



## Aboveground biomass mapping by integrating ICESat-2, SENTINEL-1, SENTINEL-2, ALOS2/PALSAR2, and topographic information in Mediterranean forests

Juan Guerra-Hernández, Lana L. Narine, Adrián Pascual, Eduardo Gonzalez-Ferreiro, Brigitte Botequim, Lonesome Malambo, Amy Neuenschwander, Sorin C. Popescu & Sergio Godinho

To cite this article: Juan Guerra-Hernández, Lana L. Narine, Adrián Pascual, Eduardo Gonzalez-Ferreiro, Brigitte Botequim, Lonesome Malambo, Amy Neuenschwander, Sorin C. Popescu & Sergio Godinho (2022) Aboveground biomass mapping by integrating ICESat-2, SENTINEL-1, SENTINEL-2, ALOS2/PALSAR2, and topographic information in Mediterranean forests, GIScience & Remote Sensing, 59:1, 1509-1533, DOI: [10.1080/15481603.2022.2115599](https://doi.org/10.1080/15481603.2022.2115599)

To link to this article: <https://doi.org/10.1080/15481603.2022.2115599>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 20 Sep 2022.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

## Aboveground biomass mapping by integrating ICESat-2, SENTINEL-1, SENTINEL-2, ALOS2/PALSAR2, and topographic information in Mediterranean forests

Juan Guerra-Hernández<sup>a</sup>, Lana L. Narine<sup>b</sup>, Adrián Pascual<sup>c</sup>, Eduardo Gonzalez-Ferreiro<sup>d</sup>, Brigitte Botequim<sup>e</sup>, Lonesome Malambo<sup>f</sup>, Amy Neuenschwander<sup>f,g</sup>, Sorin C. Popescu<sup>f</sup> and Sergio Godinho<sup>h,i</sup>

<sup>a</sup>Forest Research Centre, School of Agriculture, University of Lisbon, Tapada da Ajuda, Lisboa, Portugal; <sup>b</sup>School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA; <sup>c</sup>Department of Geographical Sciences, University of Maryland, College Park, MD, USA; <sup>d</sup>Departamento de Tecnología Minera, Topografía y de Estructuras, Universidad de León, Ponferrada, Spain; <sup>e</sup>ForestWISE - Collaborative Laboratory for Integrated Forest & Fire Management, Vila Real, Portugal; <sup>f</sup>Department of Ecosystem Science and Management, College Station, TX, USA; <sup>g</sup>Department of Ecosystem Science and Management, NASA Goddard Space Flight Center, Greenbelt, MD, USA; <sup>h</sup>Department EaRSLab—Earth Remote Sensing Laboratory, University of Évora, Évora, Portugal; <sup>i</sup>Institute of Earth Sciences (ICT), Universidade de Évora, Évora, Portugal

### ABSTRACT

The Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) provides an extraordinary opportunity to support global large-scale forest carbon mapping, but further research is needed in order to obtain wall-to-wall forest aboveground biomass (AGB) maps with this technology. The effects of vegetation structure on the performance of canopy height and AGB modeling using ICESat-2 photon-counting light detection and ranging (LiDAR) data in Mediterranean forest areas have not been previously studied in the literature. In this study, we combined recent ICESat-2 vegetation (ATL08) data, Airborne Laser Scanning (ALS)- and field-based estimates, and a multi-sensor earth observation composite for extrapolation of AGB estimates and AGB mapping. A diverse gradient of forest Mediterranean ecosystems, distributed over 19,744.15 km<sup>2</sup> of forest area in the region of Extremadura (Spain), with different species and structural complexity forming 5 different forest types (3 *Quercus* spp. dominated and 2 *Pinus* spp. dominated forests), was used to (i) evaluate the precision of ICESat-2 canopy height estimations, (ii) develop ICESat-2-based AGB models, and (iii) generate a spatially continuous prediction of AGB by using data from the satellite missions Sentinel-1 (S1), Sentinel-2 (S2), Phased Array L-band Synthetic Aperture Radar (ALOS2/PALSAR2), and Shuttle Radar Topography Mission (SRTM). First, ALS- and ICESat-2-derived metrics that best described canopy height (p98 and rh98, respectively) were compared at the ATL08 segment level. Second, ALS-based AGB values were derived at the ATL08 segment scale. Third, ALS-based AGB estimates at the ICESat-2 segment level were used as dependent variables to fit ICESat-2-based AGB models. Fourth, a multi-sensor approach was then implemented to predict ICESat-2-derived AGB, by means of a Random Forest (RF) modeling technique, with predictors retrieved from S1, S2, ALOS2/PALSAR2, and SRTM. Finally, RF was used to generate wall-to-wall AGB maps that were compared with field-, ALS- and ICESat-2-based observations. The agreement between the ALS- and ICESat-2-derived metrics related to the canopy height distribution was higher for *Pinus* spp. forest than for the *Quercus* spp.-dominated forests. The ICESat-2-based AGB models yielded model efficiency (Mef) values between 0.56 and 0.80, with a RMSE ranging from 7.76 to 17.71 Mg ha<sup>-1</sup> and rRMSE from 19.04 to 55.21%. The multi-sensor RF models provided the following results when compared with the ICESat-2- and ALS-based AGB observations:  $R^2$  values of 0.63 and 0.64, and RMSE values of 11.10 Mg ha<sup>-1</sup> (rRMSE = 28.15%) and 12.28 Mg ha<sup>-1</sup> (rRMSE = 31.45%), respectively, and an approximately unbiased result (0.03 Mg ha<sup>-1</sup> and 0.09 Mg ha<sup>-1</sup>). When applied to the field-based validation data set (4th Spanish National Forest Inventory (SNFI-4) plots = 508), the RF-derived AGB model showed a relatively lower predictive capacity ( $R^2 = 0.45$ ), a higher RMSE value (25.88 Mg ha<sup>-1</sup>) and slightly biased results (−1.47 Mg ha<sup>-1</sup>), especially for larger field-derived AGB intervals. The results of this study serve to provide an initial quantitative assessment of the ICESat-2 ATL08 data for large-scale AGB estimation. The findings suggest that a multi-sensor approach may be feasible for extrapolating ICESat-2-derived AGB estimates over areas where field or ALS reference data are not available.

### ARTICLE HISTORY

Received 9 May 2022  
Accepted 12 August 2022

### KEYWORDS

Spaceborne photon-count LiDAR; ALS; L-band SAR; C-band SAR; AGB; forest monitoring

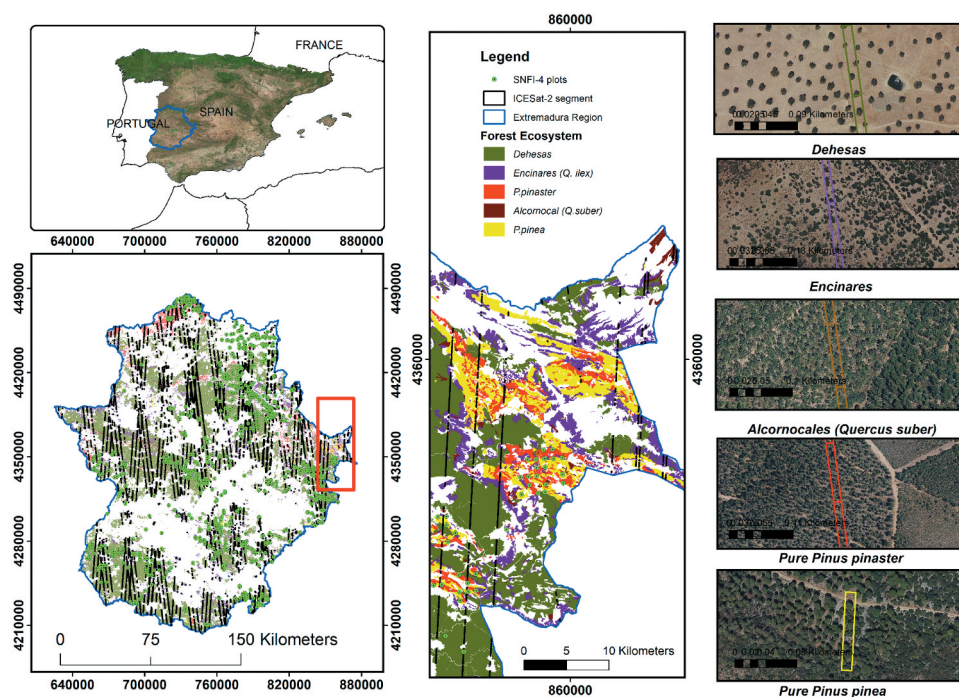
## Introduction

Regional-scale forest above-ground biomass (AGB) mapping using National Forest Inventories (NFI) is important for an effective forest management planning (Ene et al. 2013). Remote sensing data have been widely used for mapping forest inventory variables such as AGB (White et al. 2016; Gherardo et al. 2020; Saarela et al. 2020). Within this domain, active remote sensing is used to describe the three-dimensional (3D) structure of forests (Wulder et al. 2012; McRoberts, Andersen, and Næsset 2014). Airborne laser scanning (ALS) has increasingly been used in the last decade to accurately map important forest variables in Mediterranean forests, e.g. canopy height (Guerra-Hernández et al. 2018), forest inventory variables (Guerra-Hernández et al. 2021), aboveground carbon (Guerra-Hernández et al. 2016) and canopy fuel characteristics (Botequim et al. 2019), at a high spatial resolution.

Large-scale forest mapping using NFI data has relied on the Area-Based Approach (ABA) to build estimation models from which to predict spatially explicit forest attributes over forest landscapes (Chen et al. 2016; Huang et al. 2019a; Tang et al. 2021; Guerra-Hernández et al. 2022). Spatial alignment between NFI field data and co-registration of ALS surveys are being included in many NFI programs to improve the accuracy of the estimates and to enable the ABA method to be applied under robust conditions in terms of positioning. The region of Extremadura (Central-West of Spain) is a good showcase example. New and corrected positions for the fourth Spanish NFI (SNFI-4) samples have been collected, paving the way to improving the performance of ABA at the regional scale and optimizing model fitting with less training data (Pascual et al. 2020; 2021; Guerra-Hernández et al. 2022). However, at larger scales, a design-based NFI approach may not produce accurate estimates at sub-regional or local scales (Tomppo et al. 2008). On the other hand, ALS 3D mapping comes with high acquisition costs for small-scale projects, and budget constraints and the availability of technology constrain the periodic mapping of AGB worldwide with the aim of monitoring carbon stocks and fluxes. Countrywide ALS-acquisitions are increasingly common in rich countries, but they remain an unreachable goal in many less developed countries, including some countries in the Mediterranean basin.

Satellite remote sensing is an alternative approach, to supporting AGB surveys and expanding coverage over remote territories without ALS data. Current spaceborne laser scanning missions have become practical and available options (Narine et al. 2019; Narine, Popescu, and Malambo 2020). The Global Ecosystem Dynamics Investigation (GEDI) (Dubayah et al. 2020) and Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) (Markus et al. 2017; Narine et al. 2019; Neumann et al. 2019), both space-based laser altimeters, provide unprecedented opportunities for observing forest structures worldwide. The GEDI mission operates within the latitudinal limits of the International Space Station (ISS) orbit (i.e.  $\pm 52^\circ$  latitude). The GEDI laser instrument generates a total of 8 ground tracks that are spaced approximately 600 m apart in the cross-track direction. Each track consists of  $\sim 25$  m footprints spaced every 60 m along the ground track within a  $\sim 4.2$  km swath. GEDI waveforms are processed to provide canopy height and profile metrics (GEDI Level 2 product). GEDI Level 3 product includes gridded canopy height metrics products generated from Level 2 products and GEDI level 4A and 4B products provide footprint and gridded aboveground carbon estimates

ICESat-2 uses a different approach to data collection than the GEDI mission. Notably, its revisiting orbit cycle (91 days) and inclination ( $92^\circ$ ) can even yield observations in boreal forests at high latitudes ( $88^\circ$  N/S) (Montesano et al. 2015). ICESat-2 contains a multi-beam photon-counting laser altimeter (Advanced Terrain Laser Altimeter System, ATLAS), which splits a 532 nm laser beam into 6 beams arranged in 3 beam pairs. Each beam pair comprises a weak beam and a strong beam, with a detection ratio of approximately 1:4. ATLAS generates footprints at a laser repetition rate pulse of 10 kHz, resulting in a separation of 0.7 m between shots in an along-track direction, with a footprint size of 10–12 m (Neumann et al. 2019). The mission provides several along-track products, including a dedicated land and vegetation data product or ATL08 (Neuenschwander et al. 2020). The ATL08 product yields ground elevation above sea level, canopy height, and other descriptive variables estimated from a fixed step length of 100 m (as a segment) along the ground track (Neuenschwander and Pitts 2019). The sampling geometry of segment size 100 m  $\times$  12 m guaranteed the availability of sufficient numbers of photons for ground and canopy height estimation (Figure 1).



**Figure 1.** Left: the study area (blue boundary) and the Spanish forest map (SFM25) displaying the five dominant forest ecosystems in the region of Extremadura. Center: a detailed view (representing the extent of the red boundary on the left figure) of the five dominant forest ecosystems and ICESat-2 tracks. Right: a detailed view of the ATL08 100 m along-track segments, with different forest types denoted in color, and overlapping the Spanish National Plan for Aerial Orthophotography (PNOA project) aerial images (0.25 m resolution).

Accuracy assessment of terrain and canopy height estimates from the GEDI and ICESat-2 missions are essential for model calibration and the development of applications to support decision-making. Different studies have assessed the accuracy of canopy height estimates produced by GEDI and ICESat-2 (Potapov et al. 2021; Li et al. 2020; Adam et al. 2020; Lang et al. 2022; Guerra-Hernández and Pascual 2021) and their utility for AGB estimation (Narine et al. 2019; Narine, Popescu, and Malambo 2019, 2020; Duncanson et al. 2020; Neuenschwander et al. 2020; Silva et al. 2021). In most cases, temporally matched or simulated ALS data have been used to confirm canopy heights (Potapov et al. 2021; Li et al. 2020; Adam et al. 2020; Lang et al. 2022) and to calibrate GEDI- and ICESat-2-based AGB models (Narine et al. 2019; Narine, Popescu, and Malambo 2020; Duncanson et al. 2020; Neuenschwander et al. 2020; Silva et al. 2021). As such, the accuracy assessment of post-launch ICESat-2 data and products is fundamentally important (Liu, Cheng, and Chen 2021), because results might differ from those achieved using simulated data and then to use spaceborne laser data to calibrate local and regional ALS-based AGB models (Dorado-Roda et al. 2021).

Therefore, there is still a need to characterize vegetation using ICESat-2 over a wider spectrum of forest landscapes considering vertical and horizontal structural forest complexities.

The ecosystem dynamics in Mediterranean forests are particularly interesting, and the interpretation of single-photon detection technology remains more uncertain than in homogeneous boreal forest conditions (Neuenschwander et al. 2020; Martin, White, and Coops 2021). Testing the performance of ICESat-2 to describe the forest structure and to predict AGB could provide important insights into how the vegetation structure of different forest types influences the performance of ICESat-2-based estimates. The examination of AGB estimation with ICESat-2 can also facilitate comparisons with ALS-based estimation, especially when most studies have been developed for small study areas, where variation in forest complexity was not assessed. Thus, one of the main motivations for conducting this research was to analyze different types of forests over a large region.

This study investigates Mediterranean forests with canopy cover varying from scattered to dense, with the aim of comparing model performance for ALS-

predicted and ICESat-2-derived AGB. In this study, recent ALS surveys (2018–2019) and near coincident ATL08 (version 4) data (2018–2019) were used to develop and evaluate ICESat-2-based models for different forest types. Since the time interval between satellite products and ground truth information may be a source of additional errors, we validated the ICESat-2 ATL08 product (version 4) using ALS data acquired in the same time period (2018–2019). As far as we know, studies evaluating the performance of AGB models derived from the ongoing satellite ICESat-2 mission have not been conducted in Mediterranean forests. Yet, such an understanding is critical as a first step to monitoring terrestrial carbon fluxes with space-based laser data.

ICESat-2 data are acquired in transects over the landscape, and spatially continuous AGB maps are not mission data outcomes (Narine, Popescu, and Malambo 2019). For wall-to-wall estimates, ICESat-2-based AGB must be extrapolated from the segment scale to areas without ICESat-2 coverage (Shen et al. 2018). Therefore, auxiliary data from alternative spaceborne missions are needed to achieve wall-to-wall coverage (Huang et al. 2019a; Li et al. 2020). Quite a few studies have recently demonstrated the feasibility of combining ICESat-2 with multispectral imagery and SAR data from other satellites to produce spatially explicit information for key forest structural variables, such as canopy height and AGB (e.g. Jiang et al. 2021; Li et al. 2020; Nandy, Srinet, and Padalia 2021; Narine, Popescu, and Malambo 2019). Thus, the main goal of this study was to evaluate the usefulness of ICESat-2 data to estimate canopy height and aboveground biomass in five different types of Mediterranean forest in Central-West Spain. The motivations and needs for the study are summarized in the following four specific objectives:

1. Evaluate the accuracy of ICESat-2-derived canopy height statistics by comparing these with ALS-derived metrics collected in the same time period,
2. Apply previously published ALS-based models for Mediterranean forests in the region to analyze the performance of ICESat-2-derived statistics on canopy metrics (height and cover) as exogenous variables to predict AGB,
3. Construct a wall-to-wall map of AGB at 25 m resolution by integrating ICESat-2 with multi-source remotely sensed data (Sentinel-1 (S1), Sentinel-2 (S2), Advanced Land Observing Satellite-2 (ALOS2/

PALSAR2) and the Shuttle Radar Topography Mission (SRTM)), and

4. Compare generated AGB maps with field-, ALS-, and ICESat-2-based AGB observations.

## Material and methods

### Study area, NFI, and ALS data sets

This study was carried out in the region of Extremadura (Central-West Spain) (Figure 1), covering an area of about 27,300 km<sup>2</sup> of diverse forested landscapes. Elevation values vary from 116 m in the Guadiana valley to 2,405 m in the Calvitero peak located in the north-eastern part of the region, presenting a mean altitude of 425 m. The climate is a Mediterranean semi-arid, with a mean annual precipitation varying from <400 mm in the center part of the region to >1000 mm in the northern and eastern mountainous areas. The Spanish Forest Map (Scale: 1:25,000) (SFM25) (MAPA 2018) and the 4th Spanish National Forest Inventory (SNFI-4) datasets were used as field training data for the AGB model, representing a wide spectrum of forest structural complexities in the region (MAGRAMA 2017). The most recent SFM25 version was based on a refined classification of the main species present and forest stage of development using aerial photointerpretation. We selected five dominant Mediterranean forest ecosystems in the region: (i) *Dehesas*, (ii) *Encinares*, (iii), *Pinaster*, (iv) *Alcornocales*, and (v) *Pinea* (for more details about the description of the forest types, see Dorado-Roda et al. 2021).

The set of SNFI-4 training plots used to model AGB consisted of 508 concentric circular plots with 25 m of radius and accurately georeferenced plots (Table 1). These field inventory data were collected in 2017. Tree count and tree-level measurements of diameter at breast height and total height were extrapolated to area-level estimates of tree density ( $N$ , stems ha<sup>-1</sup>), basal area ( $G$ , m<sup>2</sup> ha<sup>-1</sup>) and AGB (Mg ha<sup>-1</sup>). The allometric equations of AGB used in the SNFI for each species were applied at tree level (Montero, Ruiz-Peinado, and Muñoz 2005; Ricardo, Montero González, and Del Rio 2012; Ruiz-Peinado, Del Rio, and Montero 2011). For further details about the SNFI-4 field data and processing, see (Álvarez-González et al. 2014) and Dorado-Roda et al. (2021).

The ALS datasets used were covered the regions of Extremadura: EXT-N (Extremadura north, October 2018

**Table 1.** Summarized field data from SNFI-4 for the five dominant forest types in the region of Extremadura at plot level: aboveground biomass (AGB, Mg ha<sup>-1</sup>), stand basal area (G, m<sup>2</sup> ha<sup>-1</sup>), and tree density (N, trees ha<sup>-1</sup>).

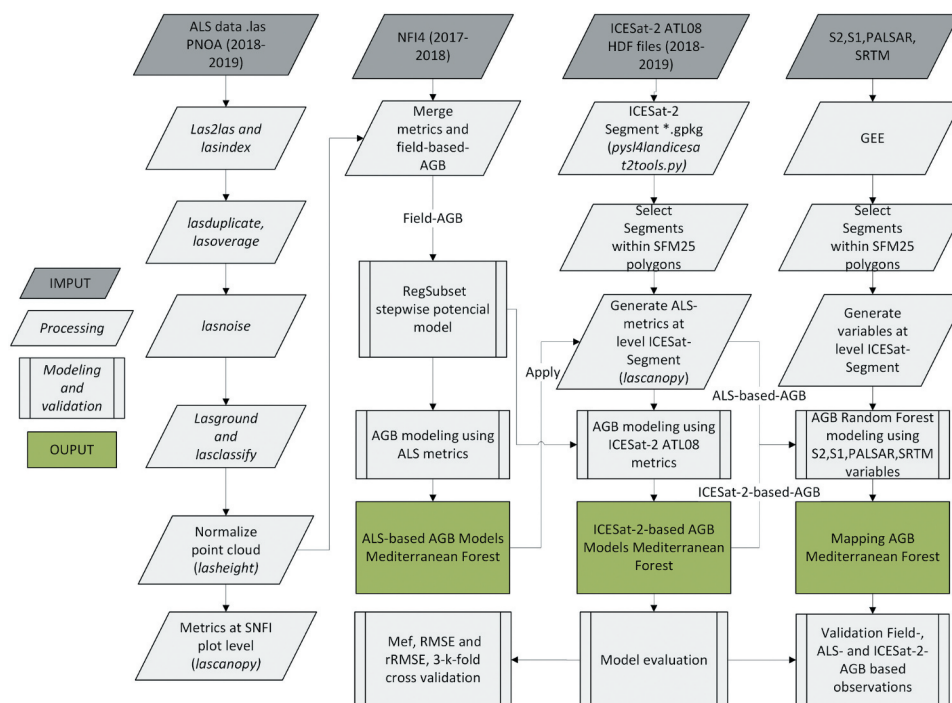
| Forest type  | SNFI-4 plots | Min AGB | Max AGB | Mean AGB | Min G | Max G | Mean G | Min N | Max N   | Mean N |
|--------------|--------------|---------|---------|----------|-------|-------|--------|-------|---------|--------|
| Dehesas      | 239          | 4.11    | 154.36  | 41.20    | 1.13  | 19.50 | 6.17   | 5.09  | 969.08  | 86.37  |
| Encinares    | 90           | 1.72    | 101.56  | 28.25    | 0.43  | 17.80 | 5.32   | 5.09  | 1310.16 | 284.88 |
| Alcornocales | 45           | 1.69    | 112.41  | 29.85    | 0.54  | 25.64 | 8.26   | 10.19 | 1457.15 | 222.21 |
| Pinaster     | 82           | 1.80    | 184.48  | 73.95    | 0.59  | 46.46 | 20.51  | 14.15 | 1464.23 | 348.38 |
| Pinea        | 52           | 11.07   | 159.90  | 49.46    | 2.77  | 39.88 | 12.41  | 29.43 | 1973.52 | 310.88 |

and March 2019) and EXT-5 (Extremadura south, October 2018 and July 2019). These datasets belong to the 2nd round of nationwide ALS measurements collected by the PNOA project (for more details about ALS data acquisition, see 2021).

### Implemented workflow

The complete workflow is shown in Figure 2. ALS point cloud data were processed using LAStools software (Isenburg et al. 2021) following the methodological steps described by Pascual et al. (2020). As a result, 23 ALS-derived metrics (Table 2) computed for the area of each SNFI4 plot were used as potential explanatory variables to build ALS-based AGB estimation models (the AGB estimation models were fitted using field-based AGB as endogenous variable and ALS-derived

metrics as exogenous variables) (see details in Dorado-Roda et al. (2021)). Using an area-based approach (ABA), these models were developed by integrating two Spanish countrywide datasets: the ALS PNOA project (the Spanish National Plan for Aerial Orthophotography and LiDAR) and the fourth Spanish National Forest Inventory (SNFI4). Using the extent and location of ATL08 segments, a range of metrics were first derived from the ALS data for inclusion as explanatory variables in previously published models to produce ALS-based estimates of AGB. *Lascanopy* was used to compute the p98 (98th height percentile), and the previously selected ALS-derived metrics included in the ALS-based AGB estimation models at the ICESat-2 segment level (Figure 2). Second, the ALS- and ICESat-2-derived metrics that best describe canopy height (p98 and rh98, respectively) were compared at



**Figure 2.** Flowchart based on Lastools software, scripts in R and Python programming languages and GIS geospatial sequences used to (i) process ALS data, (ii) extract ALS-derived metrics at the level of the ICESat-2 segments, (iii) generate ICESat-2-based AGB model estimates for each forest ecosystems, and (iv) map AGB using a multi-sensor extrapolation approach.

**Table 2.** Set of ALS-derived metrics, used as potential explanatory variables to fit the AGB estimation models at the ICESat-2 segment level.

| Metric   | Description  |
|--|--|
| <b>(A) Height metrics</b>  |  |
| $h_{mean}$   | Mean   |
| $q_{AV}$   | Quadratic mean height  |
| $h_{std}$  | Standard deviation   |
| $h_{max}, h_{min}$   | Maximum and minimum  |
| $h_{skw}$  | Skewness   |
| $h_{kurt}$   | Kurtosis   |
| $CRR$  | Canopy relief ratio ((mean height - min height) / (max height - min height))   |
| p01, p10, ..., p98, p99  | 5 <sup>th</sup> , 10 <sup>th</sup> , 20 <sup>th</sup> , 25 <sup>th</sup> , 30 <sup>th</sup> , 40 <sup>th</sup> , 50 <sup>th</sup> , 60 <sup>th</sup> , 70 <sup>th</sup> , 75 <sup>th</sup> , 80 <sup>th</sup> , 90 <sup>th</sup> , 95 <sup>th</sup> , 99 <sup>th</sup> percentiles |
| <b>(B) Canopy cover metrics: <math>cover\_cutoff</math>: 2 m</b> |  |
| $CC_{ALS}$   | Percentage of first returns above 2.00/total first returns   |
| $PARA2$  | Percentage of all returns above 2.00/total all returns   |

the ICESat-2 segment level. Third, the ALS-based AGB estimates at the ICESat-2 segment level were used as a dependent variable to fit ICESat-2-based AGB models, which were analyzed in terms of performance. Fourth, the multi-sensor approach was tested to predict AGB, by means of a Random Forest (RF) modeling technique, with predictors retrieved from S1, S2, ALOS2/PALSAR2, and SRTM. Finally, RF was used to generate wall-to-wall AGB maps that were compared with field-, ALS- and ICESat-2-based observations.

### ICESat-2 data acquisition and processing

The boundary of the Extremadura region was used to select all ICESat-2 granules in the study area. The ATL08 V004 data acquired by the ICESat-2 satellite between October 2018 and September 2019 were downloaded and processed using the NSIDC DAAC Data Access and Service API (<https://nsidc.org/data/ATL08/versions/4>) (Neuenschwander, et al. 2021). The ATL08 data product is delivered in HDF5 file format.

The `pysl4landICESat-2.py` (<https://github.com/remotesensinginfo/pysl4land>) tool was used to generate ATL08 \*.gpkg file from HDF5 files for only strong beams. A total of 95 files were created in HDF5 format. ICESat-2 segments that were completely included in

SFM25 polygons were first selected for the five forest types. A final filter was used to exclude ICESat-2 segments where forest height estimates were less than 2 m and where the 98<sup>th</sup> height percentile ( $rh_{98}$ ) values were higher than the maximum tree height measured in the SNFI4 plots within Extremadura. After spatial and data quality filtering ( $night\_flag = 1$  and no scattering or  $msw = 0$ ), we obtained 10,009 sample segments containing valid measurements (Table 3): 7,712 for *Dehesas*, 1,055 for *Encinares*, 436 for *Alcornocales*, 620 for *Pinaster*, and 186 for the *Pinea* forest ecosystem.

The ATL08 algorithm has been updated various times, and version 4 represents an improvement in canopy height estimation at high vegetation densities in order to classify photons as ground, top of the canopy, canopy, and noise (Neuenschwander, et al. 2021). ATL08 contains three canopy height variables, specifically,  $h_{max\_canopy}$  (maximum canopy height),  $h_{canopy}$  (98% relative height), and  $canopy\_h\_metrics$  (i.e. height metrics from the cumulative distribution of relative heights, calculated at 5% intervals for the range 10–95%) (Table 4). Although the  $h_{max\_canopy}$  is equivalent to the  $rh_{100}$  metric (Neuenschwander, et al. 2021), its use can lead to errors because solar background noise may not have been completely removed. Therefore,  $h_{canopy}$  ( $rh_{98}$ , the 98% relative canopy height) can be used as the height at the top of the canopy. The ATL08 algorithm can be affected by solar background noise. Thus, the solar elevation angle, seasonal variation, and beam intensity may impact the signal radiometry which will affect the number of photons in the segment ( $n_{seg\_ph}$ ) classified as ground ( $n_{te\_photons}$ ) and canopy ( $n_{ca\_photons}$ ). In addition to canopy height metrics, two canopies cover metrics (Narine, Malambo, and Popescu 2022) were computed from the ATL08 product, considering a combination of the variables  $n_{ca\_photons}$  (number of canopy photons),  $n_{toc\_photons}$  (number of top canopy photons), and  $n_{te\_photons}$  (number of terrain photons identified in segment) (Table 4).

**Table 3.** Number of strong ATL08 samples during the filtering processing.

| Forest Ecosystem                        | <i>Dehesas</i> | <i>Encinares</i> | <i>Alcornocales</i> | <i>Pinaster</i> | <i>Pinea</i> |
|---|----------------|------------------|---------------------|-----------------|--------------|
| <i>Incial</i>                           | 58,274         | 8384             | 2696                | 3189            | 1172         |
| <i>night_flag = 1 and p98 threshold</i> | 9810           | 2468             | 784                 | 1116            | 389          |
| <i>Final (Clear Skies (msw = 0))</i>    | 7712           | 1055             | 436                 | 620             | 186          |

**Table 4.** Set of number of photon statistics derived from ATL08 data for each ICESat-2 segment ranging within the study area.

| ICESat-2 ATL08 product  |   |        |  |
|---|---|--------|--|
| Label   | Variable ICESat-2-AGB Model                                 | Units  | Description  |
| <i>Height metrics computed fom ATL08</i>                            |   |        |  |
| <i>h_max_canopy (RH100)</i>   | <i>h_max ICT2</i>   | m      | rh100  |
| <i>h_canopy</i>   | <i>h_canopy ICT2</i>  | m      | rh98   |
| <i>Canopy_height_metrics_rh</i>                                     | rh25, rh50, rh60, rh70, rh75, rh80, rh85, rh90, rh95, rh100 | m      | rh 10–95, in 5% intervals. The height metrics are based on a cumulative distribution computed at 5% intervals starting from rh10 to rh95.  |
| <i>h_canopy_quad</i>  | <i>hquad<sub>ICS2</sub></i>                                 | m      | The quadratic mean relative height of relative canopy heights.   |
| <i>h_canopy_mean</i>  | <i>hmean<sub>ICS2</sub></i>                                 | m      | Mean canopy height   |
| <i>h_canopy_median</i>  | <i>hmedian<sub>ICS2</sub></i>                               | m      | Median canopy height   |
| <i>toc_roughness</i>  | <i>toc_roughness</i>  | m      | Standard deviation of relative heights of all photons classified as top of canopy within the segment   |
| Canopy cover metrics computed from ATL08                            |   |        |  |
| <i>Canopy_cover1</i>  | <i>CC1_ICST2</i>  | %      | Percentage canopy and top-of-canopy photos of the total canopy, top-of-canopy and ground photons:<br>$(n_{ca\_photons} + n_{toc\_photons}) / (n_{ca\_photons} + n_{toc\_photons} + n_{te\_photons})$ |
| <i>Canopy_cover2</i>  | <i>CC2_ICST2</i>  | %      | Percentage canopy of total canopy of canopy and ground photons:<br>$n_{ca\_photons} / (n_{ca\_photons} + n_{te\_photons})$   |
| Other parameter for spatial and data quality filtering the segments |   |        |  |
| <i>Canopy_flag</i>  | <i>Canopy_flag</i>  | 1      | Flag indicating that canopy was detected using the Landsat Continuous Cover product for the L-k segment.   |
| <i>ground_track_flag:</i>   | <i>ground_track_flag:</i>                                   | strong | <i>beam_flag:</i> used for identifying power and coverage beams  |
| <i>night_flag</i>   | <i>night_flag</i>   | 1      | <i>date_time:</i> used to determine if data is acquired during the night   |
| <i>msw_flat</i>   | <i>msw_flat</i>   | 0      | An <i>msw_flag</i> value of 0 indicates no scattering observed in the atmosphere, whereas an <i>msw_flag</i> greater than 0 indicates scattering is present.   |

To assess the strength of the relationship between the ICESat-2 and ALS-derived canopy height metrics (i.e. between rh98 and p98), Pearson's correlation coefficient ( $r$ ) (Eq. 1), overall root mean square error (RMSE, Eq. 2), relative root mean square error (rRMSE, Eq. 3), Bias (Eq. 4), and rBias % (Eq. 5) were used.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}_i) \cdot (y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (2)$$

$$rRMSE = \frac{RMSE}{\bar{x}} * 100 \quad (3)$$

$$Bias = \frac{\sum_{i=1}^n (y_i - x_i)}{n} \quad (4)$$

$$rBias = \frac{Bias}{\bar{x}} * 100 \quad (5)$$

where  $n$  is the number of ICESat-2 segments,  $x_i$  is the p98 (m) computed at the level ( $100 \times 12$ -m) of the ICESat-2  $i$ -segment,  $y_i$  is the rh98 (m) computed from ICESat-2 ATL08 product at the level of the ICESat-2

$i$ -segment, and  $\bar{x}$  is the mean of p98 observed at the levels of the ICESat-2 segment.

### ICESat-2-based AGB models

ALS-based AGB estimation models specifically developed for the five forest types are described in detail by Dorado-Roda et al. (2021) (see Section 2.5 for methodology and Table A1 (appendix A) to check the performance of ALS-based AGB prediction models and evaluation of the plot-level accuracy). The ALS-based AGB models were applied to each of the 10,009 ICESat-2 segments for which the complete set of ALS-derived metrics (Table 2) was computed. After predicting the ALS-based AGB (using the previously mentioned ALS-based AGB models), this was used as an endogenous variable to fit the ICESat-2-based AGB models, while the ICESat-2-derived metrics were used as exogenous variables in this process. The steps were similar to those described in Dorado-Roda et al. (2021) (see Section 2.6. GEDI-derived AGB models), but with ICESat-2-derived metrics instead of GEDI-derived metrics, and at the ICESat-2 segment level instead of GEDI footprints. In this case, we proposed restricting the models to two explanatory variables in order for them to be parsimonious, and thus



only two exogenous variables computed from ATL08 (one based on height, and one based on a cover metric) were retained in the final models to estimate AGB. We also applied a 3-fold cross-validation procedure to each potential regression model, using the functions available in R (R Core Team 2020). We then calculated the performance of each forest-type model at the segment level. We used the model efficiency (Mef) statistic (Eq 7), which yields a simple index of relative performance, where Mef values close to 1 indicate “a perfect” model fit, and negative values reveal poor model performance (Vanclay and Peter Skovsgaard 1997). Finally, we computed the overall root mean square error (RMSE, Eq. 8), relative root mean square error (rRMSE, Eq. 9), Bias (Eq. 10), and rBias (Eq. 11) to establish the accuracy of using ICESat-2-based models.

$$Mef = 1 - \left( \frac{(n-1) \sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n-p) \sum_{i=1}^n (y_i - \bar{y}_i)^2} \right) \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (8)$$

$$rRMSE = \frac{RMSE}{\bar{y}} * 100 \quad (9)$$

$$Bias = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{n} \quad (10)$$

$$rBias = \frac{Bias}{\bar{y}} * 100 \quad (11)$$

where  $n$  is the number of segments,  $y_i$  is the ALS-based AGB estimates at the ICESat-2  $i$ -segment level,  $\hat{y}_i$  is the value of AGB estimated from the species specific ICESat-2-based AGB model,  $\bar{y}$  is the mean observed value of the ALS-based AGB estimates at ICESat-2 segment level, and  $p$  is the number of model parameters.

### **Predictor variables for wall-to-wall AGB mapping**

Wall-to-wall statistics from three different satellite remote sensing products were retrieved from (i) multispectral data from S2, (ii) SAR data from S1 and ALOS2/PALSAR2, and (iii) topographical data from SRTM.

S1 and S2 images from October 2018 to October 2019 were processed using the Google Earth Engine (GEE) platform (Gorelick et al. 2017). The S2 data corresponds to the Level-2A product which was already atmospherically corrected using the *sen2cor* processor (Louis et al. 2016). From the 13 S2 spectral bands, only bands at 10 m and 20 m spatial resolution were used: Blue, Green, Red, Red edge 1, 2, and 3, NIR1, NIR2, SWIR1, and SWIR2. Each S2 image was pre-processed to remove clouds and poor-signal pixels using the QA60 band, which is a quality flag developed to identify and mask pixels affected by clouds and cirrus clouds. Based on the cloud-masked S2 imagery, six vegetation indices were calculated, namely the normalized difference vegetation index (NDVI), the normalized difference red edge index 1, 2, and 3 (NDRE1, NDRE2, and NDRE3), the normalized difference infrared index (NDII), and the normalized difference water index (NDWI) (Table A2, appendix A).

S1 is a C-band radar system providing images in dual polarizations (HH+HV, VV+VH) (Torres et al. 2012). In this study, the VV (transmitter-vertical and receiver-vertical) and VH (transmitter-vertical, receiver-horizontal) polarizations were extracted from S1 Level-1 Interferometric Wide Swath (IW) Ground Range Detected (GRD) at a spatial resolution of 10 m. In GEE, the S1 *imagecollection* is processed using the SNAP (Sentinel Application Platform) Toolbox to produce a calibrated and ortho-corrected product. The pre-processing steps include thermal noise removal, radiometric calibration that converts the intensity into normalized backscatter coefficient ( $\sigma^0$ ) in decibels (dB), and terrain correction using the Digital Elevation Model (DEM) data from the SRTM (Rosenqvist et al. 2007). After the pre-processing steps for both S1 and S2 data sets, a median composite image was computed for each band by temporally aggregating high-quality observations for each of the four seasons: autumn, winter, spring, and summer (Table A2, appendix A).

Two annual HH and HV mosaics (2018 and 2019) were obtained from the Phased Array L-band Synthetic Aperture Radar sensor on the Advanced Land Observing Satellite-2 (ALOS2/PALSAR2) of the Japanese Aerospace Exploration Agency (JAXA). These data, processed in GEE, were orthorectified and slope-corrected using the 90 m SRTM DEM, and they were converted to backscatter gamma-naught ( $\gamma_0$ )

values in decibels (dB) using the equation provided by (Rosenqvist et al. 2007). For the S1 and ALOS2/PALSAR2 polarizations, three textural metrics were computed using the Gray-level Co-occurrence Matrix (GLCM) algorithm (Haralick 1979), specifically Mean, Variance, and Homogeneity. These textural metrics were calculated with a window size of  $3 \times 3$  pixels using the R package *gcm* (v.1.6.1) (Chen et al. 2016). As topographical predictor variables, elevation, slope, and aspect metrics were derived from the SRTM DEM data set at a 30-m resolution using GEE. All the above-mentioned datasets were resampled to a 25-m resolution.

## 2.6. Wall-to-wall forest biomass estimation

The AGB estimates derived from ICESat-2 segment-level data (Table 3) were used as dependent variables and were related to multispectral, SAR, and SRTM topographical parameters as predictors, to construct a spatially continuous forest AGB map for the entire area. For this purpose, the Random Forest (RF) algorithm (Breiman 2001) was used to establish the relationship between ICESat-2-derived AGB estimates and up to 91 candidate predictor variables from the multi-source catalog (Table A2, appendix A). The good performance of RF in predicting forest attributes with remote sensing data as explanatory variables has been widely demonstrated (e.g. Gherardo et al. 2020; H. Huang, Liu, and Wang 2019; 2019a).

The RF algorithm has two important tuning parameters, *mtry* and *ntree*. The first refers to the number of predictor variables randomly selected at each split, and  $\sqrt{p}$ , where  $p$  the number of predictor variables, was used to determine the value of *mtry*. The second parameter is the final number of independent trees to be grown. In this study, the *ntree* was set to 1000 trees to produce stabilized variable importance estimates (Liaw and Wiener 2002). For RF modeling, the ICESat-2 segment dataset was divided into training (80%) and test (20%) sets using *createDataPartition* function in the

*caret* R package (Kuhn 2015). This function creates balanced splits in the data by ensuring that random sampling occurs within each forest type while also preserving the overall class distribution over the data set (Kuhn and Johnson 2013). The training dataset was used for RF model building, while the test dataset was used to evaluate model performance by examination of  $R^2$ , RMSE, rRMSE, Bias, and rBias. The spatially continuous AGB maps for the five ecosystems were assessed using the ALS-based AGB estimates (section 2.2) as reference data for computing  $R^2$  and RMSE.

## Results

### Accuracy of ICESat-2- vs ALS-derived metrics

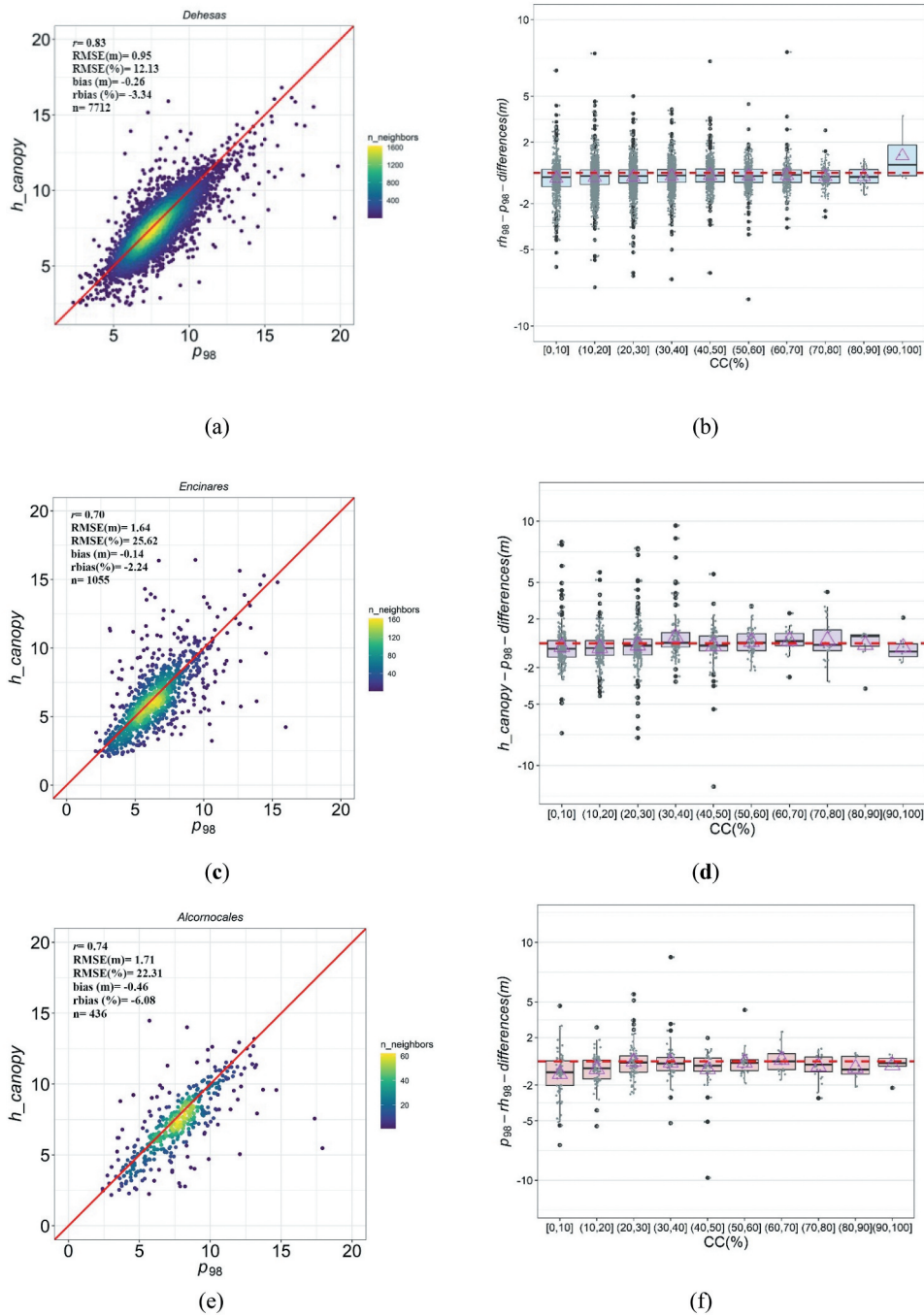
For the five forest ecosystems considered, the p98 – *h\_canopy* (rh98) comparison produced  $r$  Pearson values between 0.70 and 0.83 for *Dehesas*, *Encinares*, and *Alcornocales* and a value of 0.93 for *Pinaster* and *Pinea* forests (Table 5, Figure 3). For *Dehesas*, *Encinares*, *Alcornocales*, *Pinaster*, and *Pinea*, the RMSE values for the p98 – *h\_canopy* (rh98) comparisons were 0.95, 1.64, 1.71, 2.24, and 1.24 m, respectively. The respective rRMSE values were 12.13, 25.62, 22.31, 17.99, and 11.94%, and bias values were –0.26, –0.14, –0.46, –0.56, and –0.30 m. Finally, the mean differences between *h\_canopy* and p98 metrics for  $CC_{ALS}$ -based classification of the values (see Table 2) for all the forest types are shown in Figures 3b, d, f, h and j. For all forest types, negative differences in average bias were found for the rh98 and p98 comparison, which shows that ICESat-2 underestimates canopy height (*rh98*) relative to the estimate derived from ALS p98.

### Performance of ICESat-2-based AGB models

The performance of the ICESat-2-based AGB models for the five forest types considered is shown in Table 6. Scatterplots of ALS-observed against ICESat-2 segment-level AGB estimates are presented in Figure 4 for the best-performing forest-

**Table 5.** Comparison of ALS-based canopy height (p98) and ICESat-2 relative height *h\_canopy* (rh98) metrics.

| Forest Ecosystem    | Metrics comparison | Pearson correlation ( $r$ ) | RMSE (m) | rRMSE (%) | Bias (m) | rBias (%) |
|---------------------|--------------------|-----------------------------|----------|-----------|----------|-----------|
| <i>Dehesas</i>      | p98 – rh98         | 0.83                        | 0.95     | 12.13     | –0.26    | –3.34     |
| <i>Encinares</i>    | p98 – rh98         | 0.70                        | 2.24     | 25.62     | –0.14    | –2.24     |
| <i>Alcornocales</i> | p98 – rh98         | 0.74                        | 1.71     | 22.31     | –0.46    | –6.08     |
| <i>Pinaster</i>     | p98 – rh98         | 0.93                        | 2.22     | 17.99     | –0.56    | –4.47     |
| <i>Pinea</i>        | p98 – rh98         | 0.93                        | 1.24     | 11.94     | –0.30    | –2.93     |



**Figure 3.** Relationships of ALS- and ICESat-2-derived metrics for p98 – rh98 (labeled as  $h\_canopy$  in the box plots) and mean difference between ALS and ICESat-2 metrics by canopy cover  $CC_{ALS}$  (%): *Dehesas* (a and b), *Encinares* (c and d), *Alcornocales* (e and f), *Pinaster* (g and h), *Pinea* (i and j). The solid red line denotes a 1:1 relationship and triangles represent mean differences values between p98 – rh98. Horizontal red-dashed lines represent  $y = 0$ .

specific model, as indicated by Mef. The reported accuracy of AGB estimation determined by 3-fold cross-validation procedures for the best model summarized by forest type is shown in Table A1 and Appendix A. The final models included two variables (one related to the height distribution

and another to canopy cover), which were highly statistically significant ( $P < 0.001$ ) explanatory variables. The positive and negative mean values for bias ( $Mg\ ha^{-1}$ ) and rBias (%) indicate that ICESat-2 overestimated (*Encinares*, *Alcornocales*, *Pinaster*, and *Pinea*) and underestimated (*Dehesa*) the AGB

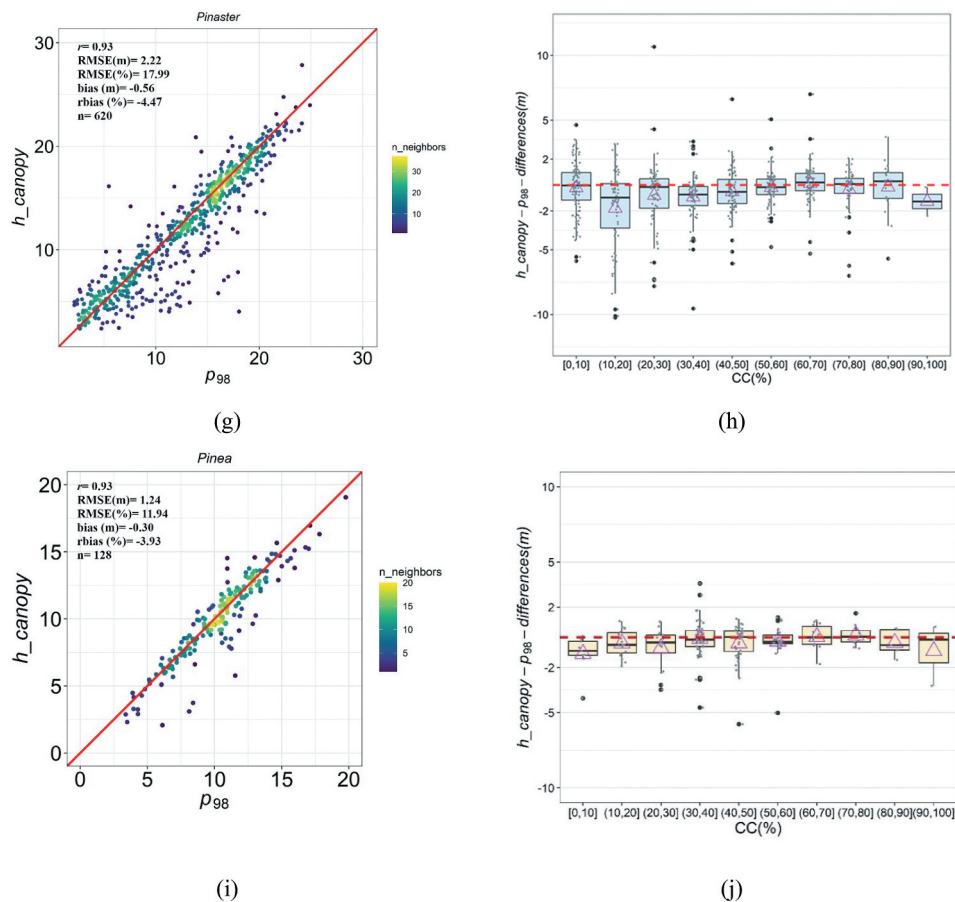


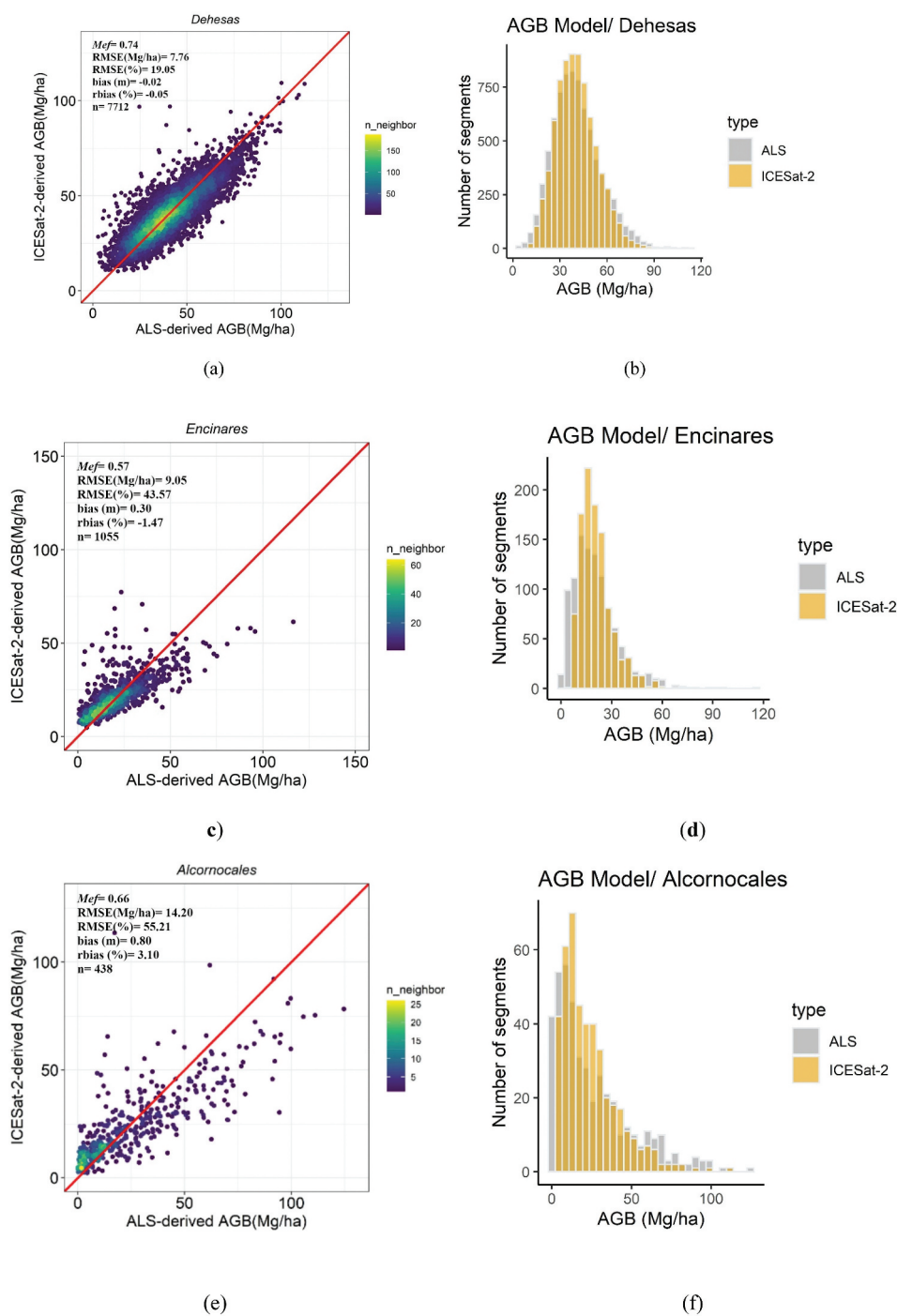
Figure 3. (Continued)

Table 6. Summary of ICESat-2-based AGB models.

| Forest type         | Model   | a          | b         | c         | Regression fitting statistics |                     |                  |                     |                  |
|---------------------|---|------------|-----------|-----------|-------------------------------|---------------------|------------------|---------------------|------------------|
|                     |   |            |           |           | Mef                           | RMSE (Mg/ha)        | rRMSE (%)        | Bias                | rBias (%)        |
| <i>Dehesas</i>      | $AGB = a \cdot hmean_{ICS2}^b \cdot CC1_{ICS2}^c$ | 25.9015*** | 0.6911*** | 0.3480*** | 0.74                          | 7.76                | 19.05            | -0.02               | -0.05            |
| <i>Encinares</i>    | $AGB = a \cdot hquad_{ICS2}^b \cdot CC1_{ICS2}^c$ | 7.4339***  | 1.0168*** | 0.1248*** | 0.57                          | 9.05                | 43.57            | 0.30                | 1.47             |
| <i>Alcornocales</i> | $AGB = a \cdot hmean_{ICS2}^b \cdot CC1_{ICS2}^c$ | 8.8859***  | 1.3421*** | 0.6489*** | 0.66                          | 14.20               | 55.21            | 0.80                | 3.10             |
| <i>Pinaster</i>     | $AGB = a \cdot hmean_{ICS2}^b \cdot CC2_{ICS2}^c$ | 20.0009*** | 0.7926*** | 0.8153*** | 0.80                          | 17.45               | 34.09            | 1.15                | 2.25             |
| <i>Pinea</i>        | $AGB = a \cdot hmean_{ICS2}^b \cdot CC1_{ICS2}^c$ | 11.5694*** | 1.1349*** | 0.9116*** | 0.74                          | 17.71               | 37.54            | 0.02                | 0.04             |
|                     | <b>All SNFI-4 combined Model</b>                  |            |           |           | <b>Mef</b>                    | <b>RMSE (Mg/ha)</b> | <b>rRMSE (%)</b> | <b>Bias (Mg/ha)</b> | <b>rBias (%)</b> |
| <i>Global</i>       | $AGB = a \cdot hmean_{ICS2}^b \cdot CC1_{ICS2}^c$ | 19.6399*** | 0.7319*** | 0.2351*** | 0.64                          | 11.90               | 30.80            | 0.236               | 0.61             |

relative to the ALS-based observations, respectively. Regression models for the five forest types produced Mef values ranging from 0.56 to 0.80. The rRMSE values were slightly lower for *Dehesas* (19.05%) *Pinaster* (34.09%) and *Pinea* (37.54%) than for *Encinares* (44.29%) and *Alcornocales* (55.29%). The results obtained for the best models (histograms Figure 4b, d, f, h, j) indicated that the model-derived AGB estimates were slightly negatively

biased at lower and higher intervals, which correspond to low and high canopy closures, respectively. The models, using mean canopy height ( $hmean_{ICS2}$ ) and  $CC1_{ICS2}$  and  $CC2_{ICS2}$  as predictors, were the most accurate and least unbiased models for all forest types, except the *Encinares* model, which included quadratic mean relative height ( $hquad_{ICS2}$ ). The models for pure homogeneous *Pinaster* and *Pinea* coniferous forest performed



**Figure 4.** Plots showing the relationships between ALS-observed and ICESat-2-segment level estimates of AGB, with associated histograms shown on the right: *Dehesas* (a and b), *Encinares* (c and d), *Alcornocales* (e and f), *Pinaster* (g and h); *Pinea* (i and j), and *Global* (k and j).

best for the observed AGB (Figure 4g, h, i, j). The 3-fold cross-validation indicated that the model performance, as indicated by  $Mef$  and  $RMSE$ , remained stable for all forest types. The non-linear global model (Table 6, Figure 4k, l) yielded  $Mef$  and  $rRMSE$  values of 0.64 and 30.80%, respectively.

### Wall-to-wall forest biomass modeling

The RF model using 91 candidate predictors showed an  $R^2$  value of 64% and a  $RMSE$  value of 10.98 Mg/ha. From that model, the importance of each independent variable in predicting AGB values was assessed by computing the percentage increase in mean

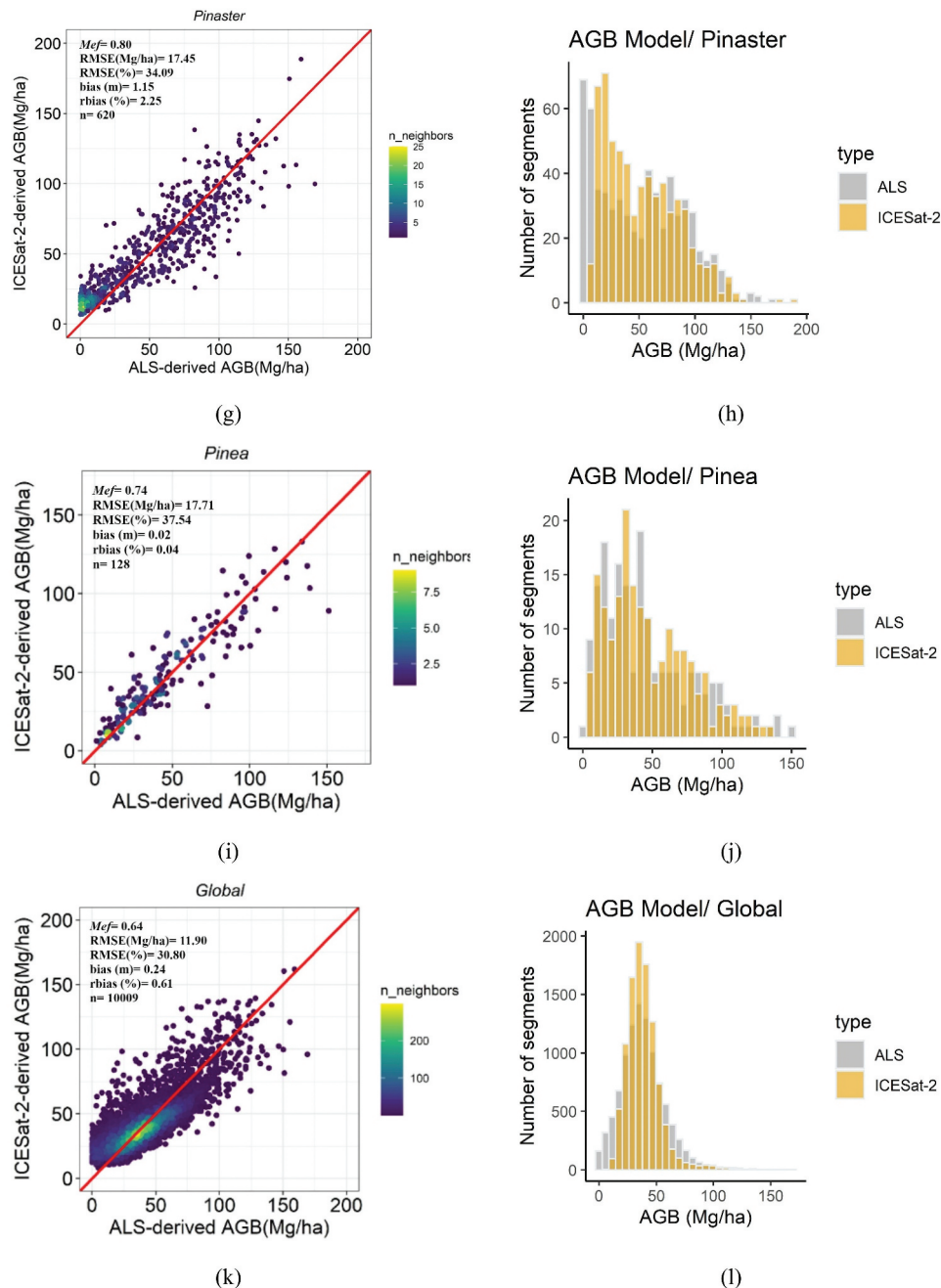
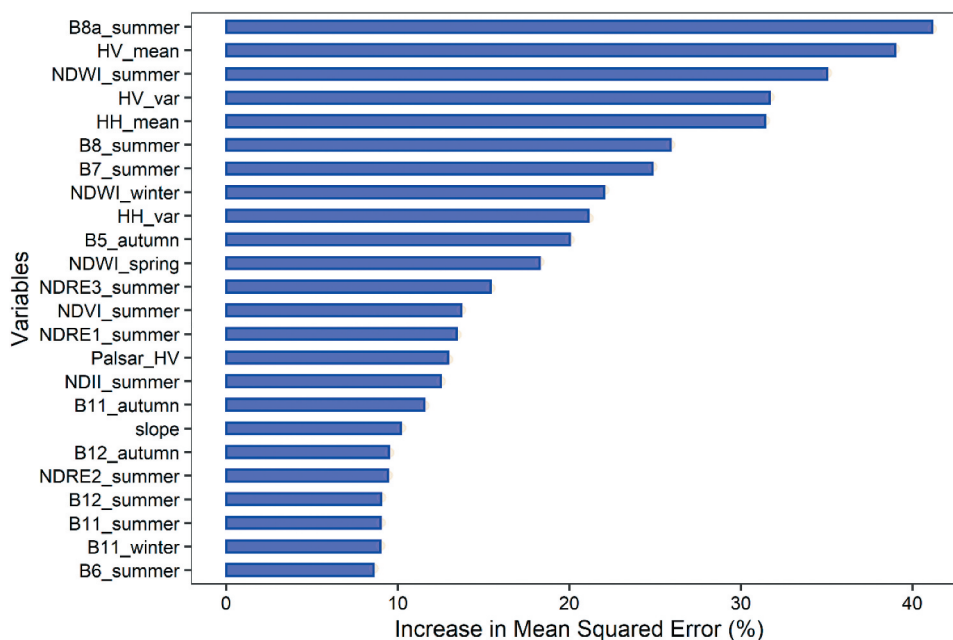


Figure 4. (Continued)

squared error (%IncMSE), where higher values of % IncMSE indicate greater importance of the independent variable. The variables that made a greater contribution to AGB estimation among the 91 independent variables (dashed vertical line in Figure A1, appendix A) are shown in Figure 5. The importance ranking showed that the summer mean composite of the near-infrared band (B8a) was the most important variable (41.15%), followed by the mean HV

SAR texture (38.99%) and the summer NDWI vegetation index (35.03%). The list of the five most important variables was completed by the variance of HV SAR texture (31.69%) and mean HH SAR texture (31.42%). Among the top five variables, the SAR L-band derived backscatter information (HV\_mean, HV\_var, and HH\_mean) had a combined importance of around 102.10%, while the two optical-derived variables (B8a\_summer and NDWI\_summer) represent 76.18%.



**Figure 5.** Importance rank of the 24 most important variables (with %MSE higher than the mean importance).

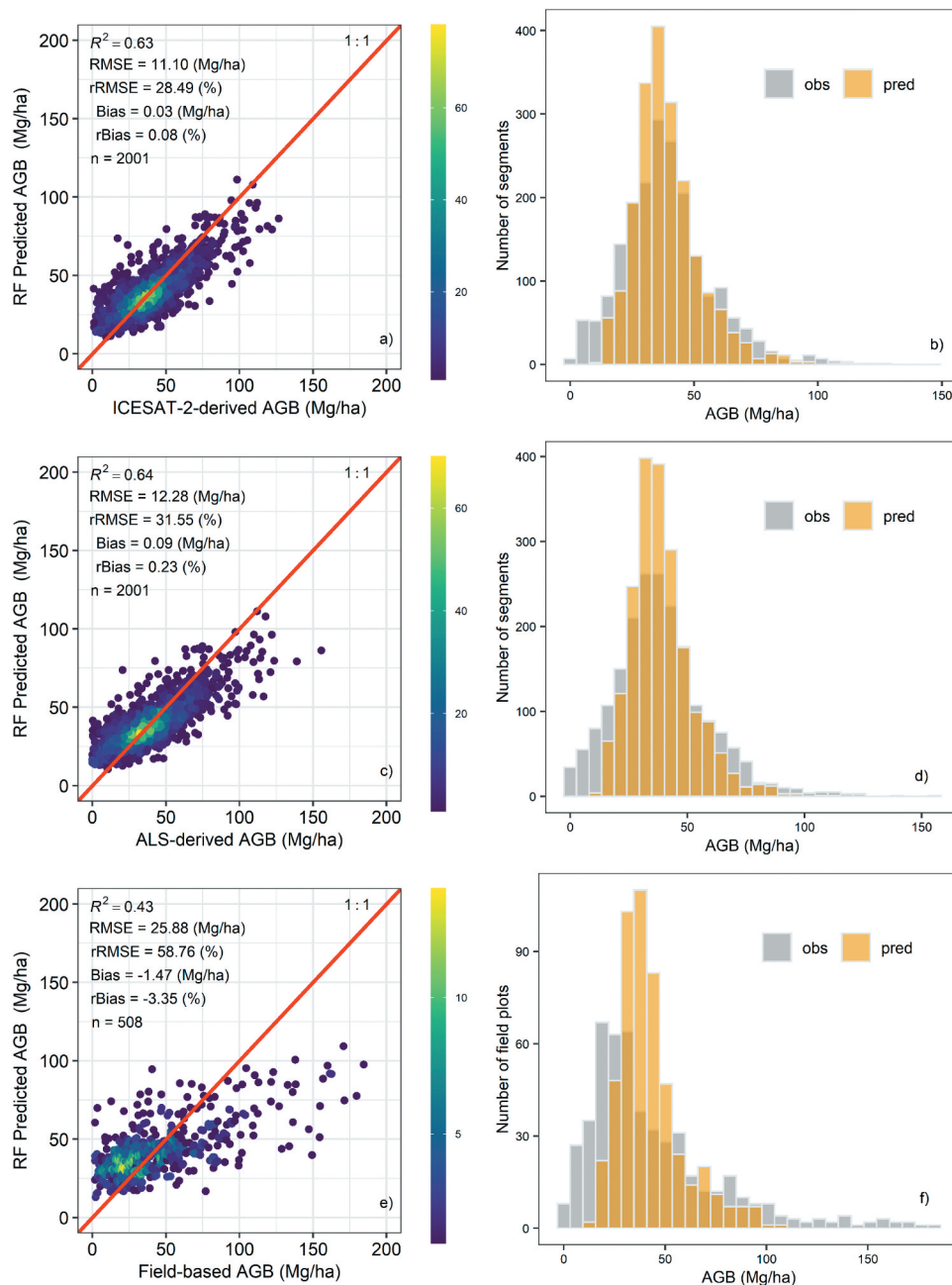
Evaluation of the relevance of the Sentinel-1 C-band polarizations (VV and VH) showed that these contributed little (<6.5%) to forest AGB prediction relative to the L-band polarizations. Regarding the topographical variables, slope was the most important, being ranked 18<sup>th</sup> among the 91 variables tested (Figure 5).

Considering the results from the relative importance assessment, a final RF model was built using only the top 24 most important variables, resulting in a substantial reduction in computation time, while suffering only a minor decrease in accuracy. The predictive accuracy of the RF model for estimating ICESat-2-derived AGB is plotted in Figure 6 for three different independent validation data sets: i) ICESat-2-derived AGB test set; ii) ALS-derived AGB; and iii) field-based AGB estimates. The estimated AGB values of the RF model showed a reasonable accuracy in predicting AGB from both ICESat-2 and ALS-derived AGB data sets, with  $R^2$  values of 0.63 and 0.64 and RMSE values of 11.10 Mg/ha and 12.28 Mg/ha, respectively (Figure 6a and 6c). When applied to the field-based validation data set, the RF model showed a relatively low predictive capacity ( $R^2 = 0.43$ ) and a higher RMSE value (25.88 Mg/ha). The observed mean bias of the RF predictions was 0.03 Mg/ha and 0.09 Mg/ha when compared with the AGB observations from the ICESat-2 and ALS-based independent data sets, respectively. The

calculation of the mean bias for RF predictions and the field-based AGB estimates showed that the RF model produced a negative bias (−1.47 Mg/ha), suggesting greater underestimation of the RF predictions. Overall, the histograms in Figure 6 (b, d, and f) reveal that the RF model tended to overestimate forest AGB in the range 0 to 45 Mg/ha, while the higher (>100 Mg/ha) levels of biomass are underestimated by the RF model. To assess the effects of spatial autocorrelation on the predictive performance of the wall-to-wall mapping model, a spatial 5-fold cross-validation was also performed. To do so, the *blockCV* R package (Roosbeh et al. 2018) was used to compute the variogram and build the spatial blocks (more details are provided in the Appendix). Testing the RF model in a spatial 5-fold cross-validation resulted in an  $R^2$  value of 0.58 and an RMSE of 11.53 Mg/ha, which are similar to those obtained from the ICESat-2-derived AGB test set ( $R^2 = 0.63$ ; RMSE = 11.10 Mg/ha). These results suggest that the existing spatial autocorrelation had an effect, yet small (0.048%), on the predictive performance of the RF model.

### Forest biomass mapping

The map of the predicted forest AGB exhibited general consistency with the ALS-based AGB map across the



**Figure 6.** Relationship between AGB RF-predictions (y-axis) and three different independent data sets (x-axis): a) ICESat-2-derived AGB test set, c) ALS-derived AGB, and e) field-based AGB estimates. Right column represents the histograms of the reference (gray) and the predicted AGB (yellow).

entire study area (Figure 7). The predicted RF AGB values were highly correlated with the ALS-AGB map, with an  $r$  value of 0.71. However, in accordance with the results shown in the histograms in Figure 6, underestimation by the RF model of higher biomass levels ( $>100$  Mg/ha) (Figure 8, local 3) and overestimation of moderate biomass levels (30–45 Mg/ha) (Figure 8, local 2) were clearly observed. Overall, the predicted RF AGB values over the study area range from 5.01 to 112.09 Mg/ha, with an

average of 36.32 Mg/ha and a standard deviation of 12.01 Mg/ha, while from the ALS-derived map the AGB values range from 1.01 to 262.33 Mg/ha, with an average of 35.31 Mg/ha and a standard deviation of 21.17 Mg/ha.

## Discussion

Our study evaluated ICESat-2's performance using robust AGB models under temporally coincident



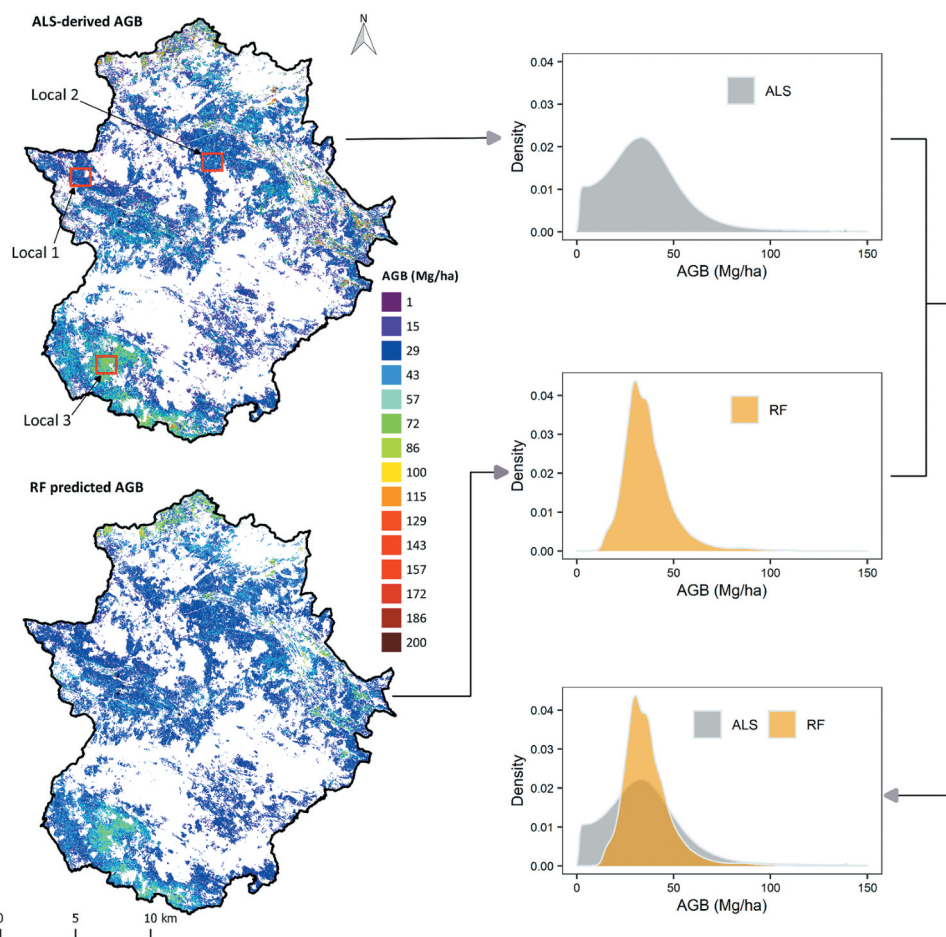


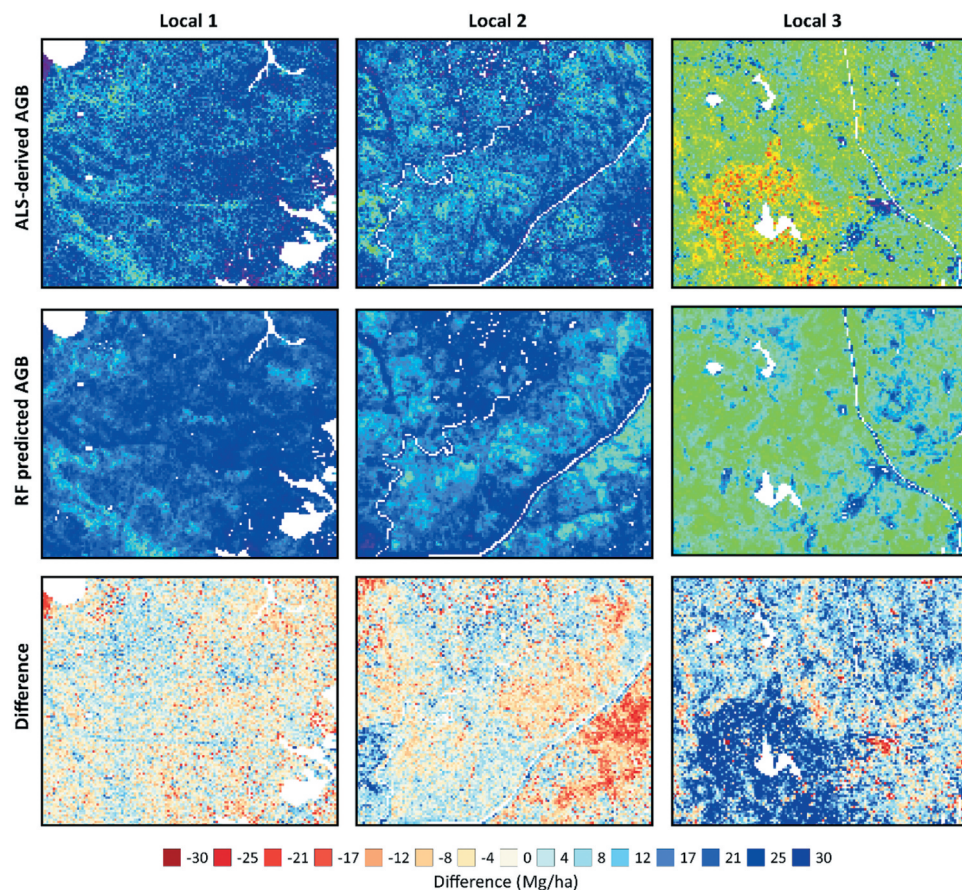
Figure 7. Forest AGB maps derived from the ALS (Guerra-Hernández et al. 2022) (left) and RF models (right)

ALS and ICESat-2 datasets, and over five forest types with different vegetation structures. Despite recent literature evaluating the accuracy of real ICESat-2 data (Liu, Cheng, and Chen 2021; Nandy, Srinet, and Padalia 2021) or ALS-simulated ICESat-2 data (Narine et al. 2019; Narine, Popescu, and Malambo 2020; Neuenschwander et al. 2020; Silva et al. 2021), to the best of our knowledge, our study is the first one evaluating the performance of on-orbit ICESat-2 ATL08 (Version 4) products for characterizing AGB by comparing with spatially and temporally coincident ALS coverage across vast areas of diverse Mediterranean forests. Our study area covers a variety of forest types and thus provides a basis to explore the impacts of forest structure on canopy height and forest AGB biomass estimations. Additionally, this work aimed to map AGB by integrating multi-sensor earth observation data to extrapolate the AGB derived

at the ICESat-2 segment level to other areas without ICESat-2 or ALS coverage.

#### **Accuracy of ICESat-2-derived canopy height**

Our results are consistent with those reported in previously published pre-launch studies that analyzed simulated ICESat-2 data (e.g. Amy and Magruder 2016), which found simulated top-of-canopy heights from ICESat-2, underestimating true top-of-canopy returns for all types analyzed with negative bias ranging from 0.28 m (1.39 m RMSE) to 1.25 m (2.63 m RMSE). However, it is important to evaluate the accuracies of post-launch ICESat-2 data across other forest types beyond boreal forests (Neuenschwander et al. 2020; Martin, White, and Coops 2021). Our results on canopy structure metrics show that the ATL08 relative height metrics (*rh98*) are reasonably accurate, with RMSE values stabilizing



**Figure 8.** Maps comparing the observed ALS-derived AGB and the RF predicted AGB for three different locations. The third row depicts the difference between ALS-derived AGB with the RF predicted AGB. White pixels represent non-forest areas.

at 1.5–2 m in Mediterranean forest (RMSE varied from 0.95 m in *Dehesas* to 2.24 m in *Pinaster*), and substantially better than the values obtained in terms of performance from p98 – rh98 at the GEDI footprint level (level 2A using Version 0001) for the same study area (Dorado-Roda et al. 2021). In terms of bias, the p98 – rh98 relationship was slightly better in *Dehesas* than the other forest types. The results highlight that the accuracy of ICESat-2 canopy heights is influenced by the vegetation structure. ICESat-2 segment estimates were better for regular-even-aged coniferous forests of *P. pinaster* and *P. pinea* species ( $r = 0.93$ ) than in sparse canopy *Quercus*-dominated forests with values of  $r$  ranging from 0.70 (irregular-uneven-aged-multi-layered *Encinares* and *Alcornocales* forests) to 0.83 (sparse homogeneous *Dehesas* forests).

The comparison between ALS- and ICESat-2-derived metrics showed that the results for

coniferous-dominated forest were better in terms of  $r$  than those obtained by Martin, White, and Coops (2021) who compared ICESat-2 ATL03 top of canopy photon height and 90<sup>th</sup> height percentile ALS in boreal forest ( $r = 0.84$ , RMSE = 2.5 m, rRMSE = 19.2%, bias = –2 m). The performance between metrics was better in our study in terms of RMSE, rRMSE, and bias, when compared to a recent study over 40 sites covering the mainland USA using nighttime acquisitions and strong beam data (same as our study) and ALS data (1–4 point  $m^{-2}$ ) collected by the National Ecological Observatory Network (NEON) in 2019 (Liu, Cheng, and Chen 2021) (RMSE = 3.93 m, rRMSE = 24.8%, and negative bias = –0.87,  $n = 7543$  segments). Only the rRMSE for *Encinares* forest produced less accurate values than those reported in Martin, White, and Coops (2021). The comparison between ALS and ICESat-2-derived metrics in this study was slightly better in terms of bias and slightly

worse in terms of rRMSE than the results reported in Neuenschwander et al. (2020). Neuenschwander et al. (2020) compared ICESat-2-derived canopy height ( $h_{canopy} = rh98$ ) with the same metric from the ALS data in boreal forests in Finland (strong beam/night/summer acquisitions, underestimate with a negative bias =  $-0.56$  m, bias% = 3.18%; RMSE % = 13.75%). Our results confirmed negative differences on average bias, between rh98 and p98 (Figure 3), indicating that ICESat-2 underestimates canopy height (rh98) at the segment level when compared with ALS p98 computed at the same extent. This negative bias compared to Neuenschwander et al. (2020), and Liu, Cheng, and Chen (2021) (see figure 11 b) is likely related to the forest vertical structure (2–20 m interval in our study). Overall, our bias and RMSE values were better than those obtained by Neuenschwander et al. (2020) and Liu, Cheng, and Chen (2021), respectively. Mediterranean vegetation has, in general, different structural complexities when compared to boreal forests and the range of type of forest analyzed by Liu, Cheng, and Chen (2021) in the USA. This may be the main reason for the relatively low errors compared with these recently published studies (Neuenschwander et al. 2020; Liu, Cheng, and Chen 2021; Lonesome and Popescu 2021).

The higher precision on the estimation of canopy cover with ICESat-2 was achieved in the 40–90% range of canopy cover. The highest errors were computed in dense canopy closure conditions (>90%) consistent with the findings of Liu, Cheng, and Chen (2021). This confirms that, at low canopy cover (<20%) conditions, both ICESat-2 photon-counting and GEDI full-waveform (FW) LiDAR sensors are more likely to record returns reflected from the terrain rather than canopies, which impedes precise estimation of canopy height (Dubayah et al. 2020a, 2020b). On the contrary, for high canopy cover conditions (CC > 90%), the terrain-reflected signal received by both sensors is weaker than the canopy signal, leading to errors in canopy height measurements (Neuenschwander et al. 2020). Therefore, rh metrics from both missions may be biased especially in low and high canopy cover conditions. Negative bias values exist between canopy height residuals for ICESat-2 when canopy cover exceeds 90%, indicating that the underestimation of canopy may be

due to the overestimation of terrain elevation (Liu, Cheng, and Chen 2021). In summary, our results on canopy height estimation confirmed that (i) ICESat-2 was more accurate in forests with canopy cover ranging between 40% and 85% and (ii) canopy height was underestimated on average by 0.14 to 0.56 m, slightly better than recent studies in boreal forest that showed that canopy height was underestimated on average by 0.5 m – 0.6 m (Amy and Magruder 2019; Neuenschwander et al. 2020; Martin, White, and Coops 2021) and by  $-0.87$  covering the US forest (Liu, Cheng, and Chen 2021).

### Performance of ICESat-2 AGB-derived models

ICESat-2-derived AGB models using mean and quadratic canopy heights and  $CC1_{ICS2}$  and  $CC2_{ICS2}$  as predictors represent a satisfactory description of vegetation structure compared to ALS-based AGB estimates. In terms of RMSE and rRMSE for the five forest ecosystems, the precision of the ICESat-2-derived AGB models were similar or better in the case of *Pinaster* (rRMSE = 36.67%) and *Dehesas* (rRMSE = 28.79%) than those values reported by Narine et al. (2019). In Narine et al. (2019), the simulated ICESat-2-derived 10<sup>th</sup> and 90<sup>th</sup> height percentiles and canopy cover statistics resulted, under linear regression modeling, an  $R^2$  value of 0.74, and an rRMSE of 32% with a RMSE of 25.50 Mg ha<sup>-1</sup>, using 85 segments as training data under a nighttime acquisition scenario. In our study, ICESat-2-based AGB estimates for *Encinares*, *Alcornocales*, and *Pinea* yielded rRMSE values slightly worse than those reported by Narine et al. (2019). The rRMSE values achieved were also similar or better than those reported by Silva et al. (2021) (rRMSE of 54% for Sonoma County (US), using GEDI and ICESat-2 fused AGB and GEDI's AGB models from Duncanson et al. (2020)). In general, the range of values in terms of Mef or  $adjR^2$  from ICESat-2-based models (Mef = 0.56–0.80) at the 100 × 12 m segment level were better than the values obtained by GEDI-based models (Mef = 0.31–0.46) at 25-m footprint level using version 001 in the same Mediterranean formations. The results confirmed the more complex and uneven-aged-multilayered forests as *Encinares* and *Alcornocales* (rRMSE = 44.29% and 55.21%) were, the lower the accuracies in modeling AGB for Mediterranean forests with GEDI and ICESat-2.

Our ICESat-2-derived AGB models were less unbiased in terms of bias and rbias than those values obtained by Duncanson et al. (2020) (bias =  $-26.3 \text{ Mg ha}^{-1}$  and bias% =  $-18.7\%$ ) using US-wide GEDI-based AGB models that rely only on rh metrics and simulated AGB estimates at the footprint level. In general, the ICESat-2-based models were less biased at lower and higher AGB intervals than GEDI-based models, although ICESat-2 models underestimated at lower and higher canopy cover, similar to findings reported with simulated or real GEDI data (Dorado-Roda et al. 2021; Duncanson et al. 2020). Our models achieved slightly better values of Mef (0.56 to 0.80) (similar to adj.  $R^2$ ) than those reported by Silva et al. (2021) (adj.  $R^2$  ranging from 0.46 to 0.51) at object-level segmentation. The Mef values for *Dehesas* (Mef = 0.74), *Pinaster* (Mef = 0.80), and *Pinea* (Mef = 0.74) were similar to the values reported by Narine et al. (2019) (Mef = 0.74) using simulated ICESat-2 data ( $n = 85$ ) in south Texas (US) with predominantly coniferous forests. For the rest of *Quercus*-dominated forests, the ICESat-2-based models from our study got AGB variations slightly worse in terms of Mef, probably due to (i) less complex variety of vegetation structure analyzed in AGB modeling by Narine et al. (2019) and (ii) Narine et al. (2019) and Silva et al. (2021) used ALS-simulated GEDI and ICESat-2 data and the ALS point cloud density was higher than in our study.

The outcomes of our ICESat-2 exercise are better than those reported by (Dorado-Roda et al. 2021) at the 25 m GEDI footprint level over the same region, using the same ALS data for benchmarking. There are important reasons for the differences in the performance between AGB models based on ICESat-2 or GEDI. The impacts of the GEDI (Version 0001) geolocation errors in the study of (Dorado-Roda et al. 2021) were higher, as already discussed, than those from ICESat-2 (geolocation horizontal accuracy of ICESat-2 is  $<5 \text{ m}$  and it gets close to 2–3 m (Neuenschwander, et al. 2021), especially in scattered tree ecosystem as *Dehesas* (Dorado-Roda et al. 2021).

The metrics derived from ATL08 products at the segment level, such as mean canopy height ( $hmean_{ICS2} = rh98$ ), quadratic mean relative height ( $hquad_{ICS2}$ ), and canopy cover metrics ( $CC1_{ICS2}$  and  $CC1_{ICS2}$ ), were found to be significant predictors of AGB. The exhaustive search step to identify the best metric included  $hmean_{ICS2}$ , which is more stable to changes in both the vertical and horizontal canopy

structures (Lefsky et al. 2002; Ni-Meister et al. 2010; Asner et al. 2012; Guerra-Hernández et al. 2016). The use of canopy height metrics alone may ignore some information in profiles with more vegetation structural heterogeneities, such as in natural Mediterranean forest. The canopy cover metric ( $CC1_{ICS2}$  and  $CC1_{ICS2}$ ) also improved the fit in all the ICESat-2-derived AGB models, consistent with a previous study using ICESat-2 simulated data (Narine et al. 2019). Our results also demonstrate that a second metric related to canopy cover from photon-count LiDAR is important for improving ICESat-2 AGB-derived models (Table 6).

### Wall-to-wall forest biomass mapping

The approach of combining wall-to-wall multispectral imagery, SAR, and topographic data to extrapolate estimates of forest AGB values from ICESat-2 tracks to obtain wall-to-wall coverage showed a reasonable agreement when benchmarking with ALS-based AGB predictions. The moderate agreement between RF predictions and ALS-based model estimates of AGB ( $R^2 = 0.64$ ) reflects the challenging study area, and a mosaic of structurally complex Mediterranean forest ecosystems which imposes limitations in acquiring vegetation structure information from space-borne lidar systems (Glenn et al. 2016; Gwenzi et al. 2016). The performance of our RF model in predicting AGB values agrees with previous studies that utilize similar methodological approaches. For instance, Narine et al. (2019), who mapped AGB by combining ICESat-2 and Landsat-8 OLI data, found an RF model explaining 58% of the variations in AGB. In a study carried out by Huang et al. 2019b in China, an AGB map was produced with an  $R^2$  value of 0.64 by integrating ICESat GLAS sensor data with Landsat and PALSAR-derived predictor variables.

To the best of our knowledge, the present study is the first to predict AGB for areas outside of ICESat-2 coverage by combining S-2, S-1, ALOS2/PALSAR2, and topographical data. The variable importance ranking indicated that both spectral information from Sentinel-2 and L-band SAR data from ALOS2/PALSAR2 played an important role in predicting AGB. From the multi-spectral source, it was found that near-infrared and shortwave region (B8a and NDWI) were key in predicting AGB, which is consistent with other studies focused on mapping AGB and tree canopy height

(e.g. (Campbell et al. 2021; Nandy, Srinet, and Padalia 2021; Zhu et al. 2020). With regard to the L-band derived variables, this study also corroborates with findings from Huang et al. (2019a), who showed that texture features of HV and HH polarizations contribute more to AGB predictions than the original backscatter PALSAR data. This may suggest that SAR texture measurements have a higher capability to discriminate spatial information, as well as to reduce the noise in the SAR data (e.g. (Sarker et al. 2012; Laurin et al. 2017). When accounting for the combined variable importance of the top five most important variables, it was found that SAR L-band outperformed the Sentinel-2 derived variables, which demonstrates the usefulness of ALOS-2/PALSAR-2 L-band in predicting AGB. In terms of the Sentinel-1 data, this study showed that SAR C-band polarizations had a limited effect on the RF model accuracy. Previous studies assessing the utility of L- and C-band SAR data for AGB estimation emphasize that L-band data generally produce better results than C-band due to its higher penetration ability through the canopy (e.g. Huang et al. 2018; Naidoo et al. 2015). Therefore, the overall  $R^2$  value of 0.63 and the outcomes from RF-based variable importance might justify the synergetic use of ICESat-2, Sentinel-2, and ALOS2/PALSAR2 L-band SAR satellite data to derive AGB over large areas lacking complete coverage of ALS data and temporally co-registered NFI data to build AGB estimation models.

The RF-approach to generate wall-to-wall maps of AGB may lead to improved temporal resolution as weekly/monthly data become available. As a result, changes from growth or silvicultural activities, for instance, can be traceable. Results from variable importance in the RF modeling exercise highlight the potentialities and capabilities of the upcoming NASA-ISRO Synthetic Aperture Radar (NISAR) satellite mission in 2023 (NISAR 2022) that will deliver denser L-band time-series data at a higher spatial resolution (12 m) with the potential to further boost the approach presented (Khati, Lavallo, and Singh 2021). Nevertheless, and although this methodological approach can be applied in other ecological contexts, the direct application of our RF model to other regions with different ecological characteristics should be approached with caution, particularly for areas where field and ALS reference data are not available.

The accuracy of satellite-based AGB estimations is influenced by many factors, such as the

characteristics of the satellite systems used, the methodological approach implemented, the edaphoclimatic and topographic characteristics of the study area, and the human and equipment errors when measuring vegetation structural parameters from the field (e.g. Arnan et al. 2022). In this study, with the exception of the expected uncertainty associated with tree measurements in the field and the allometric equations used for AGB estimations, a potential source of uncertainty may be related to the sampling nature of the ICESat-2 mission. The sparse spatial distribution and the gaps between ICESat-2 segments over the landscape may have impacted the accuracy of the RF model used for wall-to-wall AGB mapping (e.g. Huang et al. 2019). Therefore, detailed analysis should be carried out to identify the processes of error propagation in the extrapolation processes.

## Conclusions

Real ICESat-2 data were used to assess the use of canopy forest height for estimating AGB in five different Mediterranean forest types representing a wide spectrum of forest structural complexity in the region. The workflow comprises the use of ICESat-2 data and several satellite earth observation (EO) sensors (S1, S2, ALOS2/PALSAR, and STRM), in addition to two Spanish countrywide data sets (ALS PNOA project and SNFI4). The Spanish countrywide data sets were used in the training and validation phases of the ICESat-2-based AGB model development, while the EO data sets were used to scale up the AGB estimates and to generate spatially explicit AGB maps for the study area. This initial assessment of the capacity of ICESat-2 to estimate AGB indicated feasibility of the approach. The results serve as a basis for further extrapolation efforts and are of particular interest for areas or countries for which field and ALS reference data are not available. In this respect, it is important to highlight the strong correlations between the ALS- and the ICESat-2-derived metrics that best describe the top of the canopy height (p98 and rh98, respectively), together with the high percentage of the AGB variability explained by the ICESat-2-based models, since ALS-based AGB (and other key forest variables) observations are currently the most reliable ground truth reference data available for model fitting and extrapolation of estimates. These features are particularly

relevant to scientists investigating methods to better understand AGB dynamics and associated uncertainties using ongoing wall-to-wall satellite image and SAR missions, especially considering the high acquisition costs of field reference data. Moreover, the high correlation between ALS- and the ICESat-2-derived metrics opens the door to further examination of the utility of previously developed species-specific or global AGB models derived in former ALS surveys and expensive field surveys. This is important in relation to country- or region-wide sampling designs, especially in countries that cannot afford to acquire new ALS or field data. Finally, the multi-sensor EO composite was analyzed and validated, producing promising results and revealing an approach that could be considered for extrapolation of AGB estimates and mapping. Nevertheless, further research should be conducted to optimize the design of field data sampling and ALS acquisition in order to reduce costs and to maintain the robustness of the estimates.

## Acknowledgements

We acknowledge Spanish National Forest Inventory (SNFI) from MAPA to make available ground data of the 4th round (SNFI-4) and enhanced plot positioning information to model AGB using ALS statistics. The authors acknowledge the support by Forest Research Centre, a research unit funded by Fundação para a Ciência e a Tecnologia I.P. (FCT), Portugal (UIDB/00239/2021). The authors also thank to project “Apoio à Contratação de Recursos Humanos Altamente Qualificados” (NORTE-06-3559-FSE-000045). under the PORTUGAL 2020 Partnership Agreement. ForestWISE - Collaborative Laboratory for Integrated Forest & Fire Management, was recognized as a ColAB by the Fundação para a Ciência e a Tecnologia, I.P. (FCT). The authors also thank to CILIFO project with grant number 0753\_CILIFO\_5\_E.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research was supported by the project “Extensión del cuarto inventario forestal nacional mediante técnicas LiDAR para la gestión sostenible de los montes de Extremadura” from the Extremadura Forest Service (FEADER nº 1952SE1FR435) and the FUEL-SAT project “Integration of multi-source satellite data for wildland fuel mapping: the role of remote sensing for an effective wildfire fuel management”

from Fundação para a Ciência e a Tecnologia (PCIF/GRF/0116/2019).

## ORCID

Juan Guerra-Hernández  <http://orcid.org/0000-0003-3518-2978>

## Description of authors' responsibilities

Conceptualization, JGH, SG; methodology, JGH, SG, LLN; LM, BB, AM, SSP; software development, JGH, BB, SG; data curation, JGH, SG, LLN; AP, EGF, LM, BB; model validation, JGH, SG, LLN; investigation, AP, EGF, LM, BB, AM, SSP; writing—original draft preparation, JGH, SG, LLN. AP, EGF, LM, BB AM, SSP; writing—review and editing, JGH, SG, LLN; AP, EGF, LM, BB AM, SSP; visualization, JGH, SG, LLN; AP, EGF, LM, BB AM, SSP; supervision, SSP, AM; project administration, JGH, SG, BB, AM, SSP; and funding acquisition, JGH, BB, SG.

## Data Availability

The data that support the findings of this study are available from the corresponding author, Guerra-Hernandez, J., upon reasonable request.

## References

- Adam, M., M. Urbazaev, C. Dubois, and C. Schmillius. 2020. “Accuracy Assessment of GEDI Terrain Elevation and Canopy Height Estimates in European Temperate Forests: Influence of Environmental and Acquisition Parameters.” *Remote Sensing* 12 (23): 3948.
- Álvarez-González, J. G., I. Cañellas, I. Alberdi, K. V. Gadow, and A. D. Ruiz-González. 2014. “National Forest Inventory and Forest Observational Studies in Spain: Applications to Forest Modeling”. *Forest Ecology and Management* Elsevier. 316 54–64. doi:10.1016/j.foreco.2013.09.007.
- Amy, N., and L. A. Magruder. 2016. “The Potential Impact of Vertical Sampling Uncertainty on ICESat-2/ATLAS Terrain and Canopy Height Retrievals for Multiple Ecosystems.” *Remote Sensing* 8 (12): 1039.
- Amy, N., and L. A. Magruder. 2019. “Canopy and Terrain Height Retrievals with ICESat-2: A First Look.” *Remote Sensing* 11 (14): 1721.
- Annan, A., S. De Bruin, M. Herold, S. Quegan, N. Labriere, P. Rodriguez-Veiga, V. Avitabile, M. Santoro, E. T. A. Mitchard, and C. M. Ryan. 2022. “A Comprehensive Framework for Assessing the Accuracy and Uncertainty of Global Above-Ground Biomass Maps.” *Remote Sensing of Environment* 272: 112917.
- Asner, G. P., J. Mascaro, H. C. Muller-Landau, G. Vieilledent, R. Vaudry, M. Rasamoelina, J. S. Hall, and M. van Breugel. 2012. “A Universal Airborne LiDAR Approach for Tropical

- Forest Carbon Mapping." *Oecologia* 168 (4): 1147–1160. doi:10.1007/s00442-011-2165-z.
- Botequim, B., P. M. Fernandes, J. G. Borges, E. González-Ferreiro, and J. Guerra-Hernández. 2019. "Improving Silvicultural Practices for Mediterranean Forests through Fire Behaviour Modelling Using LiDAR-Derived Canopy Fuel Characteristics." *International Journal of Wildland Fire* 28 (11): 823–839. doi:10.1071/WF19001.
- Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/A:1010933404324.
- Campbell, M. J., P. E. Dennison, K. L. Kerr, S. C. Brewer, and W. R. L. Anderegg. 2021. "Scaled Biomass Estimation in Woodland Ecosystems: Testing the Individual and Combined Capacities of Satellite Multispectral and Lidar Data." *Remote Sensing of Environment* 262: 112511. doi:10.1016/j.rse.2021.112511.
- Chen, Q., R. E. McRoberts, C. Wang, and P. J. Radtke. 2016. "Forest Aboveground Biomass Mapping and Estimation across Multiple Spatial Scales Using Model-Based Inference." *Remote Sensing of Environment* 184: 350–360. doi:10.1016/j.rse.2016.07.023.
- Dorado-Roda, I., A. Pascual, S. Godinho, C. A. Silva, B. Botequim, P. Rodríguez-González, E. González-Ferreiro, and J. Guerra-Hernández. 2021. Assessing the Accuracy of GEDI Data for Canopy Height and Aboveground Biomass Estimates in Mediterranean Forests. *Remote Sensing* 13 (12): 12. doi: 10.3390/rs13122279. Multidisciplinary Digital Publishing Institute: 2279
- Dubayah, R., J. Bryan Blair, S. Goetz, L. Fatoyinbo, M. Hansen, S. Healey, M. Hofton, G. Hurtt, J. Kellner, and S. Luthcke. 2020. "The Global Ecosystem Dynamics Investigation: High-Resolution Laser Ranging of the Earth's Forests and Topography." *Science of Remote Sensing* 1. Elsevier: 100002. doi:10.1016/j.srs.2020.100002
- Dubayah, R., M. Hofton, J. B. Blair, J. Armston, and H. Tang. 2020a. "GEDI L2A Elevation and Height Metrics Data Global Footprint Level V001." Data Set. Accessed on 1 June 2020. doi: 10.5067/GEDI/GEDI02\_A.001.
- Dubayah, R., M. Hofton, J. B. Blair, J. Armston, and H. Tang. 2020b. "GEDI L2B GEDI L2B Canopy Cover and Vertical Profile Metrics Data Global Footprint Level V001." Data Set NASA EOSDIS Land Processes DAAC. Accessed on 1 June 2020. doi:10.5067/GEDI/GEDI02\_A.001.
- Duncanson, L., A. Neuenschwander, S. Hancock, N. Thomas, T. Fatoyinbo, M. Simard, C. A. Silva, J. Armston, S. B. Luthcke, and M. Hofton. 2020. "Biomass Estimation from Simulated GEDI, ICESat-2 and NISAR across Environmental Gradients in Sonoma County, California." *Remote Sensing of Environment* 242. Elsevier: 111779. doi:10.1016/j.rse.2020.111779
- Ene, L. T., E. Næsset, T. Gobakken, T. G. Gregoire, G. Ståhl, and S. Holm. 2013. "A Simulation Approach for Accuracy Assessment of Two-Phase Post-Stratified Estimation in Large-Area LiDAR Biomass Surveys". *Remote Sensing of Environment* 133: 210–224. doi:10.1016/j.rse.2013.02.002.
- Gherardo, C., F. Giannetti, R. E. McRoberts, D. Travaglini, M. Pecchi, F. Maselli, M. Chiesi, and P. Corona. 2020. "Wall-to-Wall Spatial Prediction of Growing Stock Volume Based on Italian National Forest Inventory Plots and Remotely Sensed Data." *International Journal of Applied Earth Observation and Geoinformation* 84. Elsevier: 101959. doi:10.1016/j.jag.2019.101947
- Glenn, N. F., A. Neuenschwander, L. A. Vierling, L. Spaete, L. Aihua, D. J. Shinneman, D. S. Pilliod, R. S. Arkle, and S. K. McIlroy. 2016. "Landsat 8 and ICESat-2: Performance and Potential Synergies for Quantifying Dryland Ecosystem Vegetation Cover and Biomass". *Remote Sensing of Environment* 185: 233–242. doi: 10.1016/j.rse.2016.02.039.
- Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202: 18–27. doi:10.1016/j.rse.2017.06.031.
- Guerra-Hernández, J., S. Arellano-Pérez, E. González-Ferreiro, A. Pascual, V. Sandoval Altelarrea, A. Daría Ruiz-González, and J. Gabriel Álvarez-González. 2021. "Developing a Site Index Model for P. Pinaster Stands in NW Spain by Combining Bi-Temporal ALS Data and Environmental Data." *Forest Ecology and Management* 481: 118690. doi:10.1016/j.foreco.2020.118690.
- Guerra-Hernández, J., E. Bastos Görgens, J. García-Gutiérrez, L. Carlos Estraviz Rodriguez, M. Tomé, and E. González-Ferreiro. 2016. "Comparison of ALS Based Models for Estimating Aboveground Biomass in Three Types of Mediterranean Forest." *European Journal of Remote Sensing* 49 (1): 185–204. doi:10.5721/EuJRS20164911.
- Guerra-Hernández, J., B. Botequim, S. Bujan, A. Jurado-Varela, J. Alberto Molina-Valero, A. Martínez-Calvo, and C. Pérez-Cruzado. 2022. "Interpreting the Uncertainty of Model-Based and Design-Based Estimation in Downscaling Estimates from NFI Data: A Case-Study in Extremadura (Spain)." *GIScience & Remote Sensing* 59 (1): 686–704.
- Guerra-Hernández, J., D. N. Cosenza, L. Carlos Estraviz Rodriguez, M. Silva, M. Tomé, R. A. Díaz-Varela, and E. González-Ferreiro. 2018. "Comparison of ALS-and UAV (Sfm)-derived High-Density Point Clouds for Individual Tree Detection in Eucalyptus Plantations." *International Journal of Remote Sensing* 39 (15–16): 5211–5235. doi:10.1080/01431161.2018.1486519.
- Guerra-Hernández, J., and A. Pascual. 2021. "Using GEDI Lidar Data and Airborne Laser Scanning to Assess Height Growth Dynamics in Fast-Growing Species: A Showcase in Spain." *Forest Ecosystems* 8 (1): 14. doi:10.1186/s40663-021-00291-2.
- Gwenzi, D., M. A. Lefsky, V. P. Suchdeo, and D. J. Harding. 2016. "Prospects of the ICESat-2 Laser Altimetry Mission for Savanna Ecosystem Structural Studies Based on Airborne Simulation Data". *ISPRS Journal of Photogrammetry and Remote Sensing* 118: 68–82. doi:10.1016/j.isprsjprs.2016.04.009.
- Haralick, R. M. 1979. "Statistical and Structural Approaches to Texture." In *Proceedings of the IEEE, May 1979*, 67 (5): 786–804. doi:10.1109/PROC.1979.11328.
- Huang, W., K. Dolan, A. Swatantran, K. Johnson, H. Tang, J. O'Neil-Dunne, R. Dubayah, and G. Hurtt. 2019b. "High-

- Resolution Mapping of Aboveground Biomass for Forest Carbon Monitoring System in the Tri-State Region of Maryland, Pennsylvania and Delaware, USA." *Environmental Research Letters*. Vol. 14. IOP Publishing: 095002.
- Huang, H., C. Liu, and X. Wang. 2019. "Constructing a Finer-Resolution Forest Height in China Using ICESat/GLAS, Landsat and ALOS PALSAR Data and Height Patterns of Natural Forests and Plantations". *Remote Sensing* 11 (15). Multidisciplinary Digital Publishing Institute: 1740. doi:10.3390/rs11151740.
- Huang, H., C. Liu, X. Wang, X. Zhou, and P. Gong. 2019a. "Integration of Multi-Resource Remotely Sensed Data and Allometric Models for Forest Aboveground Biomass Estimation in China". *Remote Sensing of Environment* 221: 225–234. doi:10.1016/j.rse.2018.11.017.
- Huang, X., B. Ziniti, N. Torbick, and M. J. Ducey. 2018. "Assessment of Forest Above Ground Biomass Estimation Using Multi-Temporal C-Band Sentinel-1 and Polarimetric L-Band PALSAR-2 Data." *Remote Sensing* 10 (9). Multidisciplinary Digital Publishing Institute: 1424.
- Isenberg, M., K. Friedman, D. Zhao, H. Liao, M. V. Isenberg, E. M. Keating, J. R. N. Vissoci, and C. A. Staton. 2021. "LAStools—Efficient Tools for LiDAR Processing (Version 200509, Academic)." doi:10.1371/journal.pone.0243464.
- Jiang, F., F. Zhao, M. Kaisen, L. Dongsheng, and H. Sun. 2021. "Mapping the Forest Canopy Height in Northern China by Synergizing ICESat-2 with Sentinel-2 Using a Stacking Algorithm." *Remote Sensing* 13 (8): 1535.
- Khati, U., M. Lavallo, and G. Singh. 2021. "Time-Series L-Band SAR and GEDI for Forest above-Ground Biomass Mapping of Sub-Tropical Forest." *Frontiers in Earth Science* 9: 948. doi:10.3389/feart.2021.752254.
- Kuhn, M. 2015. *Caret: Classification and Regression Training*. Astrophysics Source Code Library, ascl-1505.
- Kuhn, M., and K. Johnson. 2013. *Applied Predictive Modeling*. Vol. 26. New York, NY: Springer.
- Lang, N., N. Kalischek, J. Armston, K. Schindler, R. Dubayah, and J. Dirk Wegner. 2022. "Global Canopy Height Estimation with GEDI LIDAR Waveforms and Bayesian Deep Learning." *Remote Sensing of Environment* 268 (January): 112760. doi:10.1016/j.rse.2021.112760.
- Laurin, G., F. Pirotti, M. Callegari, Q. Chen, G. Cuzzo, E. Lingua, C. Notarnicola, and D. Papale. 2017. "Potential of ALOS2 and NDVI to Estimate Forest Above-Ground Biomass, and Comparison with Lidar-Derived Estimates." *Remote Sensing* 9 (1): 18.
- Lefsky, M. A., W. B. Cohen, D. J. Harding, G. G. Parker, S. A. Acker, and S. Thomas Gower. 2002. "Lidar Remote Sensing of Above-Ground Biomass in Three Biomes." *Global Ecology and Biogeography* 11 (5): 393–399. doi:10.1046/j.1466-822x.2002.00303.x.
- Liaw, A., and M. Wiener. 2002. "Classification and Regression by RandomForest." *R News* 2 (3): 18–22.
- Li, W., Z. Niu, R. Shang, Y. Qin, Li, Wang, and H. Chen. 2020. "High-Resolution Mapping of Forest Canopy Height Using Machine Learning by Coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 Data." *International Journal of Applied Earth Observation and Geoinformation* 92: 102163.
- Liu, A., X. Cheng, and Z. Chen. 2021. "Performance Evaluation of GEDI and ICESat-2 Laser Altimeter Data for Terrain and Canopy Height Retrievals." *Remote Sensing of Environment* 264. Elsevier: 112571. doi: 10.1016/j.rse.2021.112571
- Lonesome, M., and S. C. Popescu. 2021. "Assessing the Agreement of ICESat-2 Terrain and Canopy Height with Airborne Lidar over US Ecozones." *Remote Sensing of Environment* 266: 112711.
- Louis, J., V. Debaecker, B. Pflug, M. Main-Knorn, J. Bieniarz, U. Mueller-Wilm, E. Cadau, and F. Gascon. 2016. "Sentinel-2 Sen2Cor: L2A Processor for Users." In *Proceedings Living Planet Symposium 2016, SP-740*, 1–8. Spacebooks Online. ESA Living Planet Symposium 2016, 09 - 13 May 2016, Prague, Czech Republic. ISBN 978-92-9221-305-3.
- MAGRAMA. 2017. "4rd Spanish National Forest Inventory, Ministry of Agriculture, Food and Environment. Ministerio de Agricultura, Alimentación Y Medio Ambiente."
- MAPA. 2018. "Mapa Forestal de España. Escala 1:25.000. Ministerio de Agricultura, Pesca y Alimentación. Dirección General de Desarrollo Rural, Innovación y Política Forestal." [https://www.mapa.gob.es/es/desarrollo-rural/temas/politica-forestal/inventario-cartografia/mapa-forestal-espana/metodologia\\_mfe\\_25.aspx](https://www.mapa.gob.es/es/desarrollo-rural/temas/politica-forestal/inventario-cartografia/mapa-forestal-espana/metodologia_mfe_25.aspx).
- Markus, T., T. Neumann, A. Martino, W. Abdalati, K. Brunt, B. Csatho, S. Farrell, H. Fricker, A. Gardner, and D. Harding. 2017. "The Ice, Cloud, and Land Elevation Satellite-2 (Icesat-2): Science Requirements, Concept, and Implementation". *Remote Sensing of Environment* Elsevier. 190 260–273. doi:10.1016/j.rse.2016.12.029.
- Martin, Q., J. C. White, and N. C. Coops. 2021. "Comparing Airborne and Spaceborne Photon-Counting LiDAR Canopy Structural Estimates across Different Boreal Forest Types." *Remote Sensing of Environment* 262: 112510.
- McRoberts, R. E., H.-E. Andersen, and E. Næsset. 2014. "Using Airborne Laser Scanning Data to Support Forest Sample Surveys." In *Forestry Applications of Airborne Laser Scanning*, 269–292. Dordrecht, Netherlands: Springer.
- Montero, G., R. Ruiz-Peinado, and M. Munoz. 2005. *Producción de Biomasa y Fijación de CO2 Por Los Bosques Españoles*. Madrid: INIA-Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria.
- Montesano, P. M., J. Rosette, G. Sun, P. North, R. F. Nelson, R. O. Dubayah, K. J. Ranson, and V. Kharuk. 2015. "The Uncertainty of Biomass Estimates from Modeled ICESat-2 Returns across a Boreal Forest Gradient." *Remote Sensing of Environment* 158: 95–109. doi: 10.1016/j.rse.2014.10.029.
- Naidoo, L., R. Mathieu, R. Main, W. Kleyhans, K. Wessels, G. Asner, and B. Leblon. 2015. "Savannah Woody Structure Modelling and Mapping Using Multi-Frequency (X-, C-and L-Band) Synthetic Aperture Radar Data". *ISPRS Journal of Photogrammetry and Remote Sensing* 105: 234–250. doi:10.1016/j.isprsjprs.2015.04.007.
- Nandy, S., R. Srinet, and H. Padalia. 2021. "Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest



- Algorithm in Northwest Himalayan Foothills of India." *Geophysical Research Letters* 48 (14): e2021GL093799.
- Narine, L. L., L. Malambo, and S. C. Popescu. 2022. "Characterizing Canopy Cover with ICESat-2: A Case Study of Southern Forests in Texas and Alabama, USA." *Remote Sensing of Environment*. Received for production.
- Narine, L. L., S. C. Popescu, and L. Malambo. 2019. "Synergy of ICESat-2 and Landsat for Mapping Forest Aboveground Biomass with Deep Learning." *Remote Sensing* 11 (12): 1503.
- Narine, L. L., S. C. Popescu, and L. Malambo. 2020. "Using ICESat-2 to Estimate and Map Forest Aboveground Biomass: A First Example." *Remote Sensing* 12 (11): 1824.
- Narine, L. L., S. Popescu, A. Neuenschwander, T. Zhou, S. Srinivasan, and K. Harbeck. 2019. "Estimating Aboveground Biomass and Forest Canopy Cover with Simulated ICESat-2 Data". *Remote Sensing of Environment* 224: 1–11. doi:10.1016/j.rse.2019.01.037.
- Neuenschwander, A., E. Guenther, J. C. White, L. Duncanson, and P. Montesano. 2020. "Validation of ICESat-2 Terrain and Canopy Heights in Boreal Forests." *Remote Sensing of Environment* 251. Elsevier: 112110. doi:10.1016/j.rse.2020.112110
- Neuenschwander, A., P. K.I, B. P. Jelley, J. Robbins, Klotz, S. C. P. B, R. F. Nelson, D. Harding, D. Pederson, and R. Sheridan. 2021. "ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 4." Data Set NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/ATLAS/ATL08.003>. Accessed 2021-06-01
- Neuenschwander, A., and K. Pitts. 2019. "The ATL08 Land and Vegetation Product for the ICESat-2 Mission". *Remote Sensing of Environment* 221: 247–259. doi:10.1016/j.rse.2018.11.005.
- Neumann, T. A., A. J. Martino, T. Markus, S. Bae, M. R. Bock, A. C. Brenner, K. M. Brunt, J. Cavanaugh, S. T. Fernandes, and D. W. Hancock. 2019. "The Ice, Cloud, and Land Elevation Satellite-2 Mission: A Global Geolocated Photon Product Derived from the Advanced Topographic Laser Altimeter System". *Remote Sensing of Environment* 233: 111325. doi:10.1016/j.rse.2019.111325.
- Ni-Meister, W., S. Lee, A. H. Strahler, C. E. Woodcock, C. Schaaf, K. J. R. Tian Yao, G. Sun, and J. Bryan Blair. 2010. "Assessing General Relationships between Aboveground Biomass and Vegetation Structure Parameters for Improved Carbon Estimate from Lidar Remote Sensing." *Journal of Geophysical Research: Biogeosciences* G2 (2005–2012) 115 <http://onlinelibrary.wiley.com/doi/10.1029/2009JG000936/full>
- NISAR. 2022. "NASA-ISRO SAR Mission (NISAR). Disponible at. [https://Nisar.Jpl.Nasa.Gov/.](https://Nisar.Jpl.Nasa.Gov/)"
- Pascual, A., J. Guerra-Hernández, D. N. Cosenza, and V. Sandoval-Altelaarrea. 2021. "Using Enhanced Data Co-Registration to Update Spanish National Forest Inventories (NFI) and to Reduce Training Data under LiDAR-Assisted Inference." *International Journal of Remote Sensing* 42 (1): 126–147.
- Pascual, A., J. Guerra-Hernández, D. N. Cosenza, and V. Sandoval-Altelaarrea. 2020. "The Role of Improved Ground Positioning and Forest Structural Complexity When Performing Forest Inventory Using Airborne Laser Scanning." *Remote Sensing* 12 (3): 413. doi:10.3390/rs12030413.
- Potapov, P., L. Xinyuan, A. Hernandez-Serna, A. Tyukavina, M. C. Hansen, A. Kommareddy, A. Pickens, S. Turubanova, H. Tang, and C. Edibaldo Silva. 2021. "Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data." *Remote Sensing of Environment* 253. Elsevier: 112165. doi:10.1016/j.rse.2020.112165
- R Core Team. 2020. "R: A Language and Environment for Statistical Computing (Version 3.6. 1)[Computer Software]." In *Vienna*. Austria: R Foundation for Statistical Computing.
- Ricardo, R.-P., G. Montero González, and M. Del Rio. 2012. Biomass Models to Estimate Carbon Stocks for Hardwood Tree Species. *Forest Systems*. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) 21 (1): 42–52. doi:10.5424/fs/2112211-02193.
- Roosbeh, V., J. Elith, J. J. Lahoz-Monfort, and G. Guillera-Arroita. 2018. "BlockCV: An R Package for Generating Spatially or Environmentally Separated Folds for k-Fold Cross-Validation of Species Distribution Models." *BioRxiv. Cold Spring Harbor Laboratory* 357798.
- Rosenqvist, A., M. Shimada, N. Ito, and M. Watanabe. 2007. "ALOS PALSAR: A Pathfinder Mission for Global-Scale Monitoring of the Environment." *IEEE Transactions on Geoscience and Remote Sensing* 45 (11): 3307–3316.
- Ruiz-Peinado, R., M. Del Rio, and G. Montero. 2011. "New Models for Estimating the Carbon Sink Capacity of Spanish Softwood Species." *Forest Systems* 20 (1): 176–188. doi:10.5424/fs/2011201-11643.
- Saarela, S., A. Wästlund, E. Holmström, A. Appiah Mensah, S. Holm, M. Nilsson, J. Fridman, and G. Ståhl. 2020. "Mapping Aboveground Biomass and Its Prediction Uncertainty Using LiDAR and Field Data, Accounting for Tree-Level Allometric and LiDAR Model Errors." *Forest Ecosystems* 7 (1): 1–17.
- Sarker, M. L. R., J. Nichol, I. Huseyin Baki, B. Bin Ahmad, and A. Abdul Rahman. 2012. "Forest Biomass Estimation Using Texture Measurements of High-Resolution Dual-Polarization C-Band SAR Data." *IEEE Transactions on Geoscience and Remote Sensing* 51 (6): 3371–3384.
- Shen, W., L. Mingshi, C. Huang, X. Tao, and A. Wei. 2018. "Annual Forest Aboveground Biomass Changes Mapped Using ICESat/GLAS Measurements, Historical Inventory Data, and Time-Series Optical and Radar Imagery for Guangdong Province, China". *Agricultural and Forest Meteorology* 259: 23–38. doi:10.1016/j.agrformet.2018.04.005.
- Silva, C. A., L. Duncanson, S. Hancock, A. Neuenschwander, N. Thomas, M. Hofton, L. Fatoyinbo, et al. 2021. "Fusing Simulated GEDI, ICESat-2 and NISAR Data for Regional Aboveground Biomass Mapping." *Remote Sensing of Environment* 253:112234. doi:10.1016/j.rse.2020.112234.

- Tang, H., M. Lei, A. Lister, J. O'Neill-Dunne, L. Jiaming, R. L. Lamb, R. Dubayah, and G. Hurtt. 2021. "High-Resolution Forest Carbon Mapping for Climate Mitigation Baselines over the RGGI Region, USA." *Environmental Research Letters* 16 (3): 035011.
- Tomppo, E., M. Haakana, M. Katila, and J. Peräsaari. 2008. *Multi-Source National Forest Inventory: Methods and Applications*. Vol. 18. Switzerland:: Springer Science & Business Media.
- Torres, R., P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, B. Rommen, N. Floury, and M. Brown. 2012. "GMES Sentinel-1 Mission". *Remote Sensing of Environment* 120 9–24. doi:10.1016/j.rse.2011.05.028.
- Vanclay, J. K., and J. Peter Skovsgaard. 1997. "Evaluating Forest Growth Models." *Ecological Modelling* 98 (1): 1–12. doi:10.1016/S0304-3800(96)01932-1.
- White, J. C., N. C. Coops, M. A. Wulder, M. Vastaranta, T. Hilker, and P. Tompalski. 2016. "Remote Sensing Technologies for Enhancing Forest Inventories: A Review." *Canadian Journal of Remote Sensing* 42 (5): 619–641. doi:10.1080/07038992.2016.1207484.
- Wulder, M. A., J. C. White, R. F. Nelson, E. Næsset, H. Ole Ørka, N. C. Coops, T. Hilker, C. W. Bater, and T. Gobakken. 2012. "Lidar Sampling for Large-Area Forest Characterization: A Review". *Remote Sensing of Environment* 121: 196–209. doi:10.1016/j.rse.2012.02.001.
- Zhu, X., C. Wang, S. Nie, F. Pan, X. Xiaohuan, and H. Zhenyue. 2020. "Mapping Forest Height Using Photon-Counting LiDAR Data and Landsat 8 OLI Data: A Case Study in Virginia and North Carolina, USA." *Ecological Indicators* 114. Elsevier: 106287. doi:10.1016/j.ecolind.2020.106287