of observations.

The use of PCA and several feature selection approaches confirms that a subset of features calculated using a subset of Sentinel-1 images can be enough to achieve a near optimal GSV prediction. Interestingly, feature selection using MI and RC led to the least accurate GSV predictions. However, RC is the most feasible approach for image selection, particularly when no reference data are available, and is often used for image removal or weighting [13]. Lasso was the most suitable option for selecting the features for GSV prediction, and approached PCA-based performance in terms of how quickly the near-optimal prediction was achieved. Lasso was closely followed by the ik-NN method. Based on feature ranking with majority voting, the Sentinel-1 image #32 acquired on 8 January 2016 was the most popular. In addition to images acquired in winter, images acquired in the spring and autumn were sometimes selected. Among the most selected features, VH appeared somewhat more often than VV. No earlier benchmarking studies on feature selection with C-band data in boreal forests could be found in the literature.

5.3. Comparison with Previous Studies in Terms of GSV Accuracy and Unit Size

Presently, the number of relevant publications on C-band-based prediction of GSV is relatively limited compared to L-band and multisensor studies, primarily because of the relatively small prediction accuracy. The role of C-band data, particularly Sentinel-1, is quite minor in the multi-sensor studies where C-band images are used along with other SAR and optical datasets [23,24], although accurate results have been reported with a combination of sensors. The results are quite unsatisfactory when using C-band data alone. Our results compare favourably with other published research using C-band SAR data in terms of GSV prediction accuracy [4,5,12–14,23,24], and are more representative in terms of reference data, feature selection approaches, and studied regression models, and compared to the majority of other studies use smaller measurement units. Below, we compare our experiments with several representative studies to highlight the relevance of our results.

Use of a parametric semiempirical model in a study by [14] with set of ERS-1 VV-pol images showed that rRMSEs of 25–30% can be obtained in boreal forests with a forest block size larger than 20 ha. Combining ERS-1 images acquired in November and August 1993 was considered optimal, which supports our observation that C-band data acquired in non-winter seasons can be used. The accuracy with smaller of forest stands was much lower.

In a study by [12] with ERS-1 data in boreal forests pursuing GSV prediction using inversion of a parametric physics-based model, limiting analysis to stands larger than 10 ha led to rRMSEs of 34–71%, with the most accurate estimates for mid-December and mid-March. This result is in agreement with our Sentinel-1 images selected using different feature selection approaches. However, incorporating stands with sizes as small as 1 ha increased the rRMSE to more than 120% for separate images. Combining predictions from individual SAR images using multiple regression facilitated a strong reduction in RMSE. However, when ERS-1 images were mixed with JERS (L-band SAR) it was difficult to judge specific contributions, especially because JERS data are much more suitable for GSV prediction.

The BIOMASAR algorithm [13], which has WCM model inversion at its core, was used with hypertemporal C-band ASAR data in ScanSAR mode (inferior to Sentinel-1 in terms of resolution, but exhibiting high radiometric quality due to a large number of looks) for boreal forest GSV prediction, resulting in rRMSEs between 34.2% and 48.1% at a 1 km pixel size. Larger errors were obtained at 100 m (47.7–60.4% over Swedish test sites and 70.9–96.2% over Central Siberian test sites).

A study of boreal forests in the Krasnoyarsk region of Russia [24] tested, among other experiments, time series of strip map RADARSAT-2 data acquired in regular and ultrafine modes for AGB prediction, achieving rRMSEs of 42% and 47% in AGB prediction with a machine learning approach. The mean AGB was approximately 86 t/ha. The time series dynamics were not explored. AGB was predicted using various cross-pol ratios and Haralick features in addition to the average backscatter.

In a recent comprehensive study with several satellite C-band sensors, the rRMSE was approximately 90% when predicting AGB [76]. However, the error for branch biomass was smaller at around 80%. In the case of VV polarization, which included ERS-2 images acquired in all seasons, the error was greater than 100%. The prediction performance of C-band data depended strongly on the imaging conditions and varied in a large range. The smallest errors were associated with images acquired under stable frozen conditions in winter, which is similar to our findings. The greatest errors were generally associated with images acquired in periods of snow melt in the spring, which is consistent with our results when using separate SAR images. As ERS-2 has only VV-pol, no separate polarization analysis was possible. It is important to keep in mind that the average and peak values of GSV in Hyytiälä forest are somewhat smaller compared to Remningstorp forest [76].

When comparing our findings to previous studies, as well as to [10], where 30% RMSE was achieved using only twelve Sentinel-1 images, it is important to note that we used only stand-level backscatter as an independent variable. No ratios or textural features were used, as the aim was to analyse the seasonal dynamics and optimal timing of Sentinel-1 data backscatter. It is very likely that RMSEs would further decrease if other stand-level variables (features) were included in the analysis, which is in line with other previous studies ([10,77]). This issue needs extensive additional analysis. Furthermore, we did not pursue analysis of GSV prediction accuracy as a function of unit area size; however, we expect the prediction accuracy to improve as the unit area size becomes larger, similar to [12,14,76].

6. Conclusions

The utility of very long times series of Sentinel-1 SAR data for boreal forest GSV prediction was tested. We discovered seasonality in GSV predictions from multitemporal C-band time series, suggesting an optimal timing for image collection in early spring and autumn. Several prediction methods led to similar results. The minimum RMSE was 75.6 m³/ha (rRMSE 44%) when all 96 images were used. Temporal aggregation of images over one year greatly decreased RMSEs compared to a single image. Feature selection facilitated nearly optimal results when using only 10 images (45% relative RMSE) instead of the entire SAR stack, resulting in reduced computational complexity. Among the feature selection approaches, Lasso produced the greatest prediction accuracy, followed by ik-NN and RF. When an eight-image sliding window time series was used instead of feature selection, a relative RMSE of 47% was achieved when the timing of image acquisition was optimal.

A comparison to prior works indicates that no one has yet studied the temporal dynamics of GSV prediction in boreal forests (or elsewhere) in the fashion presented here, i.e., using assemblages of consecutive Sentinel-1 images (a “sliding window” approach). Use of a sliding window resulted in much stronger annual periodicity in GSV prediction compared to single images at the 1 ha stand level. Feature selection approaches have been used very sporadically, and the methods used here (correlation, radiometric contrast) were suboptimal, as indicated by our experiments. The possibility of combining winter and dry summer images instead of using only winter images needs further investigation, as our feature selection experiments support the idea of combining corresponding features within a multivariate model where temporal variability is inherently captured.

Author Contributions: Conceptualisation: S.G., O.A. and E.T.; methodology: S.G. and E.T.; validation: O.A. and R.E.M.; data curation: E.T. and Y.R.; writing—original draft preparation: S.G., O.A. and E.T.; writing—review and editing: R.E.M., J.P., H.G. and W.S.; supervision: O.A. and J.P.;

funding acquisition: J.P., H.G. and W.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (Grant No. 61801221, 62001229) and by the China Postdoctoral Science Foundation (Grant No. 2020M681604), as well as by Aalto University. O.A. was supported by the ESA Forest Carbon Monitoring project funded by the European Space Agency.

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

Abbreviations

The following abbreviations are used in this manuscript:

AGB	above-ground tree biomass
ALS	airborne laser scanning
CART	classification and regression trees
DEM	digital elevation model
ESA	European Space Agency
GRDH	ground-range detected high-resolution
GSV	growing stock volume
ik-NN	improved k-nearest neighbors
IW	interferometric wideswath
Lasso	least absolute shrinkage and selection operator
LSE	least squares estimation
MAE	mean absolute error
MI	mutual information
MLR	multiple linear regression
MSE	mean squared error
OOB	out-of-bag
PCA	principal component analysis
RBF	radial basis function
RC	radiometric contrast
RF	random forests
RF-Sel	random forests selection
RMSE	root mean squared error
rRMSE	relative root mean squared error
SAR	synthetic aperture radar
SFS	sequential feature selection
SVR	support vector regression
SVM	support vector machines
WCM	water cloud model

References

1. Rosenqvist, A.; De Grandi, G.; Rauste, Y.; Shimada, M.; Chapman, B.; Shimada, A.; Saatchi, S. The global rain forest mapping project a review. *Int. J. Remote Sens.* **2000**, *21*, 1375–1387. [\[CrossRef\]](#)
2. Wagner, W.; Luckman, A.; Vietmeier, J.; Tansey, K.; Balzter, H.; Schmullius, C.; Davidson, M.; Gaveau, D.; Gluck, M.; Le Toan, T.; et al. Large-scale mapping of boreal forest in SIBERIA using ERS tandem coherence and JERS backscatter data. *Remote Sens. Environ.* **2003**, *85*, 125–144. [\[CrossRef\]](#)
3. Rodríguez-Veiga, P.; Quegan, S.; Carreiras, J.; Persson, H.; Fransson, J.; 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. [\[CrossRef\]](#)
4. Sinha, S.; Jeganathan, C.; Sharma, L.K.; Nathawat, M.S. A review of radar remote sensing for biomass estimation. *Int. J. Environ. Sci. Technol.* **2015**, *12*, 1779–1792. [\[CrossRef\]](#)
5. Global Forest Observations Initiative. *Integrating Remote-Sensing and Ground-Based Observations for Estimation of Emissions and Removals of Greenhouse Gases in Forests: Methods and Guidance from the Global Forest Observations Initiative*; Group on Earth Observations: Geneva, Switzerland, 2014.
6. Sarker, M.L.R.; Nichol, J.; Iz, H.B.; Ahmad, B.B.; Rahman, A.A. Forest Biomass Estimation Using Texture Measurements of High-Resolution Dual-Polarization C-band SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3371–3384. [\[CrossRef\]](#)

7. Bourgoign, C.; Blanc, L.; Bailly, J.S.; Cornu, G.; Berenguer, E.; Oszwald, J.; Tritsch, I.; Laurent, F.; Hasan, A.F.; Sist, P.; et al. The Potential of Multisource Remote Sensing for Mapping the Biomass of a Degraded Amazonian Forest. *Forests* **2018**, *9*, 303. [\[CrossRef\]](#)
8. Chen, L.; Wang, Y.; Ren, C.; Zhang, B.; Wang, Z. Optimal Combination of Predictors and Algorithms for Forest Above-Ground Biomass Mapping from Sentinel and SRTM Data. *Remote Sens.* **2019**, *11*, 414. [\[CrossRef\]](#)
9. Reis, A.A.D.; Franklin, S.E.; de Mello, J.M.; Junior, F.W.A. Volume estimation in a Eucalyptus plantation using multi-source remote sensing and digital terrain data: A case study in Minas Gerais State, Brazil. *Int. J. Remote Sens.* **2019**, *40*, 2683–2702. [\[CrossRef\]](#)
10. Tomppo, E.; Antropov, O.; Praks, J. Boreal Forest Snow Damage Mapping Using Multi-Temporal Sentinel-1 Data. *Remote Sens.* **2019**, *11*, 384. [\[CrossRef\]](#)
11. Torres, R.; Snoei, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.; Rommen, B.; Floury, N.; Brown, M.; et al. GMES Sentinel-1 mission. *Remote Sens. Environ.* **2012**, *120*, 9–24. [\[CrossRef\]](#)
12. Kurvonen, L.; Pulliainen, J.; Hallikainen, M. Retrieval of biomass in boreal forests from multitemporal ERS-1 and JERS-1 SAR images. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 198–205. [\[CrossRef\]](#)
13. 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. [\[CrossRef\]](#)
14. Pulliainen, J.T.; Kurvonen, L.; Hallikainen, M.T. Multitemporal behavior of L- and C-band SAR observations of boreal forests. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 927–937. [\[CrossRef\]](#)
15. Rauste, Y. Multi-temporal JERS SAR data in boreal forest biomass mapping. *Remote Sens. Environ.* **2005**, *97*, 263–275. [\[CrossRef\]](#)
16. Santoro, M.; Eriksson, L.; Askne, J.; Schmullius, C. Assessment of stand-wise stem volume retrieval in boreal forest from JERS-1 L-band SAR backscatter. *Int. J. Remote Sens.* **2006**, *27*, 3425–3454. [\[CrossRef\]](#)
17. 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. [\[CrossRef\]](#)
18. Antropov, O.; Rauste, Y.; Ahola, H.; Hame, T. Stand-Level Stem Volume of Boreal Forests From Spaceborne SAR Imagery at L-Band. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 35–44. [\[CrossRef\]](#)
19. Antropov, O.; Rauste, Y.; Häme, T.; Praks, J. Polarimetric ALOS PALSAR Time Series in Mapping Biomass of Boreal Forests. *Remote Sens.* **2017**, *9*, 999. [\[CrossRef\]](#)
20. Pulliainen, J.; Mikhela, P.; Hallikainen, M.; 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. [\[CrossRef\]](#)
21. Attema, E.P.W.; Ulaby, F.T. Vegetation modeled as a water cloud. *Radio Sci.* **1978**, *13*, 357–364. [\[CrossRef\]](#)
22. Santoro, M.; Cartus, O. Research Pathways of Forest Above-Ground Biomass Estimation Based on SAR Backscatter and Interferometric SAR Observations. *Remote Sens.* **2018**, *10*, 608. [\[CrossRef\]](#)
23. Laurin, G.V.; Balling, J.; Corona, P.; Mattioli, W.; Papale, D.; Puletti, N.; Rizzo, M.; Truckenbrodt, J.; Urban, M. Above-ground biomass prediction by Sentinel-1 multitemporal data in central Italy with integration of ALOS2 and Sentinel-2 data. *J. Appl. Remote Sens.* **2018**, *12*, 016008. [\[CrossRef\]](#)
24. Stelmaszczuk-Górska, M.A.; Urbazaev, 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. [\[CrossRef\]](#)
25. Neumann, M.; Saatchi, S.S. Polarimetric backscatter optimization for biophysical parameter estimation. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 254–258. [\[CrossRef\]](#)
26. 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.* **2011**, *5*, 298–306. [\[CrossRef\]](#)
27. Schlund, M.; Davidson, M.W.J. Aboveground Forest Biomass Estimation Combining L- and P-Band SAR Acquisitions. *Remote Sens.* **2018**, *10*, 1151. [\[CrossRef\]](#)
28. Schmullius, C.; Thiel, C.; Pathe, C.; Santoro, M. Radar Time Series for Land Cover and Forest Mapping. In *Remote Sensing Time Series: Revealing Land Surface Dynamics*; Kuenzer, C., Dech, S., Wagner, W., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 323–356.
29. Zhang, N.; Chen, M.; Yang, F.; Yang, C.; Yang, P.; Gao, Y.; Shang, Y.; Peng, D. Forest Height Mapping Using Feature Selection and Machine Learning by Integrating Multi-Source Satellite Data in Baoding City, North China. *Remote Sens.* **2022**, *14*, 4434. [\[CrossRef\]](#)
30. Morin, D.; Planells, M.; Baghdadi, N.; Bouvet, A.; Fayad, I.; Le Toan, T.; Mermoz, S.; Villard, L. Improving Heterogeneous Forest Height Maps by Integrating GEDI-Based Forest Height Information in a Multi-Sensor Mapping Process. *Remote Sens.* **2022**, *14*, 2079. [\[CrossRef\]](#)
31. Li, Y.; Li, C.; Li, M.; Liu, Z. Influence of variable selection and forest type on forest aboveground biomass estimation using machine learning algorithms. *Forests* **2019**, *10*, 1073. [\[CrossRef\]](#)
32. Ataee, M.S.; Maghsoudi, Y.; Latifi, H.; Fadaie, F. Improving estimation accuracy of growing stock by multi-frequency SAR and multi-spectral data over Iran's heterogeneously-structured broadleaf Hyrcanian forests. *Forests* **2019**, *10*, 641. [\[CrossRef\]](#)
33. Dostálová, A.; Wagner, W.; Milenković, M.; Hollaus, M. Annual seasonality in Sentinel-1 signal for forest mapping and forest type classification. *Int. J. Remote Sens.* **2018**, *39*, 7738–7760. [\[CrossRef\]](#)

34. Frison, P.L.; Fruneau, B.; Kmiha, S.; Soudani, K.; Dufrêne, E.; Le Toan, T.; Koleček, T.; Villard, L.; Mougin, E.; Rudant, J.P. Potential of Sentinel-1 Data for Monitoring Temperate Mixed Forest Phenology. *Remote Sens.* **2018**, *10*, 2049. [\[CrossRef\]](#)
35. Udali, A.; Lingua, E.; Persson, H.J. Assessing Forest Type and Tree Species Classification Using Sentinel-1 C-Band SAR Data in Southern Sweden. *Remote Sens.* **2021**, *13*, 3237. [\[CrossRef\]](#)
36. Soudani, K.; Delpierre, N.; Berveiller, D.; Hmimina, G.; Vincent, G.; Morfin, A.; Éric Dufrêne. Potential of C-band Synthetic Aperture Radar Sentinel-1 time-series for the monitoring of phenological cycles in a deciduous forest. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102505. [\[CrossRef\]](#)
37. Huang, X.; Ziniti, B.; Torbick, N.; Ducey, M.J. Assessment of Forest above Ground Biomass Estimation Using Multi-Temporal C-band Sentinel-1 and Polarimetric L-band PALSAR-2 Data. *Remote Sens.* **2018**, *10*, 1424. [\[CrossRef\]](#)
38. Tomppo, E.; Halme, M. Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach. *Remote Sens. Environ.* **2004**, *92*, 1–20. [\[CrossRef\]](#)
39. Tomppo, E.; Haakana, M.; Katila, M.; Peräsaari, J. *Multi-Source National Forest Inventory—Methods and Applications: Managing Forest Ecosystems*; Springer: Dordrecht, The Netherlands, 2008.
40. Tomppo, E.; Olsson, H.; Ståhl, G.; Nilsson, M.; Hagner, O.; Katila, M. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sens. Environ.* **2008**, *112*, 1982–1999. [\[CrossRef\]](#)
41. Korhonen, L.; Hadi; Packalen, P.; Rautiainen, M. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sens. Environ.* **2017**, *195*, 259–274. [\[CrossRef\]](#)
42. Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **2017**, *199*, 415–426. [\[CrossRef\]](#)
43. Tomppo, E.; Antropov, O.; Praks, J. Cropland Classification Using Sentinel-1 Time Series: Methodological Performance and Prediction Uncertainty Assessment. *Remote Sens.* **2019**, *11*, 2480. [\[CrossRef\]](#)
44. Cui, Z.; Kerekes, J.P. Potential of Red Edge Spectral Bands in Future Landsat Satellites on Agroecosystem Canopy Green Leaf Area Index Retrieval. *Remote Sens.* **2018**, *10*, 1458. [\[CrossRef\]](#)
45. Karjalainen, T.; Kellomäki, S. Greenhouse gas inventory for land use changes and forestry in Finland based on international guidelines. *Mitig. Adapt. Strat. Glob. Chang.* **1996**, *1*, 51–71. [\[CrossRef\]](#)
46. Tomppo, E. National forest inventory of Finland and its role estimating the carbon balance of forests. *Biotechnol. Agron. Soc. Environ.* **2000**, *4*, 281–284.
47. Small, D.; Miranda, N.; Zuberbühler, L.; Schubert, A.; Meier, E. Terrain-corrected Gamma: Improved thematic land-cover retrieval for SAR with robust radiometric terrain correction. In Proceedings of the ESA Living Planet Symposium, Bergen, Norway, 28 June–2 July 2010; pp. 1–8.
48. Finnish Forest Centre *Collection of Forest Resource Information*; Finnish Forest Centre: Kuopio, Finland, 2019.
49. Luoto, M. *Forest Inventory in Finland*; Finnish Forest Centre: Kuopio, Finland, 2020.
50. Dobson, M.; Ulaby, F.; LeToan, T.; Beaudoin, A.; Kasischke, E.; Christensen, N. Dependence of radar backscatter on coniferous forest biomass. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 412–415. [\[CrossRef\]](#)
51. Rauste, Y.; Hamee, T.; Pulliainen, J.; Heiska, K.; Hallikainen, M. Radar-based forest biomass estimation. *Int. J. Remote Sens.* **1994**, *15*, 2797–2808. [\[CrossRef\]](#)
52. Rignot, E.; Way, J.; Williams, C.; Viereck, L. Radar estimates of aboveground biomass in boreal forests of interior Alaska. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 1117–1124. [\[CrossRef\]](#)
53. Kasischke, E.; Christensen, N.; Bourgeau-Chavez, L. Correlating radar backscatter with components of biomass in loblolly pine forests. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 643–659. [\[CrossRef\]](#)
54. Tsui, O.W.; Coops, N.C.; Wulder, M.A.; Marshall, P.L.; McCardle, A. Using multi-frequency radar and discrete-return LiDAR measurements to estimate above-ground biomass and biomass components in a coastal temperate forest. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 121–133. [\[CrossRef\]](#)
55. Hame, T.; Rauste, Y.; Antropov, O.; Ahola, H.A.; Kilpi, J. Improved Mapping of Tropical Forests with Optical and SAR Imagery, Part II: Above Ground Biomass Estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 92–101. [\[CrossRef\]](#)
56. 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. [\[CrossRef\]](#)
57. Baskerville, C.L. Use of Logarithmic Regression in the Estimation of Plant Biomass. *Can. J. For. Res.* **1972**, *2*, 49–53. [\[CrossRef\]](#)
58. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [\[CrossRef\]](#)
59. Vapnik, V.; Golowich, S.; Smola, A. Support vector method for function approximation, regression estimation, and signal processing. In Proceedings of the Advances in Neural Information Processing Systems, Denver, CO, USA, 2–5 December 1997; Mozer, M., Jordan, M., Petsche, T., Eds.; MIT Press: Cambridge, MA, USA, 1997; pp. 281–287.
60. Gunn, S.R. Support vector machines for classification and regression. *ISIS Tech. Rep.* **1998**, *14*, 5–16.
61. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [\[CrossRef\]](#)
62. Neumann, M.; Saatchi, S.S.; Ulander, L.M.; Fransson, J.E. Assessing performance of L- and P-band polarimetric interferometric SAR data in estimating boreal forest above-ground biomass. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 714–726. [\[CrossRef\]](#)
63. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
64. Karlson, M.; Ostwald, M.; Reese, H.; Sanou, J.; Tankoano, B.; Mattsson, E. Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and random forest. *Remote Sens.* **2015**, *7*, 10017–10041. [\[CrossRef\]](#)

65. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
66. Esteban, J.; McRoberts, R.E.; Fernández-Landa, A.; Tomé, J.L.; Naesset, E. Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. *Remote Sens.* **2019**, *11*, 1944. [[CrossRef](#)]
67. Hethcoat, M.G.; Carreiras, J.M.; Edwards, D.P.; Bryant, R.G.; Quegan, S. Detecting tropical selective logging with C-band SAR data may require a time series approach. *Remote Sens. Environ.* **2021**, *259*, 112411. [[CrossRef](#)]
68. Gleason, C.J.; Im, J. Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sens. Environ.* **2012**, *125*, 80–91. [[CrossRef](#)]
69. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [[CrossRef](#)]
70. Kinney, J.B.; Atwal, G.S. Equitability, mutual information, and the maximal information coefficient. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3354–3359. [[CrossRef](#)] [[PubMed](#)]
71. Belghazi, M.I.; Baratin, A.; Rajeswar, S.; Ozair, S.; Bengio, Y.; Courville, A.; Hjelm, R.D. Mine: Mutual information neural estimation. *arXiv* **2018**, arXiv:1801.04062.
72. Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating mutual information. *Phys. Rev. E* **2004**, *69*, 066138. [[CrossRef](#)]
73. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **1996**, *58*, 267–288. [[CrossRef](#)]
74. 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. [[CrossRef](#)]
75. Santoro, M.; Cartus, O.; Fransson, J.E. 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](#)]
76. 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](#)]
77. Wang, H.; Ouchi, K.; Watanabe, M.; Shimada, M.; Tadono, T.; Rosenqvist, A.; Romshoo, S.; Matsuoka, M.; Moriyama, T.; Uratsuka, S. In Search of the Statistical Properties of High-Resolution Polarimetric SAR Data for the Measurements of Forest Biomass Beyond the RCS Saturation Limits. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 495–499. [[CrossRef](#)]

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.