Article

Mapping Forest Growing Stock and Its Current Annual Increment Using Random Forest and Remote Sensing Data in Northeast Italy

Luca Cadez [1,2,]*, Antonio Tomao [1], Francesca Giannetti [3], Gherardo Chirici [3] and Giorgio Alberti [1]

1   Department of Agricultural, Food, Environmental and Animal Sciences, University of Udine,
    33100 Udine, Italy; antonio.tomao@uniud.it (A.T.); giorgio.alberti@uniud.it (G.A.)
2   Department of Life Sciences, University of Trieste, 34127 Trieste, Italy
3   Department of Agricultural, Food, Environmental and Forestry Sciences and Technologies, University of
    Florence, 50145 Firenze, Italy; gherardo.chirici@unifi.it (G.C.)
*   Correspondence: luca.cadez@uniud.it

Abstract: The role of forests in providing multiple goods and services has been recognized worldwide. In such a context, reliable spatial predictions of forest attributes such as tree volume and current increment are fundamental for conducting forest monitoring, improving restoration programs, and supporting decision-making processes. This article presents the methodology and the results of the wall-to-wall spatialization of the growing stock volume and the current annual increment measured in 273 plots of data of the Italian National Forest Inventory over an area of more than 3260 km^2 in the Friuli Venezia Giulia region (Northeast Italy). To this aim, a random forest model was tested using as predictors 4 spectral indices from Sentinel-2, a high-resolution Canopy Height Model derived from LiDAR, and geo-morphological data. According to the Leave One Out cross-validation procedure, the model for the growing stock shows an R^2 and an RMSE% of 0.67 and 41%, respectively. Instead, an R^2 of 0.47 and an RMSE% of 57% were obtained for the current annual increment. The validation with an independent dataset further improved the models' performances, yielding significantly higher R^2 values of 0.84 and 0.83 for volume and for increment, respectively. Our results underline a relatively higher importance of LiDAR-derived metrics compared to other covariates in estimating both attributes, as they were even twice as important as vegetation indices for growing stock. Therefore, these metrics are promising for the development of a national LiDAR-based model.

Keywords: biomass estimation; area-based approach; LiDAR; random forest; remote sensing

1. Introduction

The role of forests in providing multiple goods and services has been recognized at global, European, and national levels in many countries [1–3]. Many of these documents underline the multifunctionality of forests, the provision of energy, raw materials (wood, resin, medicinal plants, etc.), and non-wood forest products such as wild food (berries, mushrooms, honey, etc.). They also play crucial roles in regulating local and global climate, enhancing soil retention and water quality, facilitating pollination, acting as barriers to natural hazards, promoting biodiversity, and offering recreational and aesthetic values in rural and peri-urban landscapes [4,5]. The provision of updated, reliable, and spatially explicit data on forest resources on a large scale is an essential element for efficient and timely forest planning and management. In fact, it helps decision-makers to allocate resources more effectively, whether for harvesting or for biodiversity conservation, based on the latest information about forest resources. Spatially explicit data also aid in assessing risks such as forest fires, pests, diseases, and illegal activities, thus allowing proactive measures to mitigate them. Typically, information on forest resources is obtained through National Forest Inventories (NFIs). Data collected in NFI field sample plots provide

aggregated estimates of forest attributes, and their related uncertainties, by region or country and, eventually, by forest category. However, detailed spatial layers reporting the variation in the most important forest characteristics are available only at a local scale (i.e., forest property), while are lacking at a larger scale (i.e., region or country), thereby limiting their potential use in forest management and planning.

In the last decades, different methods have been used to estimate and map forest attributes (e.g., aboveground biomass, growing stock volume, basal area, etc.), and several of these rely on remote sensing (RS) and the application of Artificial Intelligence (AI) algorithms [6]. Some countries have already adopted a hybrid methodology to estimate forest attributes in the so-called Enhanced Forest Inventories (EFIs) [7–9], which combine data from different sources, like passive (optical), or active (i.e., Synthetic Aperture Radar or Light Detection and Raging—LiDAR) sensors, and from many platforms (e.g., satellite, plane, drone) [10]. Such inventories usually produce wall-to-wall spatial estimations of forest attributes such as forest volume and height, providing detailed information aimed at improving forest management and planning in the framework of the so-called precision forestry [11]. For example, optical RS (i.e., Landsat and Sentinel-2) have been largely used for forest biomass estimation over large areas [12], but they are often characterized by signal saturation problems [13] (i.e., when forest aboveground biomass reaches a certain threshold, RS imagery cannot effectively reflect the spectral change, resulting in an underestimation of forest biomass). Moreover, these applications need seamless data that are currently limited in spatial and temporal coverage due to their high costs and complexity. In particular, many European countries, including Italy, still lack comprehensive coverage of large-scale LiDAR surveys [14–16].

In terms of used approaches, the spatially explicit estimation of forest attributes follows two alternative ways: Individual Tree Detection (IDT) and area-based approach (ABA). IDT is mainly oriented at identifying the location of each single tree and the most used sources are optical RS and LiDAR data [17,18]. This approach, although very accurate, is not easy to apply in broadleaved or mixed forests [19] and could be computationally expensive due to the use of very high-resolution data, especially when applied at a large spatial scale [20]. Conversely, ABA estimates a predictive model that links a set of covariates available on the entire area (e.g., tree height, spectral indexes) to a target inventory attribute measured at selected ground plots [7,21,22]. In this case, data by global scale observation systems like Sentinel and Landsat are widely applied in the estimation of several biophysical variables [23] such as Vegetation Indices (VIs), which have been demonstrated to be effective for assessing vegetation cover, vigor, and growth dynamics [24–26].

In addition, LiDAR has often become an additional source to be included in ABA for predicting a range of forest attributes, including tree height, tree density, growing stock volume (GSV), forest stand biomass, etc. The spatialization using ABAs could follow several methods, including both parametric (multiple linear regression, geographically weighted regression) and non-parametric approaches (K-nearest neighbors, random forest) [27,28]. In particular, Random forest (RF) is a decision tree algorithm introduced by Breiman [29], well documented in the forestry sector for classifying and predicting forest attributes [30].

Specifically, in Italy, wall-to-wall mapping of GSV and current annual volume increment (CAI) at a large scale has been conducted so far using satellite data and some auxiliary features (e.g., climate and geomorphologic variables) [31,32]. In the case of CAI, the authors used a specific process-based model that simulates the storage and fluxes of water, carbon, and nitrogen within terrestrial ecosystems. While some experiences are available for local case studies [33,34], models to predict GSV and especially CAI at a large scale using multiple data sources which also include Airborne Laser Scanning (ALS) are still lacking for Italian forests.

This paper aims to fill the above-mentioned knowledge gaps by demonstrating the feasibility of mapping GSV and CAI at a large scale by combining Sentinel-2 spectral indexes, the Canopy Height Model (CHM) from ALS, and geomorphological variables. For this purpose, a specific model was developed for the Friuli Venezia Giulia Region (Italy),

where LiDAR data were available. The model was calibrated using NFI plots, and the results were validated with an independent dataset derived from forest management plans.

2. Materials

2.1. Study Area

The study area is located in Friuli Venezia Giulia (7924 km^2), in Italy. The region is considered a representative case study in Italy because it includes three different biogeographical regions (i.e., alpine, continental, and Mediterranean). Approximately 41% of Friuli Venezia Giulia is covered by forests (Figure 1). These forests have expanded over the last 70 years, particularly in marginal areas, reflecting a trend that is quite prevalent in Italy [35]. Since 1861, the forested area has doubled due to depopulation and the subsequent abandonment of traditional agro-forestry practices especially in the hilly and mountainous areas [36]. In contrast, the presence of forests in the plain is quite low, mainly due to agricultural intensification and urbanization. The most common forest categories in the region are beech (*Fagus sylvatica* L., *1753*) pure stands, black pine (*Pinus nigra J.F.Arnold*, *1785*) and Scots pine (*Pinus sylvestris* L., *1753*) mixed stands, hornbeam (*Carpinus betulus* L., *1753*) and hophornbeam (*Ostrya carpinifolia Scop.*) mixed stands, and Norway spruce (*Picea abies* (L.) *H.Karst., 1881*) pure stands. The climate ranges from the Mediterranean along the coastal areas, to a humid temperate climate in the plains and hilly areas, up to the alpine climate of the mountains.
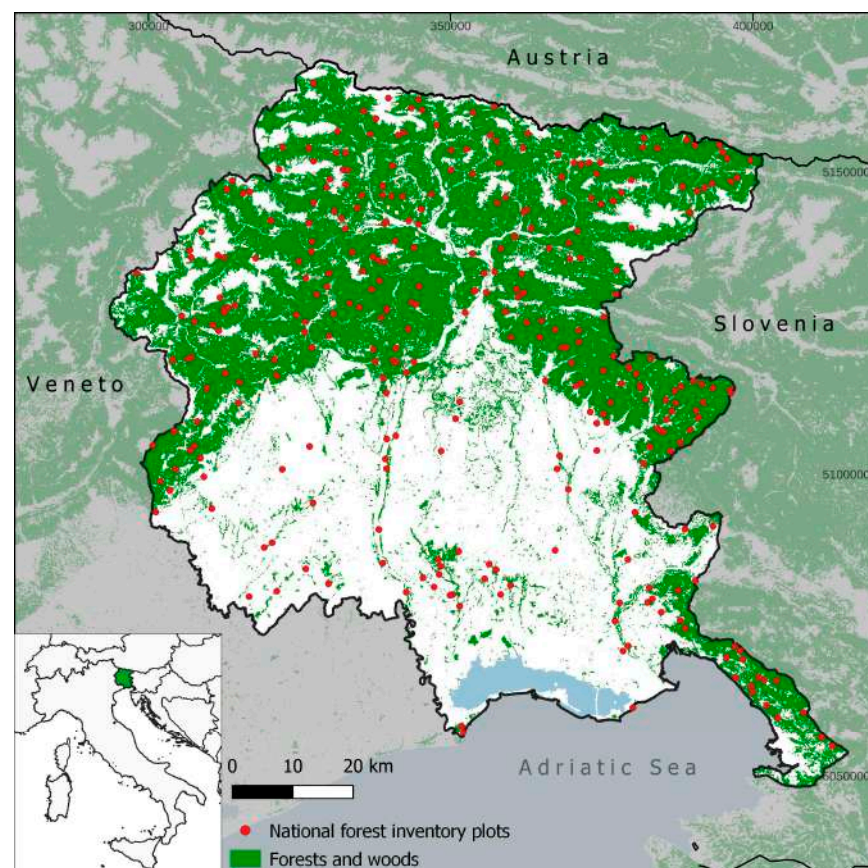
Figure 1. The Friuli Venezia region with the forest areas and the National Forest Inventory plots. Reference system: WGS84.

2.2. Data

2.2.1. Third Italian National Forest Inventory

The field reference data for this study were obtained from the third Italian NFI settled in 2015, which is based on a three-phase, non-aligned, systematic sampling design, with

data collected between 2018 and 2019 [37]. This NFI consists of three phases: the first regards classification of land use and land cover through the photointerpretation of high-resolution orthophotos following a 1 km × 1 km grid overlapping the whole county (over 301,000 points). The second phase considers a subsample of the first phase points to determine qualitative forest characteristics (i.e., composition, forest category, management, etc.) through visits to the field. In the third phase, strata are identified considering the forest category assigned in the second phase together with the land use and cover class, and the region. For each forest stratum, a subsample of points is extracted for field measurements. Dendrometric data are collected in two concentric plots (4 and 13 m radius) depending on tree diameter at 1.30 m (diameter at breast height = DBH) and used to derive forest attributes such as basal area (m^2 ha^{-1}), GSV (m^3 ha^{-1}), and carbon stock (Mg C ha^{-1}). In the NFI, CAI (m^3 ha^{-1} y^{-1}) is estimated by collecting increment cores by a subsample of trees per plot using specific allometric equations [37]. It refers to the mean value of tree growth in the last five years.

For this study, we selected all the 279 third phase's plots available for the Friuli Venezia Giulia region. After a first reliability check, six plots affected by the 2018 Vaia storm were excluded since NFI plots were surveyed before forest cover loss and RS data acquisition. Thus, the final calibration dataset comprised a total of 273 plots distributed across the entire region, as shown in Figure 1.

2.2.2. Digital Elevation Models and Derivatives

ALS data used in this study were collected between 2017 and 2019, at an average altitude of 500 m above ground level, by the Autonomous Region Friuli Venezia Giulia (https://irdat.regione.fvg.it/consultatore-dati-ambientali-territoriali/detail/irdat/dataset/11647, accessed on 29 January 2022). Some areas were also re-surveyed in 2020 after the Vaia storm in order to update the point cloud. Several flights in different years were needed to achieve wall-to-wall coverage of the entire regional area. The resulting average point density is 16 points m^{-2}, with a lower density of 10 points m^{-2} collected above the 1000 m a.s.l., with a vertical accuracy of ±15 cm. From those data, a high-resolution Digital Terrain Model (DTM) was built, and, after resampling at 10 m, the elevation, slope, and aspect were derived as geomorphological predictors to be used in the spatialization model. Moreover, a high-resolution CHM was derived as a difference between the Digital Surface Model (DSM) first pulse and the DTM, masked for the forest areas defined by the regional map of forest categories. Only some tiles in the southeastern part of the region (Karst area) were excluded due to the problems in the original LiDAR data. The importance of height models for predicting GSV and other forest attributes—also merging them with other sources like RS data—has been proven to be useful for a variety of applications in the literature, both at individual tree and area levels [38].

2.2.3. Sentinel-2 Data

In this study, four VIs commonly used in forestry research and known to be correlated to forest attributes and physiological processes [39] were selected as covariates (Table 1). Such VIs were derived by Sentinel 2 data using Google Earth Engine [40] from June, July, and August (i.e., growing season) within the 2019–2021 timeframe, starting from Level-2A, filtering for cloud presence and assembled in a median dataset (Figure 2). The considered period corresponds to the sampling period of the third Italian NFI. 2017 and 2018 were not included to avoid biases due to the loss in forest cover due to the Vaia storm in part of the study area [41], while 2021 was included to fill some small gaps due to cloud coverage and to have more data.

Table 1. Synthesis of predictor used for the estimations.

Type	Variable	Year Reference	Resolution (m)
Digital Elevation Models and Derivatives from Lidar	Elevation [median]	Original LiDAR collected between 2017 and 2020	10
	Slope [median]		10
	Aspect [median]		10
	CHM [mean, median]		0.5
SENTINEL-2 DATA	NDVI, NDII, GNDVI, EVI	2019–2021 June, July, August [median dataset]	10
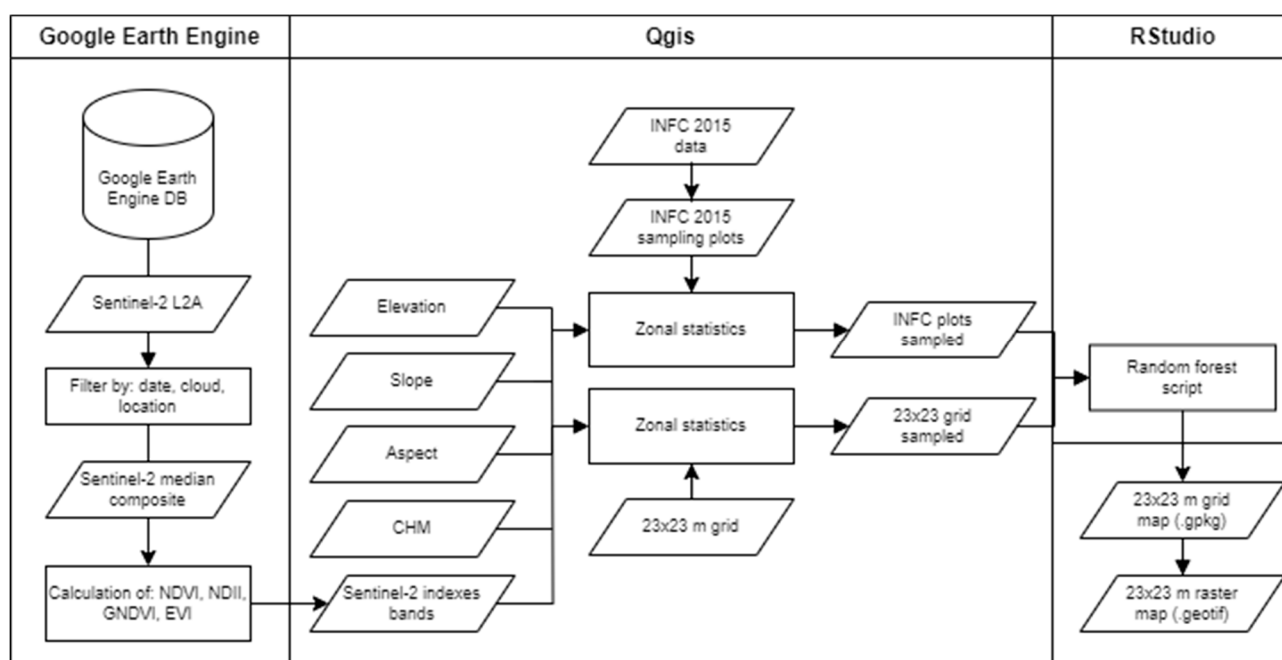
Figure 2. Flowchart of data processing with the details of used software.

The Normalized Difference Vegetation Index (NDVI) is the most widespread spectral index used in the biological field as it has been used in commercial agriculture and land-use studies to estimate different vegetation properties, including leaf area index (LAI), aboveground biomass, and plant productivity [42,43].

The Normalized Difference Infrared Index (NDII) is a reflectance measurement, sensitive to changes in water content of tree canopies. The index values increase with increasing water content. Applications of this index include agricultural crop management, forest canopy monitoring, and detection of stressed vegetation [44].

The Green Normalized Difference Vegetation Index (GNDVI) is a modified version of NDVI, more sensitive to variation in chlorophyll content than NDVI [45].

The Enhanced Vegetation Index (EVI) is like NDVI and can be used to quantify vegetation greenness and corrects for some atmospheric conditions and canopy background noise and is more sensitive in areas with dense vegetation [46]. EVI has been shown to be directly and positively associated with Gross Primary Production [47,48].

3. Methods

3.1. Random Forest Application

GSV and CAI were estimated at a regional scale following an ABA approach (Figure 2) based on an RF algorithm trained using the variables listed in Table 1, Figure 2, following Chirici et al. [31]. RF generates a set of regression trees that are aggregated across nodes to produce predictions without overfitting the data. For the present application, we considered

1000 trees to balance the calculation time with the correctness of the results [49]. To build the whole set of trees, RF uses a subset of predictors randomly chosen for each node, and each tree is built independently on samples randomly extracted from a portion of the reference set (usually 2/3 of the data), while the remaining third of the sample is randomly excluded (out-of-bag samples procedure, OOB). OOB allows an estimation of the error as well as the calculation of the relative importance of each covariate in the estimation of the attribute of interest by calculating the percentage increase in the mean squared error [29]. The predictors that produce the most accurate results are then selected at each node. According to White et al. [50], the variables that are used as predictors in the ABA exist as rasters or tessellations, with each grid cell typically representing a forest area similar in size to a standard ground plot. Thus, for this research, the RF model was estimated using a first dataset consisting of NFI data at the plot level (GSV and CAI) as dependent variables and the values of the selected covariates (Table 1) calculated for each NFI inventory plot (530 m^2) using the "zonal statistics" tool in QGIS (Figure 2). Then, the model was applied to a second dataset consisting of a 23 m × 23 m grid with the same surface of NFI plots and covering all the forest areas. Each cell contains the median value of all the covariates, and also the *mean* in the case of the CHM. Data have been exported in GeoPackage format. The RF model was first estimated using only the first dataset applying all the predictors. Then, the model was applied to the second dataset, obtaining the spatialization. Finally, the estimated attributes of interest were rasterized for the validation. All analysis were performed using the randomForest [51] within the R software version 4.2.0, and the pseudo-code is reported in the Appendix A.

3.2. Validation

A twofold validation was conducted to estimate the goodness of the approach. First, R^2, RMSE, and RMSE% were estimated, applying the leave-one-out (LOO) cross-validation technique within the RStudio code. In this method, the model was trained 273 times on all the datasets except for one point, and a prediction was made for that point. Then, the average error was computed and used to evaluate the model. Each row in the reference set was excluded in sequence to train the model and then predicted from the model established using the partial reference set. Secondly, we also performed a validation using an independent dataset, selecting 26 forest management units (mean surface = 13 ha), where a complete tree survey was performed in a period comparable to that of the LiDAR data acquisition. In these units, mean estimated GSV and CAI were compared with actual values measured in the field.

4. Results

Estimated GSV ranges between 17 and 919 m^3 ha^{-1}, while CAI shows a range from 0.46 to 15.3 m^3 ha^{-1} y^{-1}. At a regional level, the total GSV is 76,312,755 m^3 (mean = 192 m^3 ha^{-1}), while the mean CAI amounts to 1,587,265 m^3 y^{-1} (mean = 4.6 m^3 ha^{-1} y^{-1}). Figure 3 shows the distribution of GSV and CAI at a regional scale with a spatial resolution of 23 m × 23 m. The different values of both attributes are related to the heterogeneity in species composition (forest categories) and environmental conditions. The highest GSV are mostly concentrated in the mesalpic and endalpic zones on more gentle slopes where tree diversity is limited to 1–3 species (Norway spruce, fir, and partially beech). The esalpic area (Prealps) shows lower values, especially where beech and black pines are dominant and mainly on steep areas. The hilly areas and the Karst show quite low GSV values. The forests in those areas are dominated by temperate oaks, chestnut, hornbeam, hophornbeam, and some other deciduous broadleaves. CAI is higher in the endalpic area like GSV but also in the esalpic area where most of the forests have been established on fertile ex-agricultural land following secondary successions and are characterized by fast-growing species such as *Robinia pseudoacaica* L., *Acer* spp., and *Fraxinux* spp.
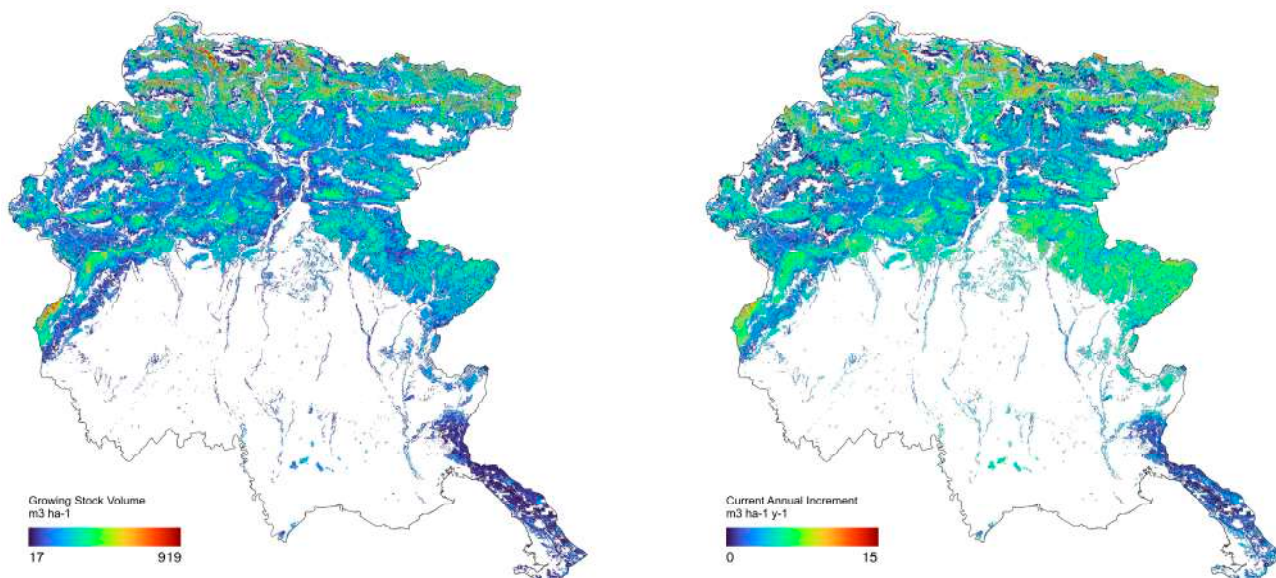
Figure 3. Growing stock volume (**left**; m^3 ha^{-1}) and current annual increment (**right**; m^3 ha^{-1} y^{-1}).

The relative importance of the considered covariates was assessed through an importance plot (Figure 4), where the code estimates the importance for each variable if removed from the model in terms of increment of MSE%. CHM-derived features are the most influential covariates for both the attributes of interest. Among the VIs, differences were detected between GSV and CAI. Indeed, GNDVI is the most important index for GSV estimation, while NDII is the most relevant for CAI. In both cases, NDVI does not result in the best predictors. With regards to geomorphology, while slope and elevation show a certain importance in estimating CAI and GSV, respectively, the aspect has a very low influence in both cases.
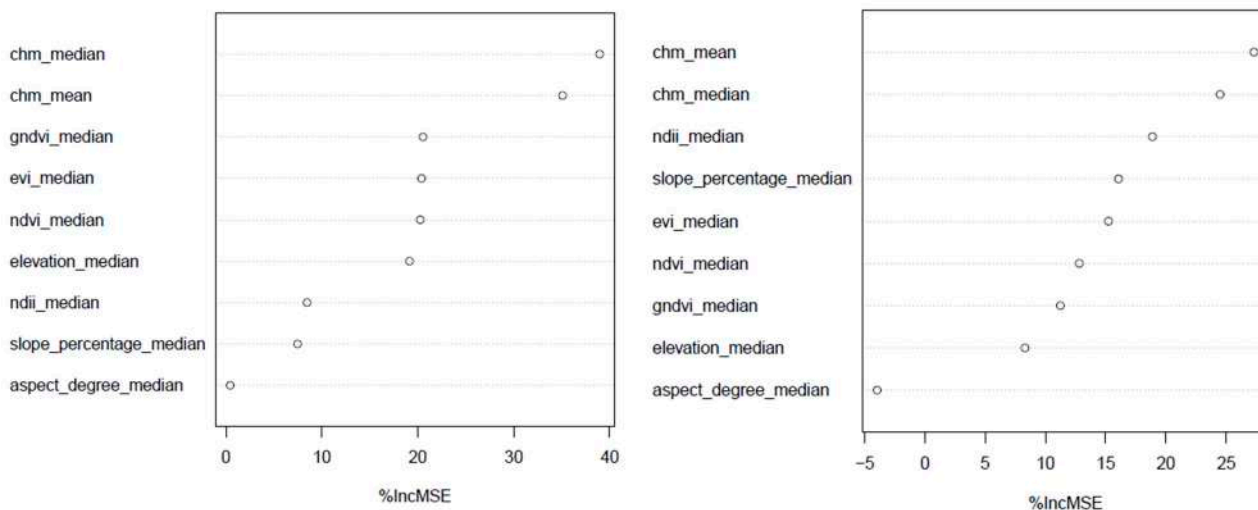

Figure 4. Importance plots for growing stock volume (**left**) and current annual increment (**right**).

The accuracy of our model in estimating GSV and CAI according to LOO is reported in Table 2 and Figure 5. The goodness of fit was higher for GSV than for CAI, with an R^2 ranging from 0.67 to 0.46, respectively. In both cases, the RMSE% was lower than 60%. The validation using the 26 forest management plans is reported in Figure 6. In this case, the R^2 shows values of 0.84 and 0.83 for GSV and CAI, respectively (Table 2, Figure 6).

Table 2. Error estimation for the growing stock volume (GSV) and current annual increment (CAI) according to the leave-one-out (LOO) procedure and the comparison with the forest management units (FMUs).

Attribute	LOO			FMU		
Attribute	R^2	RMSE	RMSE%	R2	RMSE	RMSE%
GSV	0.67	96	40	0.84	118	64
CAI	0.46	3	57	0.83	2	14
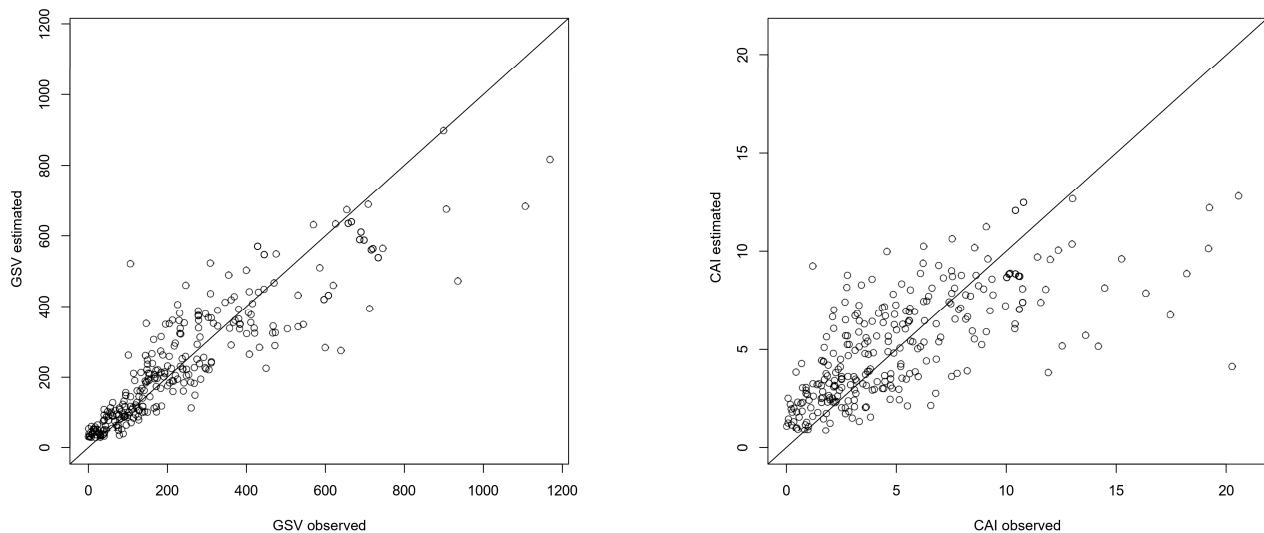
Figure 5. Scatterplots comparing estimated and observed values according to the leave-one-out (LOO) procedure (**left**: growing stock volume, GSV; **right**: current annual increment, CAI). The black line is the 1:1 line.
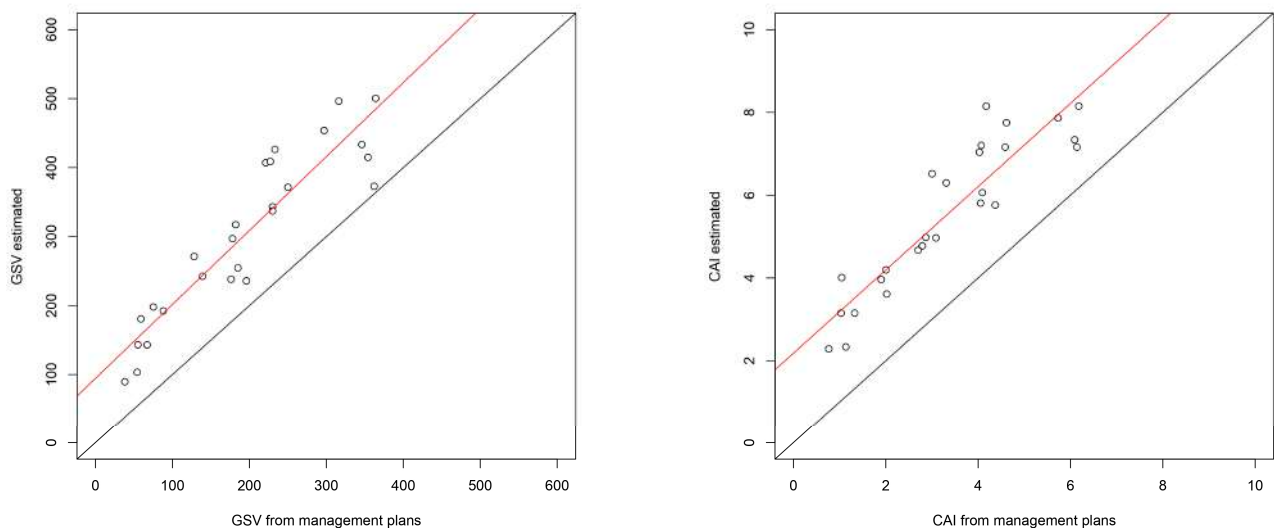
Figure 6. Scatterplots of estimated and observed (forest management units where a complete tree census was performed) values of growing stock volume (GSV) and current annual increment (CAI). The regression line is reported in red. The black line is the 1:1 line.

5. Discussion

Detailed maps on the distribution of GSV and CAI at the regional or national scale are fundamental for monitoring forest ecosystems, supporting policymakers and forest planning, as well as for the evaluation of ecosystem services (wood production, climate

regulation, soil protection). In fact, these maps are not only important for producing statistics at different spatial scales, but can also be integrated into specific forest decision support systems [52], thus supporting evidence-based strategies and decisions at a regional scale [53]. Another potential application is using them as inputs for models designed to simulate the behavior and/or probability of disturbances such as fires [54,55]. For these reasons, the main objective of the study was to develop a spatial approach to jointly obtain high-resolution GSV and CAI maps over an area of more than 3260 km^2. The proposed approach allowed, through the spatialization of the values measured in the third Italian NFI, a spatially explicit estimation of the attributes of interest with a good level of precision. When validation according to the LOO method is considered, we obtained a better accuracy for GSV than for CAI (R^2 0.67 vs. 0.46), showing how spatially explicit estimation of stand growth is still challenging. LOO has been often used for the accuracy assessment of model outputs [14], but, as it is sensitive to outliers, it may not be representative of the model's performance, and the estimation accuracy may be overoptimistic. For this reason, we validated our model results also using data from an independent dataset (i.e., forest management units). In such a case, an even better performance of the model in the estimation of the attributes of interest when averaged over forest management units' surfaces was obtained. However, despite the good agreement between estimated and measured values (slope of the regression lines around 1), a significant intercept of 94.5 m^3 ha^{-1} and 2.2 m^3 ha^{-1} y^{-1} for GSV and CAI was detected, respectively (Figure 6). This is probably related to the different sampling approaches and DBH thresholds used in NFI and forest management plans. In fact, the NFI surveys all trees with a DBH greater than 4.5 cm, while forest management plans consider only trees with a DBH greater than 20 cm. Moreover, standing snags are often not measured by forest management plans. For this reason, the NFI-based estimates show systematically greater GSV and CAI. Differences between our estimates based on NFI data and ground truth from forest management plans can also derive from differences in the used allometric equations: while NFI uses general equations for the whole country, forest management plans quite often use local and species-specific equations. Earlier studies on the spatialization of GSV conducted in Friuli Venezia Giulia using VIs only obtained a lower R^2 (0.57) when comparing estimates with ground truth forest management plan data [56], showing how combining different data sources for modeling of forest attributes (our approach) can produce much more accurate estimations of GSV [57,58]. In another recent research at the national scale, Chirici et al. [26] achieved an R^2 of 0.47 by applying the RF algorithm to RS covariates (without ALS) and an RMSE% equal to 68.70. Puliti et al. [59], including ALS among covariates, reached an accuracy similar to ours when considering management units as ground truth (R^2 = 0.83), but these authors used ALS data with a lower density than ours (7.45 points m^{-2}), and the studied area was rather small and quite homogeneous in terms of stands and species composition. Another recent study compared linear regression, K-nearest neighbors, and random forest methods for predicting GSV from ALS data, and reported that RF is the best predictor for boreal forest in Norway and Finland (R^2 of 0.91 and 0.96, respectively) even with a lower LiDAR point (0.6–4.8 points m^{-2}) [60].

Spatialization of the CAI is more challenging than GSV. Previous papers have reported very different accuracies depending on the working scale, the species composition, and the structural heterogeneity. In fact, when forest stands are more homogenous and the investigated area is small, the comparison between modeled values and field-measured ones is generally strong. For example, Tian et al. [49] reported R^2 values up to 0.65 and 0.66 in natural *Larix* and *Quercus* forests. Other authors reported values closer to the ones we estimated following the LOO procedure (i.e., [61]). When CAI estimates are compared with the ones measured in the forest management units (Table 2, Figure 6), the correlation coefficient is higher and closer to the values reported by Tian et al. [49]. Stratification for forest categories could eventually improve our CAI estimation, but this would require a larger number of well-distributed NFI plots to have enough data to train the models. On the other hand, potential improvements in our approach could come from the use of other

ALS-derived products, such as previous surveys to calculate the DSM of Differences (DoDs), or texture metrics, like Gray Level Co-Occurrence Matrix metrics (GLCMs) [62], as they can help to better characterize the spatial properties and variability of the vegetation. Previous studies have obtained different conclusions on the combined use of LiDAR and optical sensor data, depending on the data sources, applied algorithms, and the characteristics of the area under investigation (e.g., [57,63,64]). Our results show that GSV and CAI, at least in our study area, are mostly driven by CHM rather than other optical and geo-morphological information (Figure 4). Especially in the case of VIs, this could be explained by the saturation effect in dense forests, which can reduce the efficiency of these indices in predicting GSV and CAI [13]. Several other studies have underlined the importance of having ALS-derived data for a most reliable estimation of forest attributes [7,14,17]. However, few regions are actually seamlessly covered by ALS, especially in Italy [15]. A potential step forward would be to test different configurations of predictors for GSV and CAI, with or without LiDAR-derived metrics, in order to better understand to what extent the use of LiDAR data can improve the prediction of these biophysical attributes over a large area. Additionally, our approach could be also tested to assess the GSV and CAI of forested areas smaller than 2000 m^2 and not included in the NFI definition of "forest", such as hedgerows or trees outside the forests.

6. Conclusions

A reliable estimation of the spatial distribution of forest biomass and growth over large areas is fundamental to not only support forestry practice and policy, but also to estimate some other relevant ecosystem services (for example, carbon stock and flux). In this study, we modeled GSV and CAI over more than 3260 km^2 by combining LiDAR-derived metrics, optical imagery, and geo-morphological data through an RF algorithm. Our results emphasized the relatively higher importance of LiDAR-derived information than other covariates to estimate both GSV and CAI. Possible marginal improvements in the estimates could be derived by integrating other data sources into the models (e.g., other LiDAR-derived products) or by stratifying by forest category before modeling, but such an approach is still limited by the relative un-even distribution of NFI plots across forest categories at both the regional and national scale. However, our results strongly confirm the robustness of using ALS data for GSV estimation at a large scale and are promising for the development of a national LiDAR-based model as soon as data coverage is available countrywide. Even if we have found a good fit between modeled and measured CAIs, a future challenge will be a deeper integration of LiDAR data and other RS data sources with physical-based model and machine learning methods to further improve forest growth estimation.

Inventory of forests and forest carbon reservoirs released under the Creative Commons Attribution 4.0 license (www.inventarioforestale.org). Cartographic data from Autonomous Region Friuli Venezia Giulia were released under the Italian Open Data License v2.0 (IODL 2.0). Contains modified Copernicus Sentinel data for 2019, 2020, 2021.

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

Appendix A

This pseudocode used in RStudio explains the random forest algorithm here applied.
for each target variable (i.e. GSV, CAI)
   # Definition of the equation formula to be used by the model
   eq <- as.formula(dependant variable ~ covariate1 + covariate 2 + . . . + covariate *n*)

   # Estimation of the RF model, activating the option to calculate partial dependant plots
   RF_model <- randomForest(eq, data = NFI data, ntree = 1000, importance = T)

Spatialization of the results
estimating the dependent variable for the entire area (i.e. at each cell of the grid 23 × 23 m covering all the forest area)
predict (RF_model, grid)

   # LOO evaluation
for each raw in the NFI database

1. Training the model, using all the data except the raw we are evaluating using the same equation as above

LOO_rf <- randomForest(eq, data = NFI data except the raw, ntree = 1000)

2. Predicting the value of the dependent variable using the model trained on the subsample n-1 for that specific raw that is excluded from the training set

predict(LOO_rf, raw)
end for

Calculating the parameters RMSE, RMSE$_{\%}$ and R^2 confronting the predicted values of the dependent variable using the model trained on the subsample which excludes that raw and the actual value of the target
end for

References

1. United Nations United Nations Strategic Plan for Forests 2017–2030. 2017. Available online: https://www.un.org/esa/forests/wp-content/uploads/2016/12/UNSPF_AdvUnedited.pdf (accessed on 5 May 2023).
2. Ministero delle Politiche Agricole e Forestali. *Strategia Forestale Nazionale [National Forestry Strategy]*; Ministero delle Politiche Agricole e Forestali: Roma, Italy, 2021.
3. European Commission New EU Forest Strategy for 2030 COM(2021) 572 Final 2021. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0572 (accessed on 4 June 2024).
4. Grammatikopoulou, I.; Vačkářová, D. The Value of Forest Ecosystem Services: A Meta-Analysis at the European Scale and Application to National Ecosystem Accounting. *Ecosyst. Serv.* **2021**, *48*, 101262. [CrossRef]
5. Aznar-Sánchez, J.A.; Belmonte-Ureña, L.J.; López-Serrano, M.J.; Velasco-Muñoz, J.F. Forest Ecosystem Services: An Analysis of Worldwide Research. *Forests* **2018**, *9*, 453. [CrossRef]
6. Liu, Z.; Peng, C.; Work, T.; Candau, J.-N.; DesRochers, A.; Kneeshaw, D. Application of Machine-Learning Methods in Forest Ecology: Recent Progress and Future Challenges. *Environ. Rev.* **2018**, *26*, 339–350. [CrossRef]
7. White, J.C.; Coops, N.C.; Wulder, M.A.; Vastaranta, M.; Hilker, T.; Tompalski, P. Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Can. J. Remote Sens.* **2016**, *42*, 619–641. [CrossRef]
8. Stinson, G.; White, J.C. What's the Difference between EFI and NFI? Demystifying Current Acronyms in Forest Inventory in Canada. *For. Prof.* **2018**.

9. Gschwantner, T.; Alberdi, I.; Bauwens, S.; Bender, S.; Borota, D.; Bosela, M.; Bouriaud, O.; Breidenbach, J.; Donis, J.; Fischer, C.; et al. Growing Stock Monitoring by European National Forest Inventories: Historical Origins, Current Methods and Harmonisation. *For. Ecol. Manag.* **2022**, *505*, 119868. [CrossRef]

10. Toth, C.; Jóźków, G. Remote Sensing Platforms and Sensors: A Survey. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 22–36. [CrossRef]

11. Fardusi, M.J.; Chianucci, F.; Barbati, A. Concept to Practice of Geospatial-Information Tools to Assist Forest Management and Planning under Precision Forestry Framework: A Review. *Ann. Silvic. Res.* **2017**, *41*, 3–14. [CrossRef]

12. Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A Survey of Remote Sensing-Based Aboveground Biomass Estimation Methods in Forest Ecosystems. *Int. J. Digit. Earth* **2016**, *9*, 63–105. [CrossRef]

13. Mutanga, O.; Masenyama, A.; Sibanda, M. Spectral Saturation in the Remote Sensing of High-Density Vegetation Traits: A Systematic Review of Progress, Challenges, and Prospects. *ISPRS J. Photogramm. Remote Sens.* **2023**, *198*, 297–309. [CrossRef]

14. Coops, N.C.; Tompalski, P.; Goodbody, T.R.H.; Queinnec, M.; Luther, J.E.; Bolton, D.K.; White, J.C.; Wulder, M.A.; Van Lier, O.R.; Hermosilla, T. Modelling Lidar-Derived Estimates of Forest Attributes over Space and Time: A Review of Approaches and Future Trends. *Remote Sens. Environ.* **2021**, *260*, 112477. [CrossRef]

15. D'Amico, G.; Vangi, E.; Francini, S.; Giannetti, F.; Nicolaci, A.; Travaglini, D.; Massai, L.; Giambastiani, Y.; Terranova, C.; Chirici, G. Are We Ready for a National Forest Information System? State of the Art of Forest Maps and Airborne Laser Scanning Data Availability in Italy. *iForest* **2021**, *14*, 144–154. [CrossRef]

16. European Commission. *Joint Research Centre. Non-Commercial Light Detection and Ranging (LiDAR) Data in Europe*; Publications Office: Luxembourg, 2021.

17. Zhen, Z.; Quackenbush, L.J.; Zhang, L. Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data. *Remote Sens.* **2016**, *8*, 333. [CrossRef]

18. Zheng, J.; Yuan, S.; Li, W.; Fu, H.; Yu, L. A Review of Individual Tree Crown Detection and Delineation from Optical Remote Sensing Images. *arXiv* **2023**, arXiv:2310.13481. [CrossRef]

19. Kankare, V.; Räty, M.; Yu, X.; Holopainen, M.; Vastaranta, M.; Kantola, T.; Hyyppä, J.; Hyyppä, H.; Alho, P.; Viitala, R. Single Tree Biomass Modelling Using Airborne Laser Scanning. *ISPRS J. Photogramm. Remote Sens.* **2013**, *85*, 66–73. [CrossRef]

20. Syed Hanapi, S.N.H.; Shukor, S.A.A.; Johari, J. A Review on Remote Sensing-Based Method for Tree Detection and Delineation. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *705*, 012024. [CrossRef]

21. Næsset, E. Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data. *Remote Sens. Environ.* **2002**, *80*, 88–99. [CrossRef]

22. White, J.C.; Wulder, M.A.; Varhola, A.; Vastaranta, M.; Coops, N.C.; Cook, B.D. *A Best Practices Guide for Generating Forest Inventory Attributes from Airborne Laser Scanning Data Using the Area-Based Approach*; Canadian Forest Service: Victoria, BC, USA, 2013; ISBN 978-1-100-22385-8.

23. Campbell, J.B.; Wynne, R.H.; Thomas, V.A. *Introduction to Remote Sensing*, 6th ed.; Guilford Press: New York, NY, USA, 2023; ISBN 978-1-4625-4940-5.

24. Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *J. Sens.* **2017**, *2017*, 1353691. [CrossRef]

25. Mura, M.; Bottalico, F.; Giannetti, F.; Bertani, R.; Giannini, R.; Mancini, M.; Orlandini, S.; Travaglini, D.; Chirici, G. Exploiting the Capabilities of the Sentinel-2 Multi Spectral Instrument for Predicting Growing Stock Volume in Forest Ecosystems. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *66*, 126–134. [CrossRef]

26. Nandy, S.; Srinet, R.; Padalia, H. Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India. *Geophys. Res. Lett.* **2021**, *48*, e2021GL093799. [CrossRef]

27. Brosofske, K.D.; Froese, R.E.; Falkowski, M.J.; Banskota, A. A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management. *For. Sci.* **2014**, *60*, 733–756. [CrossRef]

28. Barrett, F.; McRoberts, R.E.; Tomppo, E.; Cienciala, E.; Waser, L.T. A Questionnaire-Based Review of the Operational Use of Remotely Sensed Data by National Forest Inventories. *Remote Sens. Environ.* **2016**, *174*, 279–289. [CrossRef]

29. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

30. Esteban, J.; McRoberts, R.; Fernández-Landa, A.; Tomé, J.; Næsset, E. Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. *Remote Sens.* **2019**, *11*, 1944. [CrossRef]

31. Chirici, G.; Giannetti, F.; McRoberts, R.E.; Travaglini, D.; Pecchi, M.; Maselli, F.; Chiesi, M.; Corona, P. Wall-to-Wall Spatial Prediction of Growing Stock Volume Based on Italian National Forest Inventory Plots and Remotely Sensed Data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *84*, 101959. [CrossRef]

32. Giannetti, F.; Chirici, G.; Vangi, E.; Corona, P.; Maselli, F.; Chiesi, M.; D'Amico, G.; Puletti, N. Wall-to-Wall Mapping of Forest Biomass and Wood Volume Increment in Italy. *Forests* **2022**, *13*, 1989. [CrossRef]

33. Puletti, N.; Grotti, M.; Ferrara, C.; Chianucci, F. Lidar-Based Estimates of Aboveground Biomass through Ground, Aerial, and Satellite Observation: A Case Study in a Mediterranean Forest. *J. Appl. Rem. Sens.* **2020**, *14*, 044501. [CrossRef]

34. Montagnoli, A.; Fusco, S.; Terzaghi, M.; Kirschbaum, A.; Pflugmacher, D.; Cohen, W.B.; Scippa, G.S.; Chiatante, D. Estimating Forest Aboveground Biomass by Low Density Lidar Data in Mixed Broad-Leaved Forests in the Italian Pre-Alps. *For. Ecosyst.* **2015**, *2*, 10. [CrossRef]

35. Camarretta, N.; Puletti, N.; Chiavetta, U.; Corona, P. Quantitative Changes of Forest Landscapes over the Last Century across Italy. *Plant Biosyst. Int. J. Deal. All Asp. Plant Biol.* **2018**, *152*, 1011–1019. [CrossRef]

36. Agnoletti, M.; Piras, F.; Venturi, M.; Santoro, A. Cultural Values and Forest Dynamics: The Italian Forests in the Last 150 Years. *For. Ecol. Manag.* **2022**, *503*, 119655. [CrossRef]

37. Gasparini, P.; Di Cosmo, L.; Floris, A.; De Laurentis, D. (Eds.) *Italian National Forest Inventory—Methods and Results of the Third Survey: Inventario Nazionale Delle Foreste e Dei Serbatoi Forestali Di Carbonio—Metodi e Risultati Della Terza Indagine*; Springer Tracts in Civil Engineering; Springer International Publishing: Cham, Switzerland, 2022; ISBN 978-3-030-98677-3.

38. Balestra, M.; Marselis, S.; Sankey, T.T.; Cabo, C.; Liang, X.; Mokroš, M.; Peng, X.; Singh, A.; Stereńczak, K.; Vega, C.; et al. LiDAR Data Fusion to Improve Forest Attribute Estimates: A Review. *Curr. For. Rep.* **2024**, *10*, 281–297. [CrossRef]

39. Glenn, E.P.; Huete, A.R.; Nagler, P.L.; Nelson, S.G. Relationship between Remotely-Sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors* **2008**, *8*, 2136–2160. [CrossRef] [PubMed]

40. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

41. Chirici, G.; Giannetti, F.; Travaglini, D.; Nocentini, S.; Francini, S.; D'Amico, G.; Calvo, E.; Fasolini, D.; Broll, M.; Maistrelli, F.; et al. Forest damage inventory after the "Vaia" storm in Italy. *Forest* **2019**, *16*, 3–9. [CrossRef]

42. Vicente-Serrano, S.M.; Camarero, J.J.; Olano, J.M.; Martín-Hernández, N.; Peña-Gallardo, M.; Tomás-Burguera, M.; Gazol, A.; Azorin-Molina, C.; Bhuyan, U.; El Kenawy, A. Diverse Relationships between Forest Growth and the Normalized Difference Vegetation Index at a Global Scale. *Remote Sens. Environ.* **2016**, *187*, 14–29. [CrossRef]

43. Huang, S.; Tang, L.; Hupy, J.P.; Wang, Y.; Shao, G. A Commentary Review on the Use of Normalized Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing. *J. For. Res.* **2021**, *32*, 1–6. [CrossRef]

44. Buschmann, C. Fernerkundung von Pflanzen: Ausbreitung, Gesundheitszustand und Produktivitat. *Naturwissenschaften* **1993**, *80*, 439–453. [CrossRef]

45. Gitelson, A.A.; Merzlyak, M.N. Remote Sensing of Chlorophyll Concentration in Higher Plant Leaves. *Adv. Space Res.* **1998**, *22*, 689–692. [CrossRef]

46. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]

47. Sims, D.A.; Rahman, A.F.; Cordova, V.D.; El-Masri, B.Z.; Baldocchi, D.D.; Flanagan, L.B.; Goldstein, A.H.; Hollinger, D.Y.; Misson, L.; Monson, R.K.; et al. On the Use of MODIS EVI to Assess Gross Primary Productivity of North American Ecosystems. *J. Geophys. Res.* **2006**, *111*, 2006JG000162. [CrossRef]

48. Sims, D.; Rahman, A.; Cordova, V.; Elmasri, B.; Baldocchi, D.; Bolstad, P.; Flanagan, L.; Goldstein, A.; Hollinger, D.; Misson, L. A New Model of Gross Primary Productivity for North American Ecosystems Based Solely on the Enhanced Vegetation Index and Land Surface Temperature from MODIS. *Remote Sens. Environ.* **2008**, *112*, 1633–1646. [CrossRef]

49. Tian, H.; Zhu, J.; He, X.; Chen, X.; Jian, Z.; Li, C.; Ou, Q.; Li, Q.; Huang, G.; Liu, C.; et al. Using Machine Learning Algorithms to Estimate Stand Volume Growth of Larix and Quercus Forests Based on National-Scale Forest Inventory Data in China. *For. Ecosyst.* **2022**, *9*, 100037. [CrossRef]

50. White, J.C.; Tompalski, P.; Vastaranta, M.; Wulder, M.; Saarinen, N.; Stepper, C.; Coops, N. A Model Development and Application Guide for Generating an Enhanced Forest Inventory Using Airborne Laser Scanning Data and an Area-Based Approach. 2017. Available online: https://www.researchgate.net/publication/323166566_A_model_development_and_application_guide_for_generating_an_enhanced_forest_inventory_using_airborne_laser_scanning_data_and_an_area-based_approach?channel=doi&linkId=5a83c50245851504fb3a78df&showFulltext=true (accessed on 5 May 2022).

51. Liaw, A.W. Matthew Classification and Regression by RandomForest. *Forest* **2002**, *2*, 18–22.

52. Cadez, L.; Giannetti, F.; De Luca, A.; Tomao, A.; Chirici, G.; Alberti, G. A WebGIS Tool to Support Forest Management at Regional and Local Scale. *iForest* **2023**, *16*, 361–367. [CrossRef]

53. Corona, P. Communicating Facts, Findings and Thinking to Support Evidence-Based Strategies and Decisions. *Ann. Silvic. Res.* **2018**, *42*, 1–2. [CrossRef]

54. Aragoneses, E.; Chuvieco, E. Generation and Mapping of Fuel Types for Fire Risk Assessment. *Fire* **2021**, *4*, 59. [CrossRef]

55. Maciel, E.A.; Martins, V.F.; De Paula, M.D.; Huth, A.; Guilherme, F.A.G.; Fischer, R.; Giles, A.; Barbosa, R.I.; Cavassan, O.; Martins, F.R. Defaunation and Changes in Climate and Fire Frequency Have Synergistic Effects on Aboveground Biomass Loss in the Brazilian Savanna. *Ecol. Model.* **2021**, *454*, 109628. [CrossRef]

56. Travaglini, D.; Simon, G.D.; Puletti, N.; Alberti, G.; Peressotti, A.; Chirici, G.; Corona, P. Stime non Parametriche di Attributi Forestali Con Dati Inventariali e Immagini Telerilevate. In Proceedings of the Atti 14a Conferenza Nazionale ASITA, Brescia, Italy, 9–12 November 2010.

57. Almeida, C.T.D.; Galvão, L.S.; Aragão, L.E.D.O.C.E.; Ometto, J.P.H.B.; Jacon, A.D.; Pereira, F.R.D.S.; Sato, L.Y.; Lopes, A.P.; Graça, P.M.L.D.A.; Silva, C.V.D.J.; et al. Combining LiDAR and Hyperspectral Data for Aboveground Biomass Modeling in the Brazilian Amazon Using Different Regression Algorithms. *Remote Sens. Environ.* **2019**, *232*, 111323. [CrossRef]

58. Jiang, F.; Kutia, M.; Sarkissian, A.J.; Lin, H.; Long, J.; Sun, H.; Wang, G. Estimating the Growing Stem Volume of Coniferous Plantations Based on Random Forest Using an Optimized Variable Selection Method. *Sensors* **2020**, *20*, 7248. [CrossRef]

59. Puliti, S.; Saarela, S.; Gobakken, T.; Ståhl, G.; Næsset, E. Combining UAV and Sentinel-2 Auxiliary Data for Forest Growing Stock Volume Estimation through Hierarchical Model-Based Inference. *Remote Sens. Environ.* **2018**, *204*, 485–497. [CrossRef]

60. Cosenza, D.N.; Korhonen, L.; Maltamo, M.; Packalen, P.; Strunk, J.L.; Næsset, E.; Gobakken, T.; Soares, P.; Tomé, M. Comparison of Linear Regression, k-Nearest Neighbour and Random Forest Methods in Airborne Laser-Scanning-Based Prediction of Growing Stock. *For. Int. J. For. Res.* **2021**, *94*, 311–323. [CrossRef]

61. Sánchez-Ruiz, S.; Moreno-Martínez, Á.; Izquierdo-Verdiguier, E.; Chiesi, M.; Maselli, F.; Gilabert, M.A. Growing Stock Volume from Multi-Temporal Landsat Imagery through Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *83*, 101913. [CrossRef]

62. Giannetti, F.; Chirici, G.; Gobakken, T.; Næsset, E.; Travaglini, D.; Puliti, S. A New Approach with DTM-Independent Metrics for Forest Growing Stock Prediction Using UAV Photogrammetric Data. *Remote Sens. Environ.* **2018**, *213*, 195–205. [CrossRef]

63. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; Koch, B. Importance of Sample Size, Data Type and Prediction Method for Remote Sensing-Based Estimations of Aboveground Forest Biomass. *Remote Sens. Environ.* **2014**, *154*, 102–114. [CrossRef]

64. Luo, S.; Wang, C.; Xi, X.; Pan, F.; Peng, D.; Zou, J.; Nie, S.; Qin, H. Fusion of Airborne LiDAR Data and Hyperspectral Imagery for Aboveground and Belowground Forest Biomass Estimation. *Ecol. Indic.* **2017**, *73*, 378–387. [CrossRef]