



**HAL**  
open science

## High-resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2 data with a deep learning approach

Martin Schwartz, Philippe Ciais, Catherine Ottlé, Aurelien de Truchis, Cedric Vega, Ibrahim Fayad, Martin Brandt, Rasmus Fensholt, Nicolas Baghdadi, François Morneau, et al.

### ► To cite this version:

Martin Schwartz, Philippe Ciais, Catherine Ottlé, Aurelien de Truchis, Cedric Vega, et al.. High-resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2 data with a deep learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 2024, 128, pp.103711. 10.1016/j.jag.2024.103711 . hal-04489474

**HAL Id: hal-04489474**

**<https://hal.science/hal-04489474v1>**

Submitted on 5 Mar 2024

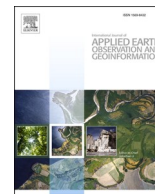
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Contents lists available at ScienceDirect

# International Journal of Applied Earth Observation and Geoinformation

journal homepage: [www.elsevier.com/locate/jag](http://www.elsevier.com/locate/jag)

## High-resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2 data with a deep learning approach

Martin Schwartz<sup>a,\*</sup>, Philippe Ciais<sup>a</sup>, Catherine Ottlé<sup>a</sup>, Aurelien De Truchis<sup>b</sup>, Cedric Vega<sup>c</sup>, Ibrahim Fayad<sup>d</sup>, Martin Brandt<sup>e</sup>, Rasmus Fensholt<sup>e</sup>, Nicolas Baghdadi<sup>d</sup>, François Morneau<sup>f</sup>, David Morin<sup>g</sup>, Dominique Guyon<sup>h</sup>, Sylvia Dayau<sup>h</sup>, Jean-Pierre Wigneron<sup>h</sup>

<sup>a</sup> Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris Saclay, 91191 Gif-sur-Yvette, France

<sup>b</sup> Kayrros SAS, Paris 75009, France

<sup>c</sup> IGN, Laboratoire d'Inventaire Forestier, 54000 Nancy, France

<sup>d</sup> INRAE, UMR TETIS, Université de Montpellier, AgroParisTech, CEDEX 5, 34093 Montpellier, France

<sup>e</sup> Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark

<sup>f</sup> IGN, Service de l'Information Statistique, Forestière et Environnementale, 45290 Nogent-sur-Vernisson, France

<sup>g</sup> CESBIO, Université de Toulouse, CNES/CNRS/INRAE/IRD/UPS, 18 Av. Edouard Belin, bpi 2801, CEDEX 9, 31401 Toulouse, France

<sup>h</sup> ISPA, UMR 1391, INRA Nouvelle-Aquitaine, Bordeaux Villenave d'Ornon, France

### ARTICLE INFO

#### Keywords:

Forest height  
GEDI  
Sentinel-1  
Sentinel-2  
U-Net  
Deep Learning  
Landes forest  
Forest Inventory  
3D Stereo

### ABSTRACT

In intensively managed forests in Europe, where forests are divided into stands of small size and may show heterogeneity within stands, a high spatial resolution (10–20 m) is needed to capture the differences in canopy height. In this work, we developed a deep learning model based on multi-sensor remote sensing measurements to create a high-resolution canopy height map over the “Landes de Gascogne” forest in France, a large maritime pine plantation of 13,000 km<sup>2</sup> with flat terrain and intensive management. This area is characterized by even-aged and mono-specific stands, of a typical length of a few hundred meters, harvested every 35 to 50 years. Our deep learning U-Net model uses multi-band images from Sentinel-1 and Sentinel-2 with composite time averages as input to predict tree height derived from GEDI waveforms. The evaluation is performed with external validation data from forest inventory plots and a stereo 3D reconstruction model based on Skysat imagery available at specific locations. We trained seven different U-Net models based on combinations of Sentinel-1 and Sentinel-2 bands to evaluate the importance of each sensor in the dominant height retrieval. The model outputs allow us to generate a 10 m resolution canopy height map of the whole “Landes de Gascogne” forest area for 2020 with a mean absolute error of 2.02 m on the test dataset. The best predictions were obtained using all available bands from Sentinel-1 and Sentinel-2 but using only one satellite source also provided good predictions. For all validation datasets in coniferous forests, our model showed better metrics than previous canopy height models available in the same region.

### 1. Introduction

Forest biomass plays a crucial role in the global carbon cycle, and its preservation or increase is an essential element of land-based mitigation policies (Griscom et al., 2017; IPCC, 2019; Pan et al., 2011; UNFCCC, 2015). In intensively managed European forests, which are divided into small stands, a high spatial resolution (10–20 m) is needed to capture the difference in biomass between adjacent stands or within stands in case of heterogeneous forest structure. Canopy height, combined with other

forest parameters such as tree species, may be a good proxy for forest biomass estimation (Duncanson et al., 2022). Forest inventories have been the only method to estimate forest biomass and height in the past. They provide reliable statistical information on forests over large regions but are not designed to produce high-resolution maps. In recent decades, remote sensing data opened the possibility of height estimation at a finer scale with more data collected. Airborne Laser Scanning (ALS) may provide accurate height estimations consistent with forest inventory measurements at a high spatial resolution. These data, used as

\* Corresponding author.

E-mail address: [martin.schwartz@lsce.ipsl.fr](mailto:martin.schwartz@lsce.ipsl.fr) (M. Schwartz).

<https://doi.org/10.1016/j.jag.2024.103711>

Received 9 March 2023; Received in revised form 30 January 2024; Accepted 8 February 2024

Available online 21 February 2024

1569-8432/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

reference height for models using space-borne images, have a high potential to accurately describe forest structures (Wilkes et al., 2015). However, their acquisition is costly, and measurement campaigns are sparse in time, which does not enable up-to-date maps. On the other hand, space-borne measurements have a lower spatial resolution but a higher temporal resolution and global coverage. They have the potential to map forest properties at a global scale monthly or yearly like the product developed by Hansen et al. (2013), which derived global forest annual loss and probability of gain at 30 m resolution from 2000 to 2012 from Landsat imagery. Additionally, spaceborne LiDARs such as ICESat have demonstrated the potential to map forest height (El Hajj et al., 2019; Fayad et al., 2014; Pourrahmati et al., 2015). Global and local maps of forest biomass and height have been developed based on various remote sensing approaches, often with worldwide coverage and a medium to low resolution (~100 to 1000 m). They are often based on spaceborne or airborne observations, calibrated with in-situ measurements.

The GEDI (Global Ecosystem Dynamics Investigation) LiDAR mission, developed and operated by NASA onboard the International Space Station (ISS) since 2019, has produced accurate point-wise observations of forest structure (Dubayah et al., 2020). Combined with other space-borne and airborne data, this instrument has shown promising capabilities for height mapping (Fayad et al., 2021a; Lang et al., 2022, 2023; Potapov et al., 2021). Sentinel-1 (S1) and Sentinel-2 (S2) are two satellite missions of ESA's Copernicus program for Earth observation. They provide measurements of Earth's radar (Synthetic Aperture Radar) backscattering coefficients (Sentinel-1) or multi-spectral reflectance (Sentinel-2) at a 10 m resolution with a revisit interval of approximately five days since 2015 and have already been used to estimate biomass and canopy height at high resolution (Lang et al., 2019; Li et al., 2020; Morin et al., 2019). GEDI and Sentinel-2 combinations have been proposed for mapping crop height (Tommaso et al., 2021) and forest canopy height (Lang et al., 2023), and its potential to estimate forest height was evaluated by Pereira-Pires et al. (2021) with linear and exponential regressions.

Machine learning has proven some solid results for forest parameter estimations in previous remote sensing studies (Fayad et al., 2014; Morin et al., 2019; Potapov et al., 2021). More recently, deep learning and, more specifically, convolutional neural networks (CNN) have provided a new set of tools allowing remote sensing research to process large amounts of training data for more accurate predictions (Ball et al., 2017; Zhu et al., 2017). CNNs (LeCun et al., 2015) have significantly increased accuracy in image interpretation tasks. Thanks to a series of linear operations (convolutions) and non-linear "activation" functions, these models can learn multi-scale image features, including image texture, which are then used to carry out predictions. CNNs are already widely used for object detection or scene classification tasks and have proven their efficiency in remote sensing (Zhu et al., 2017). However, few studies have used these models for regression tasks like tree height mapping (Dalagnol et al., 2022; Illarionova et al., 2022), and, to our knowledge, no studies have used simultaneously GEDI, Sentinel-1, Sentinel-2 and a CNN model to estimate canopy height up to now.

Here, we introduce a new methodology that leverages the potential of GEDI to be used as reference height to train a deep-learning model. The methods combine GEDI's height pointwise measurements and S1 and S2 images at a high temporal and spatial resolution to create wall-to-wall height maps at 10 m scale, based only on space-borne data, over a large forest area in France. The region analyzed is the largest western European plantation forest: the *Landes de Gascogne* (referred to as Landes forest in the following), a maritime pine plantation located in the southwest of France. Our model is trained on seven combinations of S1 and S2 layers. The retrieval results are evaluated using several in-situ datasets: a dense maritime pine inventory performed in 2016 (GLO-RIE), the French National Forest Inventory (NFI) distributed all over the study area, and a height map based on 3D stereo height reconstruction from Skysat imagery at a specific location within the Landes forest.

Additionally, the model is compared to three canopy height maps available in the area of interest.

## 2. Material and methods

This study relies on the use of canopy height measured by the GEDI sensor, a space-borne LiDAR onboard the ISS (Dubayah et al., 2020). We used GEDI's RH<sub>95</sub> variable defined in 2.2.1 as reference height samples and Sentinel-1 and Sentinel-2 images (2.2.2 and 2.2.3) as predictors for a deep learning U-Net framework (2.4.1) to generate a gridded map of the *Landes de Gascogne* area at a resolution of 10 m. Fig. 1 describes this general workflow.

### 2.1. Study area

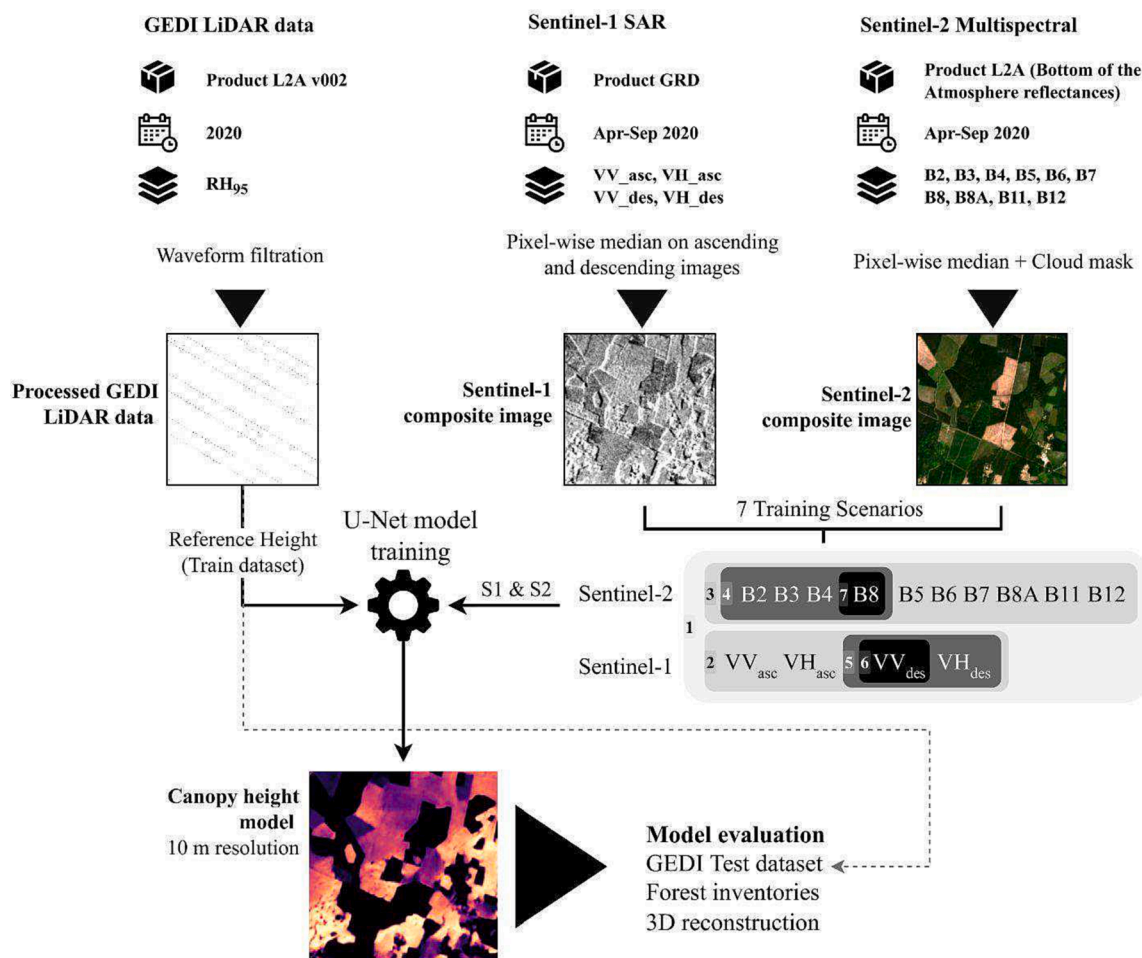
The *Landes* forest is located in a flat region with an oceanic climate and sandy soils in the South-West of France (Fig. 2). It is the largest European plantation (~1 million hectares) composed of 90 % of maritime pine (*Pinus Pinaster*). The remaining forested part consists of broadleaved forests, including several species of oaks, mainly located around rivers. This forest is intensively managed: thinning occurs every 5–10 years, clear-cuts are performed after 35–50 years, and they are replanted within 2–3 years. Forest management leads to very homogeneous tree repartition within stands but also to high heterogeneity between separate forest parcels. Understorey vegetation comprises various woody shrubs and perennial herbaceous plants (fern). It is regularly cleared during the first years following reforestation.

### 2.2. Datasets used to train the model

#### 2.2.1. GEDI

Data from the Global Ecosystem Dynamics Investigation (GEDI) are used as a reference variable to model a continuous height map at 10 m resolution. GEDI is operated by NASA and produces high-resolution LiDAR observations of the Earth's vertical structure (Dubayah et al., 2020). This spaceborne infrared LiDAR, deployed in December 2018 on the ISS, provides energy return waveforms (L1B product) and derived metrics such as canopy relative height (RH) and plant area index (L2A/L2B product) that describe the vertical forest structure within 25 m diameter circular footprints. The instrument acquires data over eight tracks with a footprint spacing of 60 m along the track and 600 m across tracks. Due to the ISS orbit, it covers latitudes between 51.6° South and 51.6° North. The precision of the horizontal footprint location, initially around 20 m, has been improved in GEDI's second release to a value of 10 m, as shown in Appendix 1 (Dubayah et al., 2021).

The GEDI L2A product provides RH metrics representing the height relative to the ground of n% (RH<sub>n</sub> with n ranging from the 0 to 100) of the total returned energy between the top of the canopy and the signal end. These metrics are extracted from the raw waveforms with six different algorithms representing different combinations of thresholds and smoothing settings. They can vary from one algorithm to another depending on the forest type (Adam et al., 2020). In our study area, most footprints had the same RH values for most algorithms. Hence, we chose to use the algorithm selected by NASA in the "selected\_algorithm" variable in GEDI data, which corresponds mainly to algorithm 1 over our study area. In theory, RH<sub>100</sub> should represent the top of canopy height. Still, this metric is affected by noise from atmospheric disturbances, uncertainties on the position of the detected ground return, vegetation, and ground variability. Here, we used RH<sub>95</sub>, as it has showed a better correlation with other height sources and has been used as a proxy for height in previous studies (Fayad et al., 2021b; Potapov et al., 2021) even though other similar metrics such as RH<sub>98</sub> have also been used (Lang et al., 2022). Due to the LiDAR properties, this measure is intended for vegetation and may represent confusing results for bare soil or water bodies (Beck et al., 2020). Indeed, these surfaces mirror the transmitted waveforms that have a pulse width of ~ 15 ns which



**Fig. 1.** General workflow showing the data preprocessing steps, the U-Net model training, and the evaluation strategy. More details for the U-Net training process are presented in Fig. 5. The seven training scenarios correspond to the layers from S1 (VV<sub>asc</sub>: Vertical-Vertical polarization for ascending orbits, VH<sub>asc</sub>: Vertical-Horizontal polarization for ascending orbits, VV<sub>des</sub>: Vertical-Vertical polarization for descending orbits, VH<sub>des</sub>: Vertical-Horizontal polarization for descending orbits) and S2 (B2: Blue, B3: Green, B4: Red, B5-B6-B7: Red edge, B8: Near Infrared (NIR), B8A: "narrow" NIR, B11-B12: Short Wave Infrared (SWIR)) used in the training process: 1 - all layers; 2 - all S1; 3 - all S2; 4 - S2 RGB + NIR, 5 - S1 VV<sub>des</sub> + VH<sub>des</sub>; 6 - VV<sub>des</sub>; 7 - S2 NIR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

corresponds to a ~ 4.5 m wide waveform and hence to a RH<sub>100</sub> value of 2.25 m (Dubayah et al., 2020).

In total, 526,449 footprints from the GEDIv002 L2A product (Dubayah et al., 2021) were downloaded from NASA’s EarthDataSearch website (<https://search.earthdata.nasa.gov/search>) for this study, covering the entire area of interest for 2020. Due to atmospheric perturbations, some waveforms could not be used to give information on the vertical forest structure. Therefore, several filtering criteria were applied to remove unusable waveforms: (1) When the *quality\_flag* provided in the GEDI data was set to zero. (2) When one of *toploc*, *botloc*, *num\_detectedmodes*, *RH<sub>100</sub>* provided metrics had a null value. (3) When the ratio defined as the maximum amplitude of the waveform divided by the standard deviation of noise was lower than 30 (arbitrary value that removed 3 % of the remaining waveforms). As the study area is mostly flat, terrain was not considered in the data filtering. Moreover, recent studies (Fayad et al., 2021c) showed that GEDI canopy height retrievals were not affected by small slopes (<45 %). After these operations, 175,511 GEDI valid waveforms (33 %) were kept in 2020.

These GEDI footprints were then spatially separated into train, validation, and test datasets. For this, the study site was divided into 117 tiles of 100 km<sup>2</sup> and randomly separated into 91 train tiles, 15 validation tiles, and 11 test tiles so that it corresponds respectively to 75 %, 15 %, and 10 % of the GEDI footprints (Fig. 3a). The temporal distribution of the data is uneven and some months (July, ~ 16 % of the data) are more

represented than others (January, ~ 4 % of the data). The spatial distribution of the data is also uneven due to the ISS trajectory and the northern tiles of the study area contain more footprints than other tiles. The spatial and temporal distributions of these data are shown in Appendix 5. To evaluate our model on forested areas only, in test and validation tiles, we removed the GEDI footprints located in pixels where the tree cover has a null value in the Copernicus tree cover density map (Copernicus Land Monitoring Service, 2018). For the train dataset, we kept all data, thus leading to more realistic results over non-forest areas. The height repartition of data in train, validation, and test datasets is presented in Fig. 3b.

### 2.2.2. Sentinel-1

Sentinel-1 (S1) is a C-band Synthetic Aperture Radar (SAR) mission composed of 2 satellites with a sun-synchronous orbit and 12 days repeat cycle (Sentinel-1A, Sentinel-1B) launched respectively in 2014 and 2016 by ESA. Here, we used the Ground Range Detected (GRD) scenes with dual-band cross-polarization (Vertical-Vertical + Vertical-Horizontal bands at 10 m resolution). They were preprocessed with the Sentinel-1 toolbox in Google Earth Engine, which includes thermal noise removal, radiometric calibration, and terrain orthorectification as specified at <https://developers.google.com/earth-engine/guides/sentinel1>. The S1 data set provides backscattering coefficients (dB) which are a measure of the backscattered microwave radiations



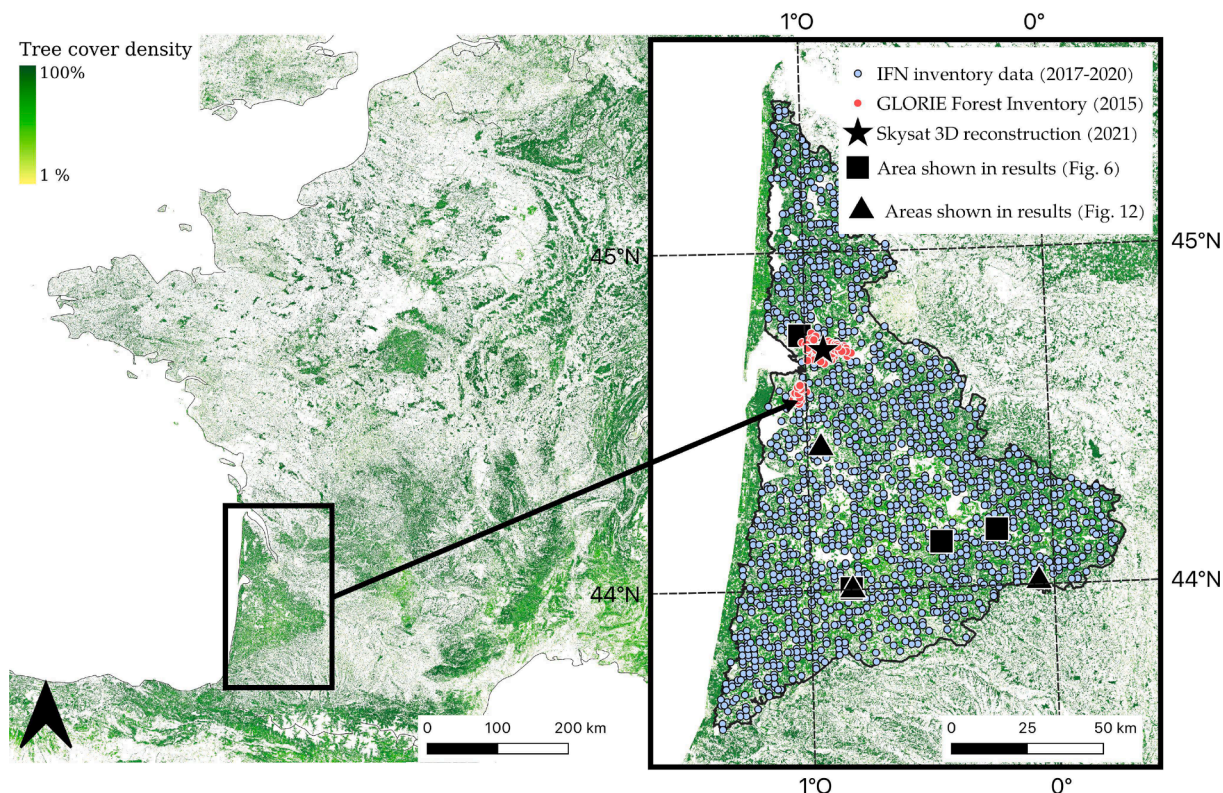


Fig. 2. Study site, known as “Les Landes de Gascogne” (referred to as Landes forest). Forest inventory plots, GLORIE inventory, Skysat 3D reconstruction, and forest areas considered in the results are shown. Greenness indicates tree cover density (Copernicus Land Monitoring Service, 2018).

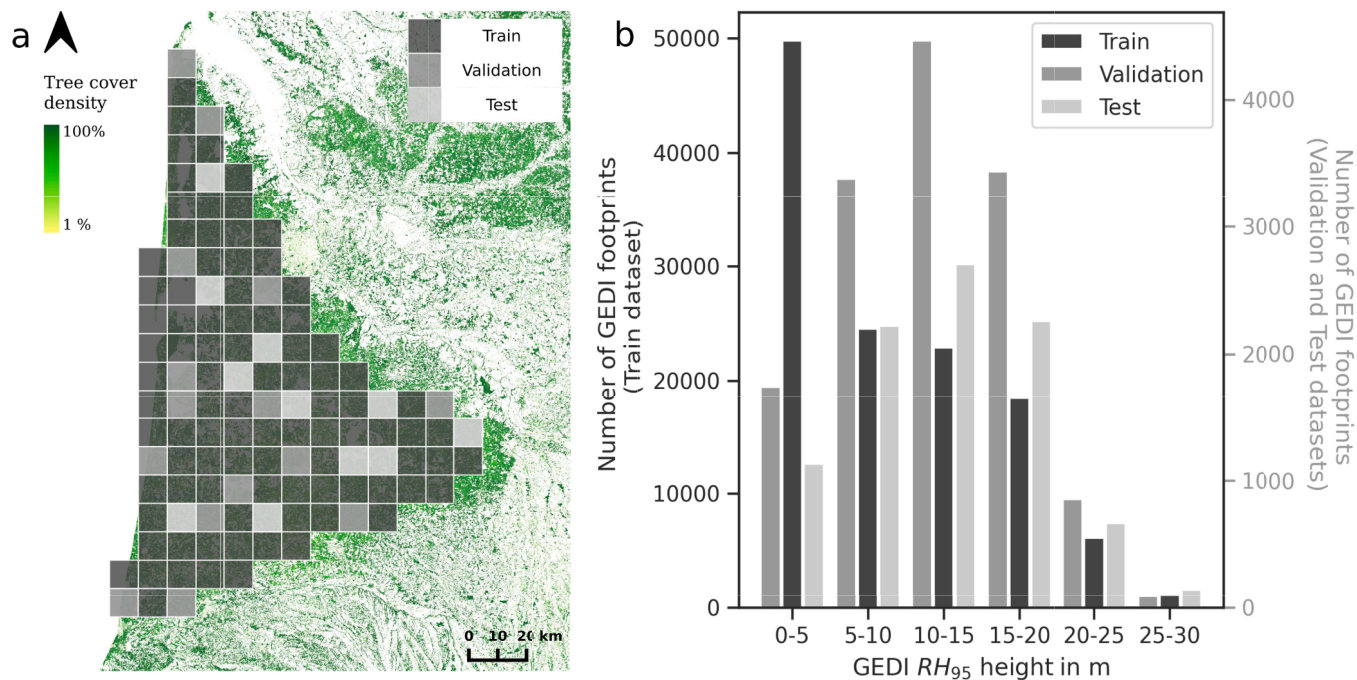


Fig. 3. (a) Repartition of the area of interest into 100 km<sup>2</sup> tiles divided into train, validation, and test datasets. (b) Count of GEDI waveforms per height in train, validation, and test datasets. Gedi footprints from non-forest areas were removed from validation and test datasets.

(backward direction) emitted by the radar system.

Using the Google Earth Engine (GEE) online platform, we selected 280 images covering the entire area of interest over a five-month period in 2020 (2020-05-01 to 2020-10-01). This period was selected to have

enough data to reduce the speckle observed in raw S1 images but still not too large to keep a similar aspect of the forest canopy (leaf-on season) to calculate a median composite. We then separated them between ascending and descending orbits (140 images per category) and

calculated the per-pixel median of these image time series, thus creating a single composite image of 4 layers: VV\_ascending, VH\_ascending, VV\_descending, and VH\_descending at 10 m resolution. Each pixel of the composite image corresponds to the median value of pixels from  $\sim 25$  images at different dates. The median is little influenced by extreme values and reduces the potential soil and vegetation moisture-related effect that could be observed on raw S1 images. These bands were then restrained to values between  $-30$  dB and  $0$  dB and scaled to a  $0$  to  $1$  interval in order to have a similar range of values as the Sentinel-2 input data (See 2.2.3).

### 2.2.3. Sentinel-2

The Sentinel-2 (S2) mission is part of the European Spatial Agency's (ESA) Copernicus program for Earth Observation (EO). It is composed of two satellites, Sentinel-2A, launched in 2015, and Sentinel-2B, launched in 2017, on a sun-synchronous orbit. S2 provides multi-spectral images of the Earth surface reflectance with a low revisit interval of  $\sim 5$  days, including images from both satellites, thanks to a large swath of  $290$  km. The L2A product used here provides bottom-of-the-atmosphere reflectance, processed from L1C (Top-of-the-atmosphere reflectance) with ESA's Sen2Cor processor (Main-Knorn et al., 2017). It comprises 13 spectral bands ranging from  $10$  m to  $60$  m resolution in visible, near-infrared (NIR), and short-wave infrared (SWIR).

Based on the Google Earth Engine (GEE) online platform, we selected 172 images within the entire area of interest during the same period as the S1 images (2020-05-01 to 2020-10-01) with less than  $50\%$  of clouds. After applying a cloud mask, we created a single composite image by taking the per-pixel median value over the image time series. We chose the median composite to be less sensitive to outlier pixels such as pixels affected by clouds or cloud shadows that could remain even after the cloud mask. The median value computed for each pixel comes from 16 to 72 images depending on the pixel location because of differences in image overlap. The ten following spectral bands ( $10$  and  $20$  m resolution) were kept and resampled to  $10$  m if necessary: B2: Blue, B3: Green, B4: Red, B5-B6-B7: Red edge, B8: Near Infrared (NIR), B8A: "narrow" NIR, B11-B12: Short Wave Infrared (SWIR). We did not calculate any hand-crafted features such as vegetation indices or texture features insofar as we used a deep learning network (See 2.4.1) which calculates its own features from the raw images thanks to convolutional filters. This approach is very flexible and encompasses non-linear combinations of features derived from raw satellite data that may be less rigid than setting fixed indexes. However, adding hand-crafted features to the prediction layers may be tested in further research to evaluate their potential to improve the model's accuracy. The input data of a neural network should be on a similar scale to help stabilize the gradient descent step in the training phase. Thus, it is important to have the same range of values for the different inputs of our model (S1 and S2). Sentinel-2 reflectance values in GEE are given in digital units (DN) between  $0$  and  $10,000$  ([https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S2\\_SR](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR)) where  $DN = 10000 * \text{Reflectance}$ . Forest reflectance typically ranges from  $0$  to  $0.3$  (DN from  $0$  to  $3000$ ) depending on the spectral bands. To have a better contrast and a more suited distribution for deep learning without a long tail, we clipped the DN values in the  $0 \sim 5000$  interval (values higher than  $5000$  were set to  $5000$ ). Then we divided all the values by  $5000$ , creating a range of values from  $0$  to  $1$ . Thanks to filtering, cloud mask, and median average, the resulting composite image was exempt from cloudy areas.

### 2.3. Evaluation datasets

To compare our height predictions with independent sources, we used three complementary datasets covering different spatial scales: a very high-resolution 3D stereo reconstruction from Skysat imagery, a dense local forest inventory (GLORIE) performed near the Bordeaux area, and the French national forest inventory data from IGN (Institut national de l'information géographique et forestière) spanning the

entire area of interest (Fig. 2). Additionally, we compared our predictions to canopy height maps computed from previous local and global studies using alternative techniques or data sources (Lang et al., 2023; Morin et al., 2019; Potapov et al., 2021).

#### 2.3.1. Forest inventory data

The French national forest inventory (NFI), performed yearly by the French Geographical Institute (IGN, Institut Géographique National), carries out forest measurement campaigns in France (IGN, 2022). Every year,  $30$  m diameter circular plots are sampled in one-tenth of a grid that covers the national territory. These forest plots are distributed in the train, validation, and test tiles (see 2.2.1). The dominant height variable provided by IGN for these forest plots corresponds to the estimated mean height of the 100 highest trees within a surface area of  $1$  ha. In order to avoid a too-long time difference between the date of inventory and the date of our prediction, we selected only plots that were measured between 2017 and 2021. To compare these heights to our predicted heights, we used the mean value of our  $10$  m x  $10$  m pixels that intersect the  $30$  m circular plots of the NFI, which corresponds to the mean of  $\sim 9$  pixels. We also tested other comparison methods, e.g., taking the maximum value of these pixels and taking the single  $10$  m x  $10$  m pixel value at the center of the plot. These different comparison methods yielded very similar results and we used the mean value for this study.

The GLORIE local forest inventory was acquired in the winter of 2015–2016 before the beginning of tree growth (Motte et al., 2016; Zribi et al., 2019). It includes measurements of, among others, tree heights, DBH, and tree density. The 99 forest plots (location indicated in Fig. 2) were selected from forest stands of maritime pines, trying to cover the entire range of forest structures in this region. The GLORIE forest plots are located within the train tiles (see 2.2.1). The dominant height variable provided in this dataset was calculated as the mean height of the two largest trees (higher diameter at  $1.30$  m) among the ten trees closer to the center of the forest plot. Here, we computed the mean value of the height pixels predicted by our model within a  $25$  m diameter circle (typical size of these forest plots) around the location of the forest inventory plots and compared it to the dominant height measured in the inventories.

Although forest inventory canopy height data were considered as ground truth to evaluate our model, uncertainties are associated with them. Forest inventory accuracy on height measurements strongly relies on the surveyors and the method used (Berger et al., 2014; Jurjević et al., 2020; Kitahara et al., 2010). Internal IGN studies based on forest inventory control, re-measuring  $\sim 4000$  trees each year, have shown that the typical standard error on tree height estimation ranges from  $1.4$  m to  $1.75$  m.

#### 2.3.2. Stereo 3D reconstruction from Skysat imagery

Skysat is a constellation of optical sub-meter resolution EO satellites owned by Planet Labs, providing high-resolution imagery in panchromatic, visible, and infrared bands at  $0.9$  m resolution. Its geolocation accuracy is  $3.4$  m for all images (Saunier et al., 2021), but it should be better on the cloud-free images used here. The stereo 3D height reconstruction technique of de Franchis et al. (2014) uses multiple image acquisitions with different view angles to reconstruct 3D objects with a dispersion error of  $\sim 0.5$  m. Here, we applied this technique to reconstruct canopy height in a small area of  $11$  km<sup>2</sup> (Location indicated in Fig. 2) and compared it to our model outputs. We first created a digital surface model (DSM) at  $0.8$  m resolution from point-cloud 3D reconstruction based on Skysat images acquired in 2021. Then we used a cloth simulation algorithm (Zhang et al., 2016) to select ground points from the 3D point cloud. Finally, we used a Laplace interpolation to create a terrain model from those points that we subtracted from the raw DSM to obtain a canopy height model (CHM). To compare it to our model, we resampled this CHM to a  $10$  m grid aligned with our prediction map. For this, we took the maximum value of the CHM within each  $10$  m x  $10$  m grid cell, which corresponds to the top canopy height and is thus



relevant for a comparison with our RH<sub>95</sub>-based model. As forest stands in the Landes region are very homogeneous, we carried out two types of comparison with our model: (1) a simple per-pixel comparison of the two maps. 2 A comparison at the forest stand level by calculating the median value over manually delimited forest stands (Fig. 4).

### 2.3.3. Canopy height maps from previous studies

We compared our model predictions to three canopy height maps from independent studies available in the Landes forest (Lang et al., 2023; Morin et al., 2019; Potapov et al., 2021). Maps from Potapov et al. (2021) and Lang et al. (2023) are two global canopy height maps, hence non-specific to the study area, based on GEDI and optical data only. Potapov et al. (2021) have mapped forest height globally in 2019 at 30 m resolution, using Landsat-8 images to extrapolate canopy height from GEDI footprints with a bagged regression trees ensemble method. Lang et al. (2023) have used Sentinel-2 images to extrapolate canopy height from GEDI footprints with a deep fully convolutional network and create a global canopy height map for 2020 at 10 m resolution. The map from Morin et al. (2019) is a local canopy height product specific to maritime pine homogeneous plantations in the Landes forest. Contrary to the two previous maps that used GEDI, reference height measurements were based on a small local forest inventory of maritime pine only (GLORIE inventory, see 2.3.1). These forest inventory measurements were used with remote sensing data (Sentinel-2, Sentinel-1, ALOS-PALSAR...) and a machine learning model (Support vector machine) to produce a canopy height map of the Landes forest in 2016, only valid for maritime pine plantations.

## 2.4. Methodology

### 2.4.1. U-net model description

In this study, we performed a pixel-wise regression, which is the process of attributing a particular value (tree height) to each pixel instead of attributing a label to the whole image. This task can be addressed by fully convolutional networks (FCNs) (Long et al., 2015) that have been adapted from classical CNNs. Here, we use a U-Net model adapted from Milesi (2022) which is a FCN that outperformed previous models in speed and accuracy and requires fewer training examples thanks to its U-shape architecture (Ronneberger et al., 2015). This model consists of a contracting path (left) and an expansive path (right) which gives it its “U” shape and enables the model to extract relevant information at different spatial scales. In the contracting path, the input

image goes through two 3x3 convolutions followed by a rectified linear unit (ReLU) and is then downsampled with a 2x2 max pooling operation with stride 2. This is repeated four times, and at each step, the number of feature channels is multiplied by two. In the expansive path, the image first goes through a bilinear upsampling, then it is concatenated with the corresponding image of the contracting path and finally goes through two 3x3 convolutions followed by a ReLU like in the contracting path. This step is also repeated four times and at each step, the number of feature channels is divided by four instead of two because of the concatenated image. The final layer consists of a 1x1 convolution that creates an output image with the same height and width as the input image. After the training process, this output will represent the canopy height map derived from the input image. In total the network has 18 convolutional layers and ~ 17 Million trainable weights. The exact architecture of the prediction model we developed can be found in Appendix 2 (it will be referred to as the “FCN model” in the following).

### 2.4.2. Training process

The objective of the training process was to adjust all the trainable weights of the FCN model so that when the FCN model takes a multi-channel image composed of S1 and S2 layers as input, it outputs an estimated canopy height map. One single FCN model was trained on all the train tiles. To achieve this, we followed the process described in Fig. 5: (1) We randomly selected one of the 91 train tiles. This random selection was weighted by the number of GEDI footprints in each tile. (2) We took a random 2560x2560 m subset from this tile with at least one GEDI footprint inside. This technique increases the number of different images used as input of the model and contributes to reducing model overfitting (3) The corresponding 256x256 pixels image composed of S1 and S2 layers was used as input of the U-Net. In the case where we chose to use all S1 and S2 bands, this image was composed of 14 layers. (4) The GEDI RH<sub>95</sub> values in this 2560x2560 m subset were rasterized on a 10 m grid aligned with S1 and S2. For each GEDI footprint, the RH<sub>95</sub> value was rasterized on the pixel where the center of the circular footprint was located. As the GEDI footprints correspond to 25 m diameter circles with a 10 m uncertainty on their location, we acknowledge that our rasterization process at 10 m could create label noise, especially in forests with a very heterogeneous canopy. Still, training at 10 m was shown to give better results than training at 20 m (see Appendix 7). (5) The FCN model output was compared to the reference height from the rasterized GEDI data with a mean absolute error (MAE) loss only for pixels where a RH<sub>95</sub> height value was available. (6) The gradient of this loss was calculated

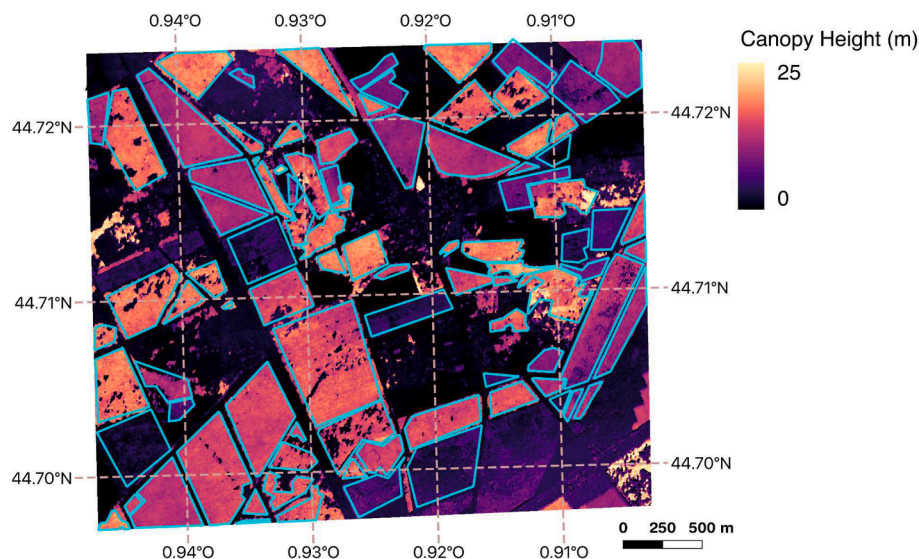
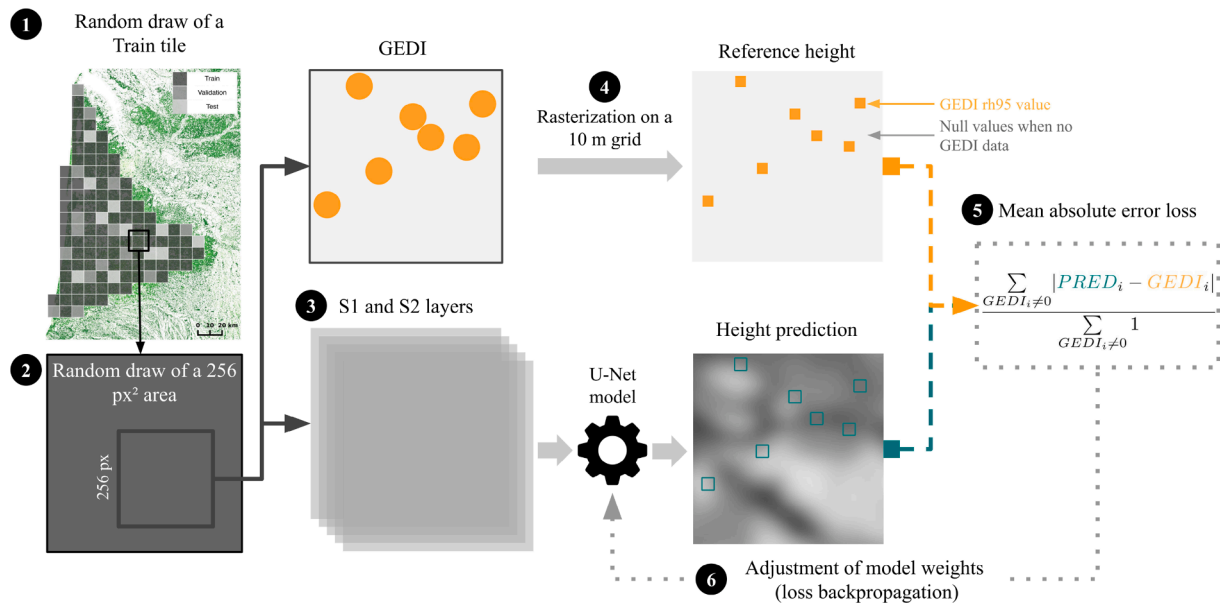


Fig. 4. Stereo 3D reconstruction from Skysat imagery (0.8 m resolution). Forest stands estimated from manual delimitation are represented by blue polygons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Computation of the mean absolute error (MAE) loss in the FCN prediction model training process. The numbers in the black circles correspond to the steps described in section 2.4.2 (training process).

with respect to each of the model weights and weights were then adjusted accordingly. This process is called “loss backpropagation” and is a key element in the training of neural networks. Here it was performed with a SGD optimizer (momentum of 0.9) with a cyclic learning rate scheduler “triangular2” (Smith, 2017). The base value was set to  $1.10^{-7}$ , the max value to 0.1, with 320 steps for a half cycle. After each epoch of 960 images (32 batches of 30 images), we calculated the MAE loss on all the validation tiles. After  $\sim 100$  epochs, which corresponded to  $\sim 2$  h, this validation loss was stabilized and we manually stopped the training process. We also tried other types of learning rate schedulers but this one led to faster convergence and similar results. The training process was done with Amazon AWS cloud platform on a GPU NVIDIA Tesla T4 (16 GB). We used the Pytorch library, an open-source machine learning framework in Python, to implement the FCN model.

#### 2.4.3. Training scenarios

To understand the importance of each spectral (Sentinel-2) and polarization (Sentinel-1) band, we designed seven FCN models that followed exactly the same training process but with seven different combinations of S2 and S1 bands as input. The first FCN model was trained on all layers (Scenario 1: 10 from S2 and 4 from S1). The second and third FCN models were based on one source of data only (either S2 or S1). The other FCN models were trained on subsets of S2 (only 10 m resolution bands or only B8 (NIR band)) and S1 bands (only descending orbit, only VV polarization for descending orbit) respectively (See Fig. 1 for a summary of the 7 scenarios used).

#### 2.4.4. Used metrics

To evaluate our FCN model against the validation datasets, we used several metrics, namely the mean absolute error (MAE), the root mean squared error (RMSE), the mean error (ME) that indicates the bias, and the coefficient of determination ( $R^2$ ) that is a measure of the correlation between predicted height and true height values. Additionally, in order to have more information on the source of errors, we decomposed the mean squared error (MSE) into three additive terms: squared bias (SB), squared difference between standard deviations (SDSD), and lack of correlation weighted by the standard deviations (LCS) as proposed by (Kobayashi and Salam, 2000) where  $MSD = SB + SDSD + LCS$ . SB indicates the bias, SDSD shows how the FCN model is able to simulate the magnitude of the fluctuation between the  $n$  measurements and LCS is the

ability of the FCN model to simulate the fluctuations across the  $n$  measurements. The detailed formulas for these metrics are provided in Appendix 3.

### 3. Results

#### 3.1. Canopy height model

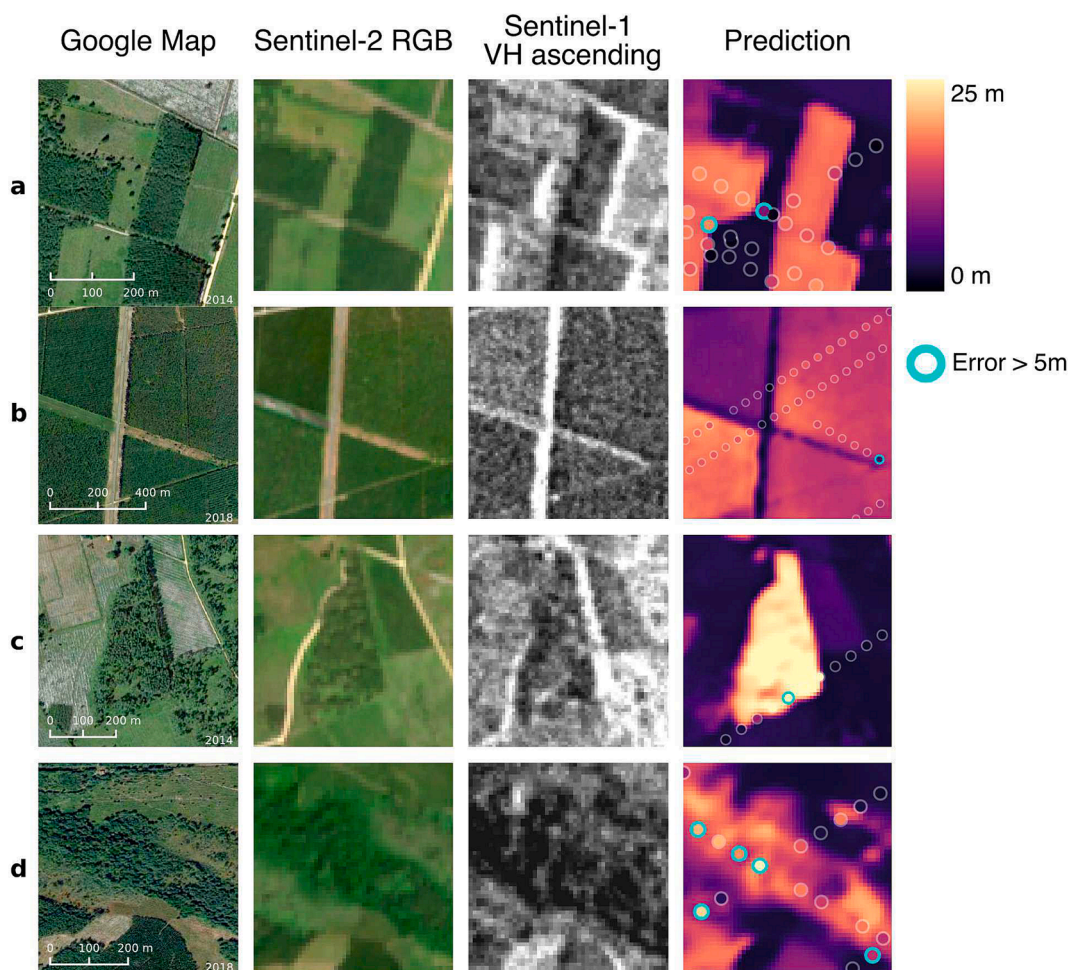
Here, we present the results for Scenario 1 where the FCN model was trained using all Sentinel-1 and Sentinel-2 layers. In Section 3.4, we present the results for Scenarios 2–7, where different subsets of the layers were used; all of which performed less well than Scenario 1. Fig. 6 illustrates four examples of predictions for different types of forested areas, while Fig. 7 shows our prediction map for the whole Landes forest in 2020.

Overall, even though our FCN prediction model was trained on sparse reference data (eg. rasterized GEDI RH<sub>95</sub> values in the train tiles), it is still capable of producing a continuous canopy height map, where forest structures and other landscape features are recognizable and look similar to the image input. Fig. 6a shows that the FCN model is able to capture high height differences between forest and non-forested areas. When compared to the GEDI test dataset for maritime pine plantations (Fig. 6a, b, c), most errors are below 5 m except in areas close to forest borders. The within-stand homogeneity and across-stand heterogeneity are well captured by the FCN model over coniferous forest stands with different heights (Fig. 6b). Fig. 6d shows an example obtained in an area including deciduous trees along a river path. There is a higher number of differences greater than 5 m when comparing to the GEDI test footprints and the predicted forest delimitations are less precise.

#### 3.2. GEDI test set evaluation

We compared the predictions of our FCN model (Scenario 1) to the RH<sub>95</sub> values for the test dataset. We obtained a mean absolute error (MAE) of 2.02 m, RMSE = 2.98 m, and  $R^2 = 0.73$  (Fig. 8). The FCN model tends to slightly underestimate higher heights (ME =  $-2.5$  m for trees between 20 m and 25 m) and predictions for lower heights start around 2.5 m due to the RH<sub>95</sub> properties (see 2.2.1). Predictions are relatively well scattered around the  $y = x$  axis with the exception of a cluster of points located on the left side of the graph where the





**Fig. 6.** Model input data and model predictions on four different areas in the test tiles. The first column shows Goggle Maps images for reference, not used by the FCN prediction model. The two middle columns show some of the input bands to the FCN prediction model. The last column shows our predicted map of forest height where GEDI height values ( $RH_{95}$ ) can be identified by circles. When the predicted and the GEDI heights are different by at least 5 m, the GEDI footprints are shown with cyan circles. (a) and (b): Maritime pine forest stands of different heights. (c) old forest stands of maritime pine. (d) deciduous forest. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

predictions overestimated the reference height values. This discrepancy is likely due to GEDI footprint location errors (See 2.2.1). Boxplots in Fig. 8a reveal almost no bias for height values between 5 m and 20 m ( $ME = -0.36$  m,  $MAE = 1.80$  m) which accounts for 79 % of the GEDI footprints.

### 3.3. Evaluation with independent datasets

#### 3.3.1. Forest inventory data

The comparison with the GLORIE dataset (Fig. 9a), a dense forest inventory of homogeneous stands of maritime pine in Les Landes carried out in 2015–2016 shows a good correlation with our predicted height with a  $R^2$  coefficient of 0.93 ( $ME = 1.56$  m,  $MAE = 2.43$  m,  $RMSE = 2.84$  m). We observe a higher prediction bias for lower heights from 0 to 10 m. The NFI dataset (Fig. 9b), composed of more diverse inventory plots (broadleaves, coniferous, distributed all over the study area and not only with homogeneous forest stands, see Fig. 2), is also well correlated to our canopy height retrievals ( $R^2 = 0.71$ ,  $ME = -1.18$  m,  $MAE = 2.67$  m,  $RMSE = 3.55$  m), especially for coniferous forests ( $R^2 = 0.79$ ,  $RMSE = 3.09$  m). Broadleaved forests show a lower correlation and poorer error metrics:  $R^2 = 0.38$ ,  $RMSE = 5.74$  m. Points highlighted in red in Fig. 9 indicate clear-cuts (we checked it visually with Sentinel-2 images time series) and were not considered in the calculation of the metrics.

#### 3.3.2. 3D reconstruction from stereo satellite acquisition

The pixel distribution in height (Fig. 10a) of our FCN model follows globally the same pattern as the 3D reconstruction canopy height model (referred to as 3D CHM hereafter). The 3D CHM shows a higher number of low-height pixels between 5 m and 10 m, while the FCN model has a higher number of pixels for all heights above 4 m, especially for heights around 16 m. The comparison at forest stand level (Fig. 10b) reveals a very good correlation ( $R^2 = 0.90$ ,  $MAE = 1.17$  m) with no bias ( $ME = -0.03$  m) between the 3D CHM and the FCN predictions of forest heights. Visually, the 3D reconstruction (Fig. 10c) presents sharper delimitations between the different forest units compared to the FCN model (Fig. 10d). This is confirmed by the height profile (Fig. 10e) which reveals that the FCN model tends to smooth the transitions between forest units. Within forest stands, the height also seems to be averaged and more homogeneous with the FCN model compared to 3D CHM which shows more fluctuations.

#### 3.4. Influence of Sentinel-1 and Sentinel-2 bands on height predictions

The results presented above are all retrieved from our best training scenario (Scenario 1: based on all S1 and S2 layers). Here, we evaluate other combinations of the S1 and S2 layers as inputs to the FCN model to understand the influence of the S1 and S2 layers in the FCN retrievals of forest heights (Fig. 11). Overall, the two first scenarios (1: All bands, 2:

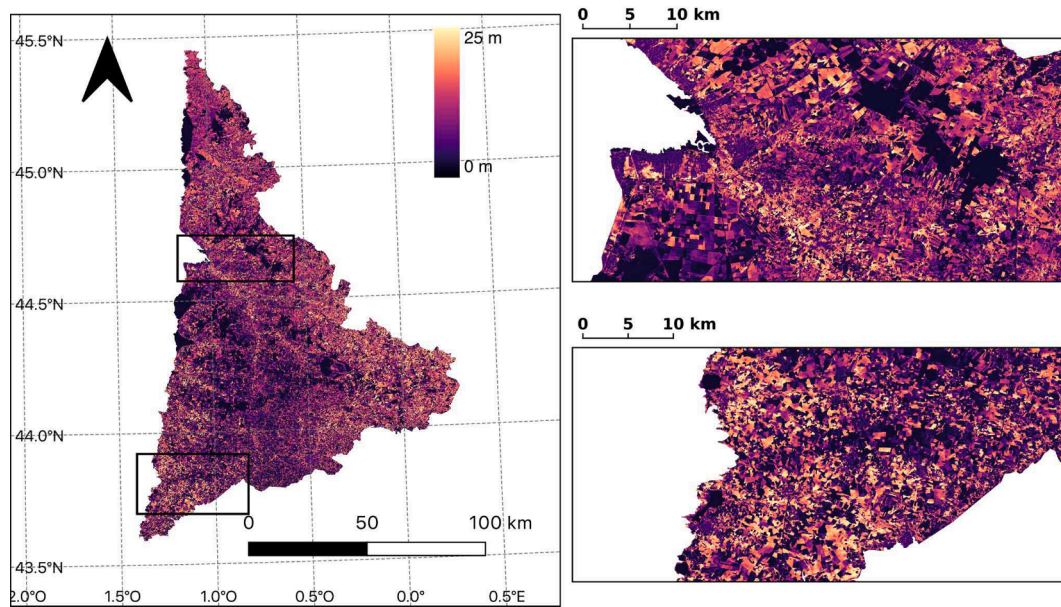


Fig. 7. Map of forest canopy height of the Landes forest for 2020, predicted by the FCN model from training Scenario 1: 10 layers from Sentinel-2 + 4 layers from Sentinel-1.

All Sentinel-1 bands) perform better, with a lower RMSE value when compared to any of the four evaluation datasets (e.g. 3.55 m and 3.76 m when compared to the French NFI while the results obtained for the other scenarios led to a  $RMSE > 4$  m). Scenarios including only Sentinel-2 bands have lower performances (Scenarios 3, 4, and 7) but still, lead to relatively good error metrics. Similar results are observed between Scenarios 3 (all S2 bands), 4 (only 10 m resolution i.e., 4 bands: RGB + NIR), and 5 (VV and VH descending from S1) except for the GLORIE dataset where Scenario 5 is better with no bias and lower SDS. Lastly, Scenarios 6 and 7, respectively based on one Sentinel-1 (VV descending) and one Sentinel-2 (B8: NIR) band lead to larger errors for all validation datasets ( $RMSE > 3$  m when compared to 3D CHM while other scenarios lead to a  $RMSE < 2.5$  m).

Overall, SB is higher when the FCN model is compared to forest inventory datasets (GLORIE and IFN). SDS indicates how the FCN model retrieves the magnitude of the spatial fluctuations and shows higher values only for the GLORIE dataset. LCS indicates the ability of the FCN model to simulate the fluctuations across the  $n$  measurements and is the dominant term of the error for most validation datasets. For all scenarios, the correlation is better when compared to the 3D Skysat height reconstruction model.

### 3.5. Comparison with other canopy height maps

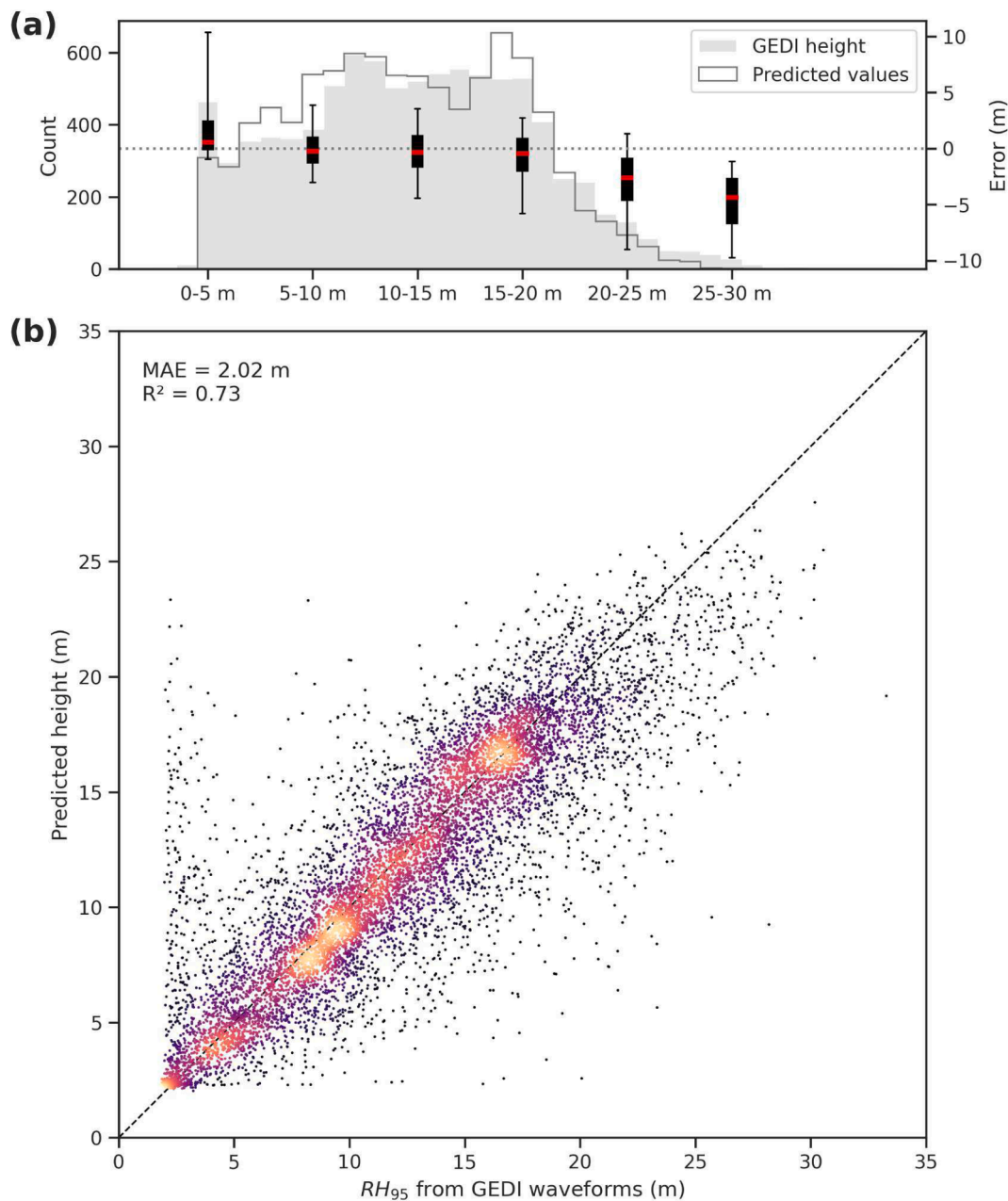
To evaluate our height predictions against existing models, we compared the predictions from the FCN model that performed best on the validation datasets (Scenario 1, trained with all S1 and S2 layers, see Sect. 3.4) with three different canopy height maps: L23 (Lang et al., 2023), P21 (Potapov et al., 2021), and M19 (Morin et al., 2019) at various locations in the test tiles.

As an illustration, a visual comparison (Fig. 12) reveals that the FCN and L23 models predict a higher homogeneity within the forest stands while the two other models (P21 and M19) show a higher variability between adjacent pixels. Fig. 12a shows two adjacent planted forests, surrounded by a non-forested area. The GEDI footprints in this region indicate (unshown results) that the forest in the southern part of the forest plot is higher ( $RH_{95} \sim 14$  m) than that in the northern part ( $RH_{95} \sim 10$  m). This height difference is well captured by the FCN model and by L23 to a lesser extent. On the contrary, P21 shows opposite height predictions and M19 does not seem to differentiate forest heights in the

southern and northern parts of the forest plot. Similarly, Fig. 12b shows an example of a more complex landscape structure with forest patches of various heights. In this case, all models are able to predict different heights for different forest stands. However, the FCN model produces more plausible outputs with clearer delimitations between forest stands of different heights. Over a broadleaved forest along a river path (Fig. 12c), L23 and P21 present higher predictions than the FCN model. M19 has not been trained on broadleaved forests, and therefore the comparison for this particular case is not relevant.

The comparison of the three height models (L23, P21 and M19) with the GLORIE forest inventory, the French NFI (separated into coniferous and broadleaved) and the Skysat 3D CHM (Fig. 13 and Fig. 14) reveals that the FCN model presents better error metrics except for broadleaved forests. The FCN model has a RMSE of 3.09 m when compared to the French NFI coniferous stands (Fig. 13b) while L23 ( $RMSE = 5.01$  m), P21 ( $RMSE = 6.57$  m), and M19 (6.75 m) show higher errors. However, L23 has better performances over broadleaved forests ( $RMSE = 5.26$  m,  $R^2 = 0.42$ , Fig. 13c) compared to the FCN model ( $RMSE = 5.74$ ,  $R^2 = 0.38$ ). For coniferous forests, the FCN model, along with M19 shows a much lower bias (almost null for the 3D CHM) than the two global canopy height maps (P21 and L23). L23 has a high SB value (high bias) when compared to the GLORIE forest inventory and to the Skysat 3D CHM but still has a good correlation with these datasets ( $R^2 = 0.81$  for GLORIE with a low LCS). This can also be visually seen in predictions in Fig. 12 and scatterplots in Fig. 14. All canopy height models have better performances on the Skysat 3D reconstruction (all RMSE values below 5 m, Fig. 13.d). P21 presents a saturation effect for coniferous forests at around 15–20 m for coniferous forests only (Fig. 14).

L23 shows a higher correlation ( $R^2 = 0.52$  for coniferous forests) than the FCN model (Fig. 15) but tends to predict higher heights ( $ME = 3.75$  m for coniferous forests and  $ME = 6.0$  m for broadleaved forests). The bias observed in the previous figures is also visible in Fig. 15, especially for lower heights. A lower correlation is obtained with M19 ( $R^2 = 0.28$  for coniferous forests only) but with almost no bias ( $ME = -0.3$  m) and the predictions seem to follow the same pattern but with a high variability around the  $x = y$  axis. The comparison with broadleaved forests for M19 is not relevant as M19 has not been trained on such types of forests. P21 predicts most coniferous forests at  $\sim 15$  m, most broadleaved forests at  $\sim 22$  m and most other pixels at  $\sim 5$  m.



**Fig. 8.** Comparison of the GEDI height ( $RH_{95}$ ) in the test dataset and predicted values from the FCN prediction model for Scenario 1. (a) Histogram with box plots that show the differences between predicted and GEDI height per height range of 5 m. The red lines represent median values. The upper and lower edges are the upper and lower quartiles and whiskers symbolize the 5th and 95th percentiles. (b) Density scatterplot. The dashed line corresponds to  $x = y$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

#### 4. Discussion

- Forest border errors

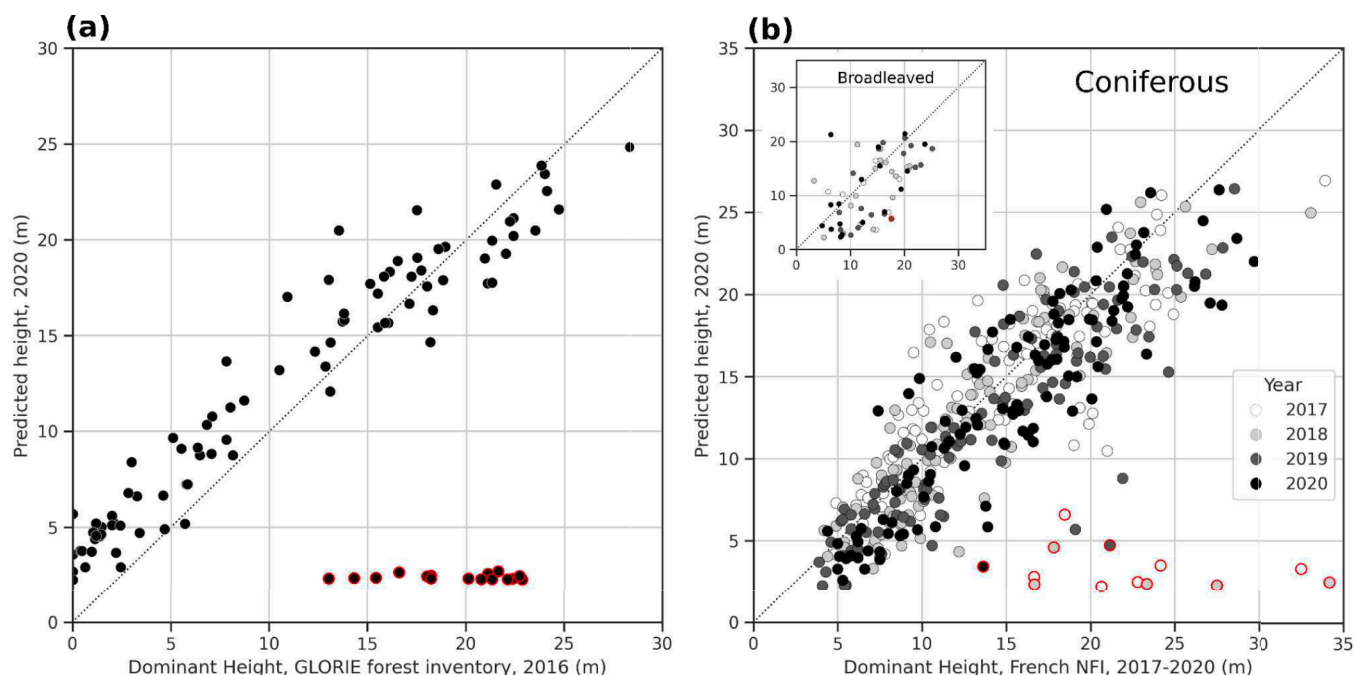
The visual analysis of the FCN height predictions shows higher errors at the forest borders (Fig. 6). These errors are likely related to the GEDI uncertainty of  $\sim 10$  m for the ground location. Some GEDI footprints are located within a forest, close to the border, but the waveform corresponds to the reflection on bare soil outside the forest. This effect may also explain the high FCN height predictions when the  $RH_{95}$  values are close to zero in Fig. 8. However, the opposite phenomenon of low predictions for high  $RH_{95}$  values is less frequently observed. This is likely because the GEDI test dataset we used only includes footprints located within forests. As the GEDI’s footprint diameter is 25 m, small gaps in

the canopy that are invisible in the S1 and S2 images cannot be reflected in GEDI’s  $RH_{95}$  and cannot explain these high prediction errors. To a lesser extent, clearcuts occurring between the date of the GEDI acquisition and the date of the S1 and S2 images could be responsible for some of these errors.

- Smoothness

At the border of landscape units, a smoothing effect seems to occur in the predictions. This can be particularly observed in the comparison with the 3D reconstruction from stereo Skysat images (Fig. 10c,d,e). Unlike per-pixel predictions, such as those performed in P21, each pixel of the FCN canopy height map is the result of multiple convolutions involving the surrounding pixels (See 2.4.1 for U-Net structure). Pixels





**Fig. 9.** Comparison between the predicted height from the FCN model (Scenario 1) for the year 2020 with the forest inventory dominant height from the GLORIE project (measurements made in 2016) (a), and with French National Forest Inventory (NFI, measurements made over 2017–2020) (b). For both graphs, the dotted line represents  $x = y$ . The predicted height corresponds to the mean height of the pixels within the forest plot area. The French NFI (b) dataset is colored by year of sample and separated into broadleaved and coniferous forests. Red circles indicate plots that are not considered in the calculation of the error metrics because of forest clear-cuts between the date of the forest inventories and the date of the S1 and S2 images. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

close to forest borders have neighbor pixels from forest and bare soil which could lead to this average height prediction, creating smoothness between landscapes units. Additionally, the 10 m x 10 m S1 and S2 pixels on forest borders contain also average information from both forest and non-forest areas and are also “smoothed” which can explain the smoothing effect in our predictions (Fig. 16). Finally, it is likely that the FCN model tends to predict average values at forest borders to avoid making high errors during the training process because of the GEDI location uncertainty.

- Bare soil

The  $RH_{95}$  properties result in bare soil being retrieved at a height of  $\sim 2.25$  m (See 2.2.1). As a consequence, the FCN canopy height model is not suitable for heights below this threshold. The height profile (Fig. 10. e) used to compare the 3D reconstruction with the FCN model highlights this phenomenon where we can see that low non-forest heights are overestimated. The overestimation of lower heights visible in Fig. 9a for the GLORIE dataset is mainly related to this  $RH_{95}$  effect and explains the higher SDS values in Fig. 11b. Indeed, this dataset contains several forest plots with tree height measurements close to zero (no forest in some cases or newly planted forest in other cases). Additionally, a tree growth effect may be present due to the four-year difference between the inventory (2016) and the prediction (2020), which could further accentuate this phenomenon.

- Underestimation at high forest heights

Our model underestimates heights above 20 m ( $ME = -2.5$  m for trees between 20 and 25 m, Fig. 8a). This phenomenon is common in most studies that try to predict forest height from machine learning algorithms but do not always occur at the same height (Lang et al., 2023, 2019; Morin et al., 2019; Potapov et al., 2021). Rather than a saturation effect related to the information contained in S1 and S2, it is more likely

due to an imbalanced distribution of reference height labels. Indeed, in the test dataset (only forests), GEDI  $RH_{95}$  values above 20 m account only for 8.9 % of the total number of footprints and 1.5 % for  $RH_{95} > 25$  m. Several techniques such as a weighted cost function or a different sampling strategy among the height labels can be used to reduce this effect and will be investigated in future developments of the FCN model.

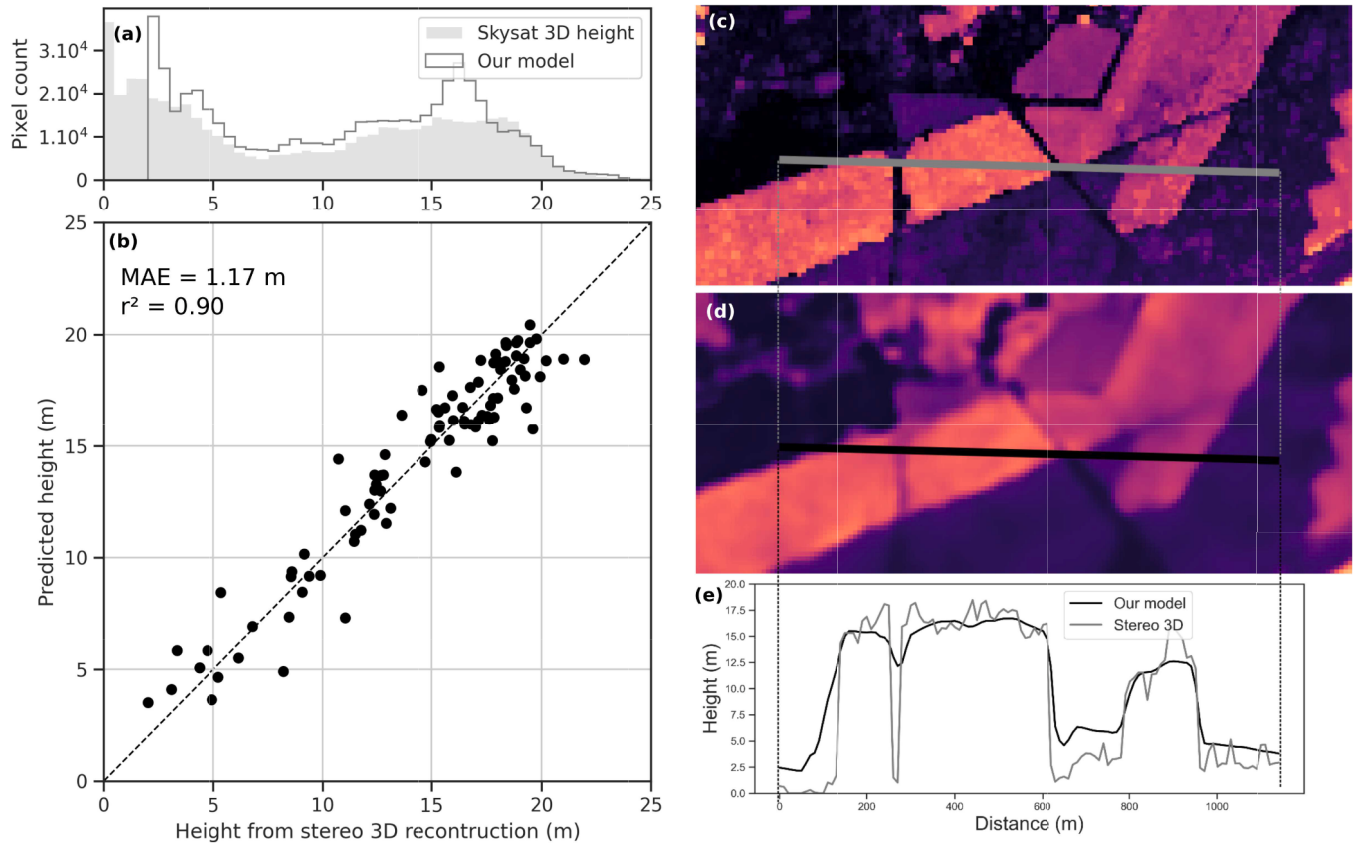
- Better performance over coniferous forests

Thanks to the French NFI data, we were able to evaluate separately the performance of the FCN model over broadleaved and coniferous forests. The error metrics are much better for coniferous forests for several reasons. First, as well as for higher heights, the number of GEDI footprints in broadleaved forests was much lower in our dataset than for maritime pine forests (only 10 % of the forested areas in the Landes forest is not covered by maritime pines). This lower number of reference height data for broadleaved forests could likely explain the lower performances of the FCN model on this type of forest. Additionally, broadleaved forests have a more complex structure, with a higher spatial heterogeneity. Within one 25 m GEDI footprint, several tree species, with different heights and shapes are summed up into one  $RH_{95}$  value. Therefore, an error on the GEDI footprint location has a higher impact on model training for this type of forest. Finally, this complexity also makes the on-site measurement of tree height more difficult which results in less precise validation data and therefore degraded performance metrics for broadleaved forests.

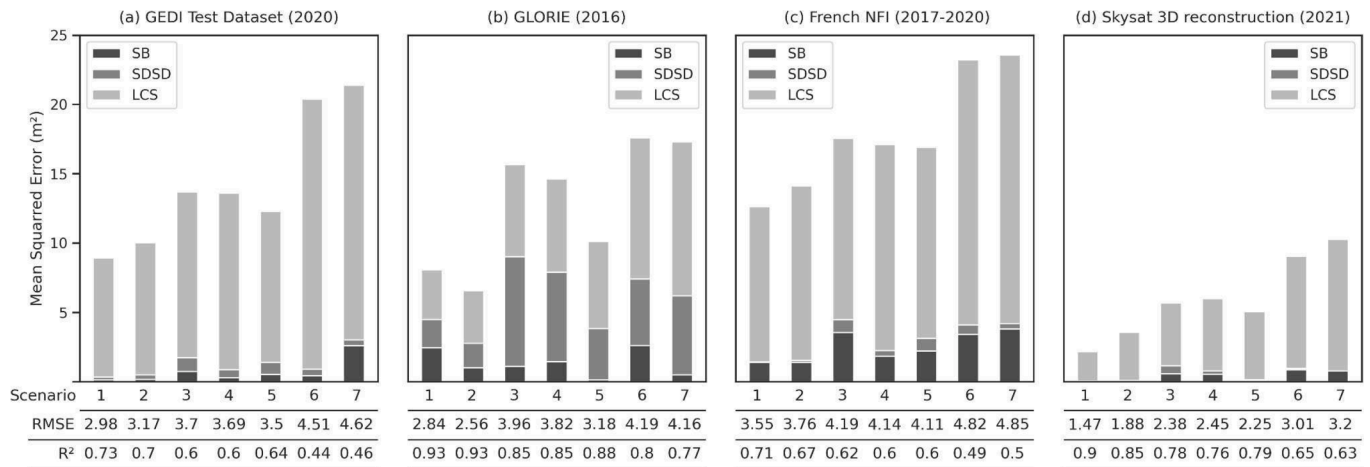
- Influence of the S1 and S2 layers

The analysis of the seven training scenarios (Fig. 11) reveals that the more S1 & S2 layers are available as inputs, the better the prediction is. However, training the model only with a subset of the 14 layers proposed as inputs (10 for S2, 4 for S1) still leads to quite satisfying error metrics. The S1 bands seem to be the most interesting predictors in our





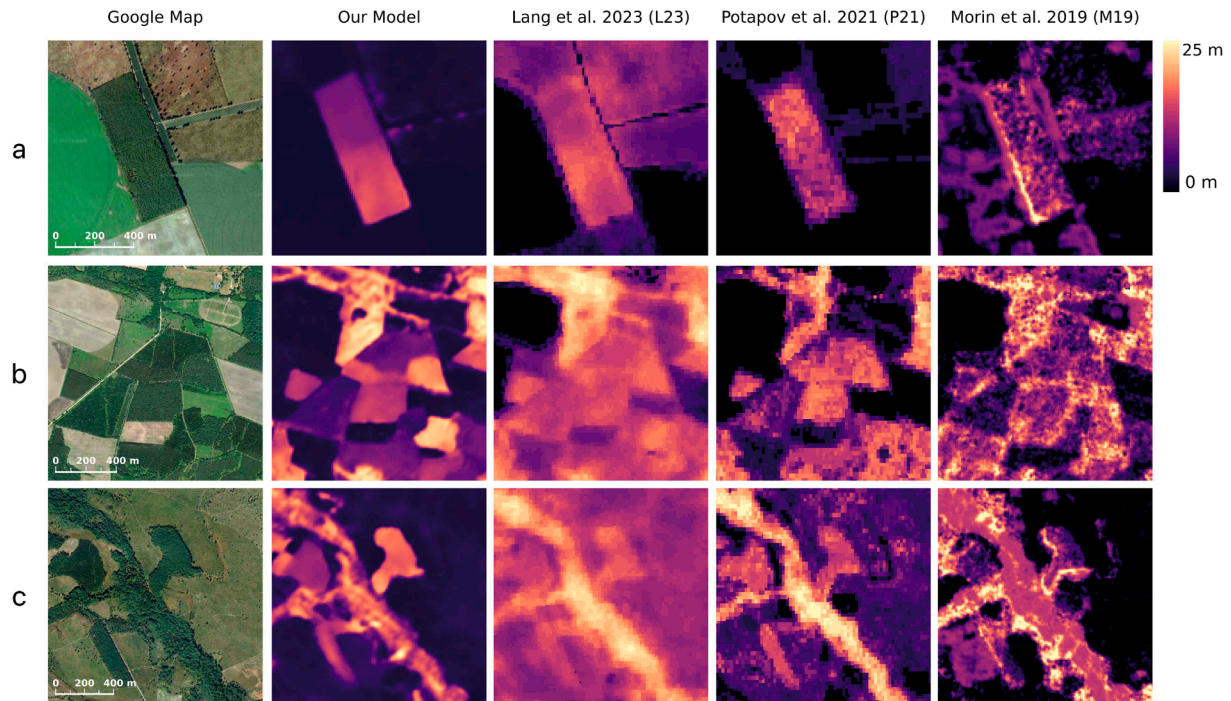
**Fig. 10.** Comparison of the FCN prediction model for Scenario 1 (2020) and a 3D canopy height model (3D CHM) based on Skysat imagery (2021). (a) Pixel distribution of the FCN and 3D CHM height predictions resampled at 10 m with a max method. (b) Scatter plot of the comparison of the pixel median over forest stands which were labeled manually (Fig. 4). The dashed line corresponds to  $x = y$ . (c) Stereo 3D reconstruction from Skysat imagery resampled at 10 m (2021). (d) FCN model (2020). (e) Comparison of the FCN and 3D CHM height predictions over a 1-km height profile shown in Fig. 10c and 10d.



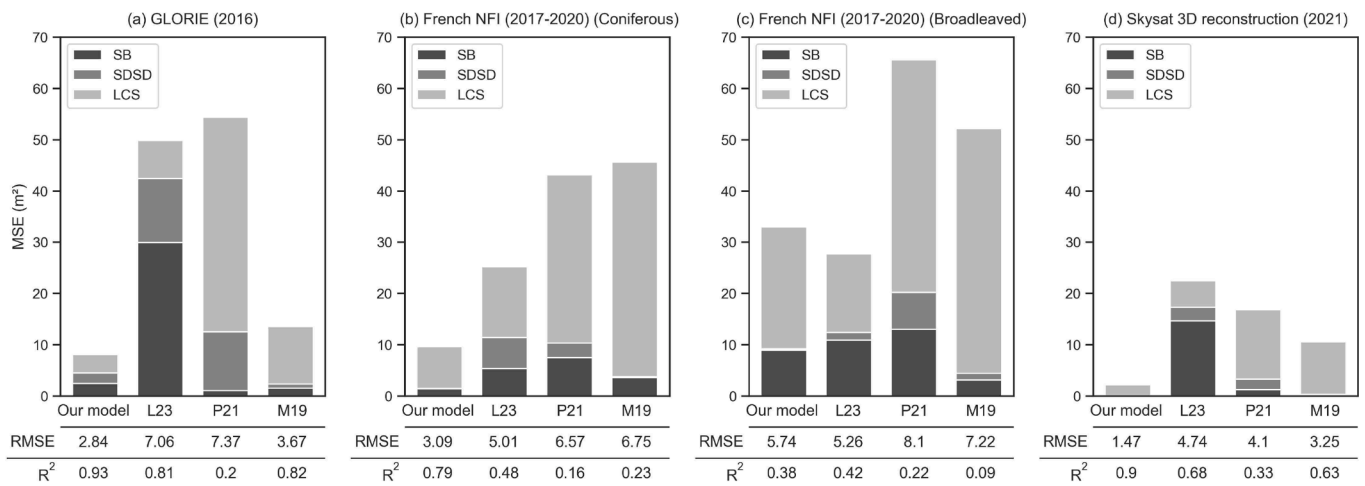
**Fig. 11.** Comparison of the error metrics on four evaluation datasets for the 7 training scenarios. Mean squared error is decomposed into three additive terms (see 2.4.4). Squared Bias (SB) is the squared difference of the mean of both datasets. Squared Difference between Standard Deviations (SDSD) indicates how the model is able to simulate the magnitude of the fluctuations between the  $n$  measurements. The lack of correlation weighted by the standard deviations (LCS) indicates the ability of the model to simulate the fluctuations across the  $n$  measurements. RMSE and R<sup>2</sup> values are presented in the tables below the graphs. Description of the scenarios: 1 - all bands; 2 - all S1; 3 - all S2; 4 - S2 RGB + NIR; 5 - S1 VV<sub>des</sub> + VH<sub>des</sub>; 6 - VV<sub>des</sub>; 7 - S2 NIR.

case (Scenario 2, RMSE = 3.76 m on French NFI), and adding 10 S2 bands (Scenario 1) only decreases the RMSE to 3.55 m. But a model trained only with S2 (Scenario 3) is still able to carry out good predictions (RMSE = 4.19 m on French NFI). Scenario 4 highlights that only the S2 spectral bands with a 10 m resolution are necessary to obtain such

results which can be related to the higher resolution of these bands but also to the contained information (RGB + NIR). A combination of two S1 layers (Scenario 5) leads to much better results than a single layer (Scenario 6). In other results that were not presented here, we found that any combination of two S1 bands led to better results than the best S1



**Fig. 12.** Comparison of our model (FCN model) with three independent canopy height models (Lang et al., 2023; Morin et al., 2019; Potapov et al., 2021) at three different locations in the test tiles. (a) Two homogeneous forest stands of maritime pines with different heights. (b) Several forest stands of maritime pines with different heights. (c) Broadleaved forest along a river path..



**Fig. 13.** Comparison of the error metrics for the different models: FCN (our model), L23 (Lang et al., 2023), P21 (Potapov et al., 2021), and M19 (Morin et al., 2019) against data from the GLORIE forest inventory (a), the French NFI separated into coniferous (b) and broadleaved (c) forests and the stereo 3D reconstruction of Skysat imagery (d). The mean squared error (MSE) is decomposed into three additive terms (see 2.4.4). RMSE and R<sup>2</sup> values are presented in the tables below the graphs.

band alone (VH descending). Hence, increasing the number of input bands improves the predictive power of the model. Even though they might be very specific to the very particular forest structure of the Landes forest, these results suggest that, for highly cloudy regions where cloud-free S2 images are rare, a training based only on S1 could be sufficient to retrieve forest height.

• Better validation with Skysat 3D reconstruction

The validation datasets used in this study, as presented in Fig. 2, show significant discrepancies in the error metrics, as shown in Fig. 11 and Fig. 13. For example, the RMSE for training scenario 1 is 3.55 m when compared to the French NFI data, while it is only 1.47 m for the

Skysat 3D reconstruction. Each validation dataset has its own specific characteristics related to the quality of the measurements, the date of acquisition, or the representativeness of the validation area. The validation results for the 3D reconstruction from Skysat imagery are superior for our model (Fig. 11) and other canopy height maps (Fig. 13). This may be due to three reasons. First, the Skysat imagery area only covers homogeneous coniferous pine stands, which are more easily retrieved by our FCN model, as well as P21, L23, and M19, as shown in Fig. 13b and Fig. 13c. Then, this 3D reconstruction technology is likely to be more reliable than ground-based measurements for height retrieval and it was performed in 2021, while GLORIE and the French NFI comprise measurements that have a longer time difference with maps to evaluate. Lastly, the height at the forest stand level was aggregated before doing

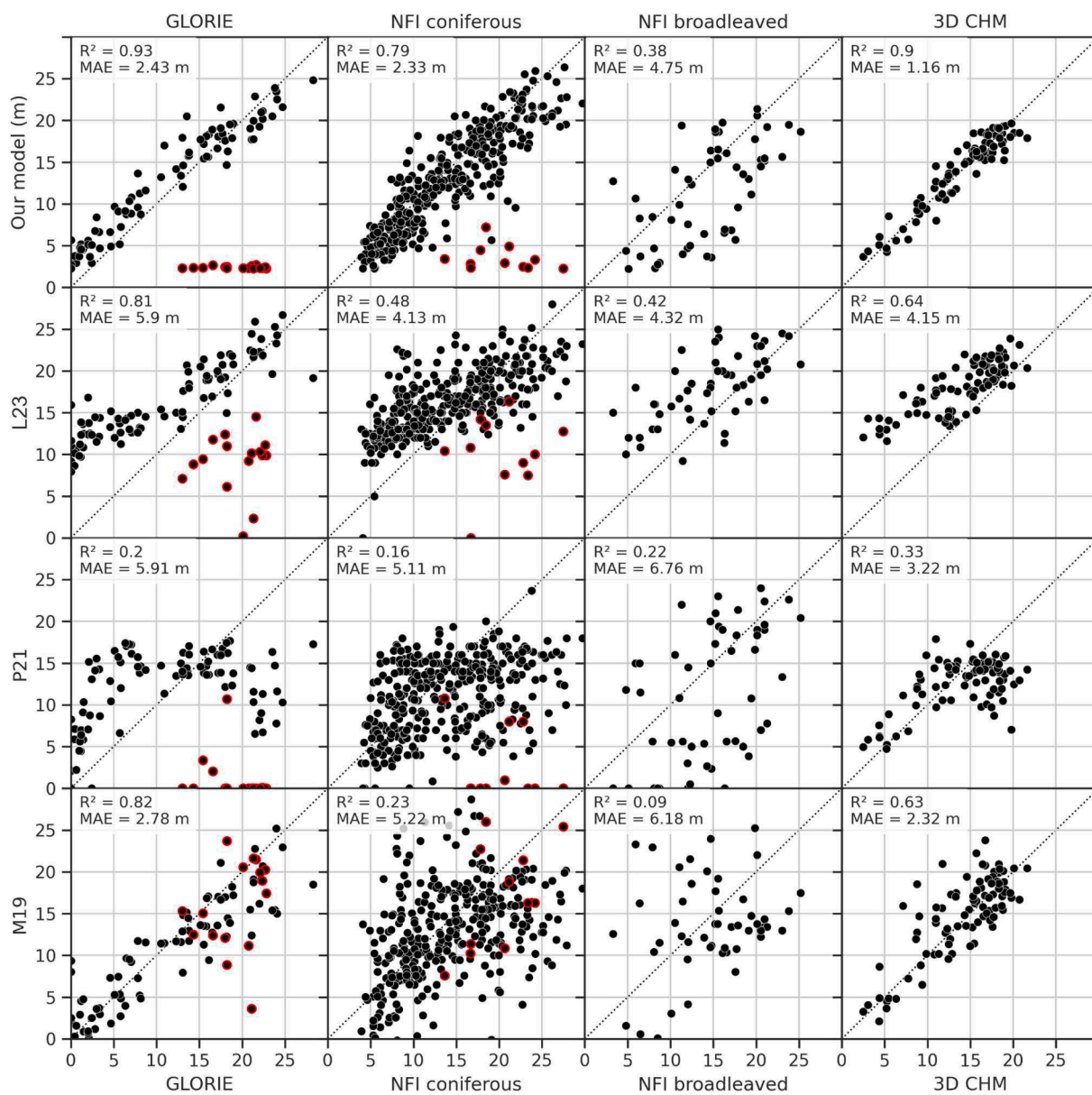


Fig. 14. Evaluation of the FCN model (Scenario 1), L23 (Lang et al., 2023), P21 (Potapov et al., 2021), and M19 (Morin et al., 2019) against data from the GLORIE forest inventory, the French NFI (separated into broadleaved and coniferous) and the 3D height reconstruction from Skysat imagery. Points circled in red indicate forest clearcuts between the date of inventory and 2020. The latter data are not considered in the calculation of  $R^2$  and MAE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the comparison to evaluate the median canopy height within each parcel. The error metrics provided by this comparison with the Skysat dataset thus offer insight into the accuracy of our FCN models in predicting forest height at the forest stand level, rather than at the pixel level.

- Comparison with other canopy height estimation models

For most evaluation metrics, the FCN model shows improved performance compared to available canopy height maps for the Landes forest. Better delimitation between landscape units, lower bias, better correlation, and better RMSE are obtained on all validation datasets except for the broadleaved part of the NFI data set where L23 is better. These results underline the importance of the scale at which the training is performed. P21 and L23 are the outputs from global models, trained with optical imagery (Landsat-8 for P21, S2 for L23) and GEDI metrics as reference height. Even though the number of GEDI data used specifically

for the Landes forest by P21 and L23 had the same order of magnitude as our train dataset, they only represent a very small fraction of the global GEDI dataset. The P21 and L23 models were trained to optimize height prediction globally and can make predictions on more diverse forest types but they are less specific to the Landes forest. For instance, L23 performs slightly better on broadleaved forests which is most likely related to the higher number of reference height data for this type of forest available in the global dataset used for training. It is likely that the bias observed for lower heights comes also from this global training process.

The comparison with M19 highlights the importance of representativity and quantity of reference height data. In this case, the reference data used for training M19 (GLORIE dataset only) is only representative of homogeneous areas within maritime pine plantation stands. Therefore, the M19 model did not learn to deal with different forest surfaces such as borders, gaps or broadleaved forests. This model shows good performances on the 3D Skysat reconstruction which is close to the area



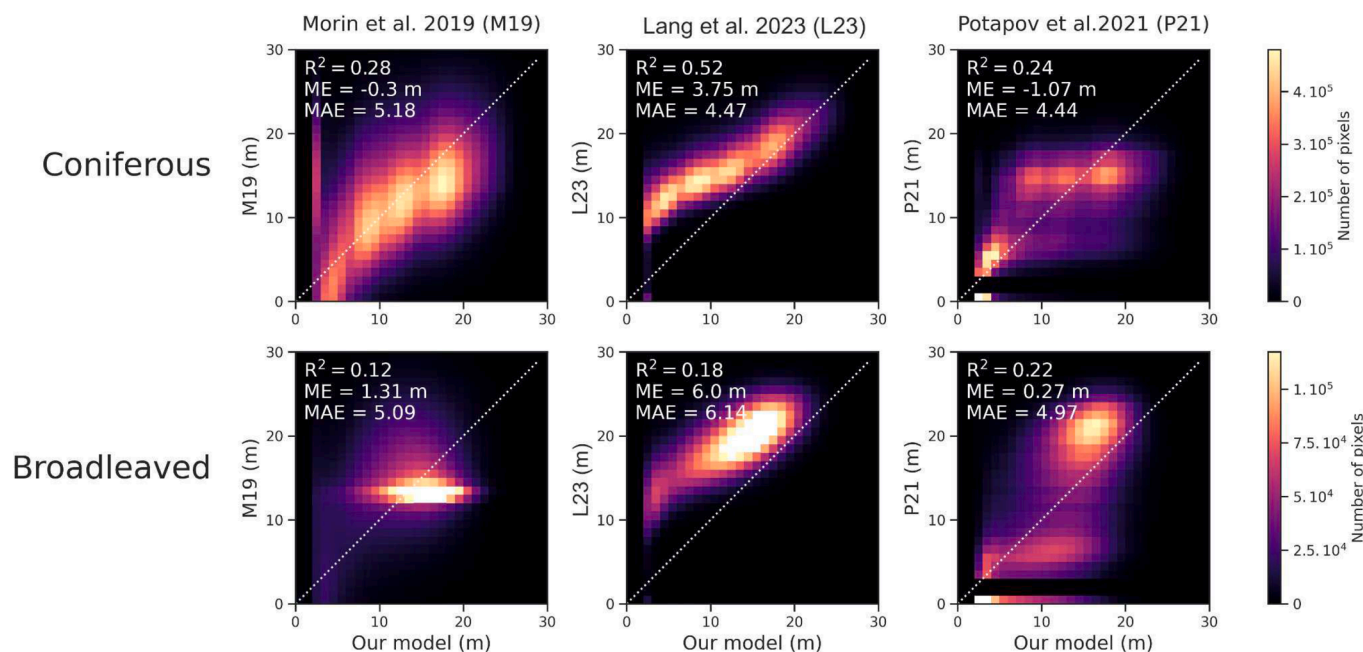


Fig. 15. Pixel-wise comparison of the FCN model with three independent canopy height models: L23, M19, and P21 (Lang et al., 2023; Morin et al., 2019; Potapov et al., 2021). Forest pixels were categorized into coniferous or broadleaf forests with the Copernicus forest type map. The white line represents  $x = y$ .

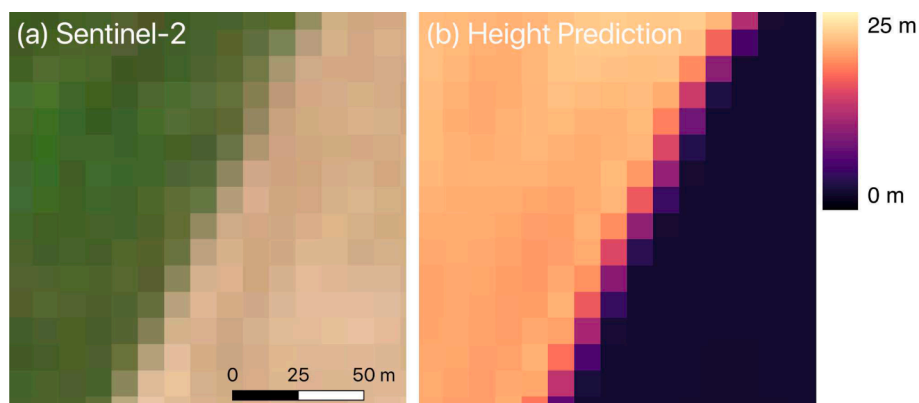


Fig. 16. (a) Sentinel-2 composite image used for prediction. The transition between the forest and non-forest area is not sharp. Some 10 m x 10 m pixels contain average information from both landscape units. (b) Height prediction from the FCN model (Scenario 1). The transition between the two landscape units is “smoothed”.

where it was trained and where we aggregated the height values at the forest stand level. Lower performances are observed on the French NFI, probably because some of the forest plots are close to forest borders and not located close to the training area. Additionally, the GLORIE forest inventory (see 2.3.1) contains only 99 height samples while we used 175,511 valid GEDI waveforms to train our FCN model.

Deep learning techniques applied to images (CNN) have the advantage of being “spatially aware” while other classical machine learning methods (Random Forest for P21, Support Vector Machine for M19) are not. In P21 and M19, forest texture metrics were added to the prediction variables in order to give this spatial awareness to the model but it is limited compared to the high number of convolution filters within the U-Net model that we used. In a newer study, Morin et al. (2022) have integrated GEDI features into their forest parameters retrieval method. Another comparison with these new methods would be interesting to understand whether it is deep learning (compared to simpler machine learning algorithms) or the number of training data (from GEDI) that makes our current model better than M19.

- On the use of GEDI

The GEDI mission provides an unprecedented database of LiDAR waveforms with high precision and has a high potential for forest parameter estimation. However, some caveats remain when using this type of data as ground truth for deep learning models. Indeed, the ability of GEDI to retrieve canopy height properly in more complex forest structures is still uncertain. In denser forests, some GEDI laser beams could not penetrate deep enough in the trees to reach the ground thus leading to a height underestimation. Moreover, the uncertainty associated with the GEDI footprint location could potentially lead to large errors. Here, we rasterized the GEDI footprint in a 10 m x 10 m pixel that corresponds to the center of this footprint. However, the height information contained in this GEDI footprint encompasses a 25 m diameter circle. Additionally, the 10 m uncertainty on the footprint location extends the area where the height information is potentially captured by GEDI to a 45 m diameter circle around the supposed center of the footprint. Hence the height information could potentially come from a point that is the third neighbor of the actual pixel that was rasterized.



These errors are randomly distributed and, thanks to the large number of GEDI samples, they only create noise in the reference height data used for training. So, it seems that these errors can be well handled by the FCN model. But in other more complex regions, this noise can be higher and the FCN model could have difficulties to carry out correct height predictions. To evaluate the impact of this uncertainty, we retrained our model with the previous GEDI release (GEDI v001). In this previous version, the standard deviation on the GEDI footprint location was 20 m (GEDI v001) and it was improved to 10 m in the GEDI v002 version. The MAE on the test dataset improved from 2.43 m to 2.02 m when using the more recent version of GEDI. FCN model outputs show clearer patterns and transitions between forested and non-forested areas for v002 vs v001 (See Appendix 4). The FCN model based on GEDIV001 produces blurred transitions between landscape units, thus trying to avoid high losses related to location errors. A further improvement in GEDI location has the potential to increase even more the accuracy of the FCN height maps we produced in this study.

• Influence of the number of training samples

The GEDI footprints are unevenly distributed globally because of the ISS trajectory. Therefore, in some other regions, fewer waveforms would be available to apply the same methodology. In order to assess how this can affect the reproducibility of the proposed method, we trained three additional models like the one from scenario 1 (all S1 and S2 bands) by randomly keeping only 10 %, 1 % and 0.1 % of the footprints from the training dataset. As shown in Table 1, a model trained with only 10 % of the original train dataset still leads to good results (MAE = 2.43 m on the test dataset) and a canopy height map that looked very similar to the original one by visual inspection. However, when the number of training samples drops to 1 % of the original sample size, the output map looks a little noisier with a lower MAE (3.02 m).

These results stress the importance of having large training datasets, especially for deep learning algorithms like our FCN model. However, the model still has good performances with only 10 % of the original train dataset size which makes this methodology robust and suitable for being applied in other regions with a lower amount of GEDI data.

5. Concluding remarks

This study highlights the potential of deep learning models to continuously map forest height at high resolution with sparse references, e.g. GEDI LiDAR data, and continuous images from Sentinel-1 SAR and Sentinel-2 multispectral imager. Additionally, it confirms the potential of GEDI data to produce a good estimation of forest height, especially when integrated into a deep learning prediction model that reduces uncertainty related to the measurement of a single LiDAR footprint. Our FCN model is able to retrieve forest height at 10 m resolution in a French coniferous plantation with a relatively good accuracy (MAE = 2.02 m) for the GEDI test dataset and a high correlation with independent height measurement sources. These results remain


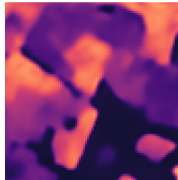
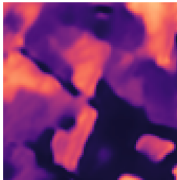
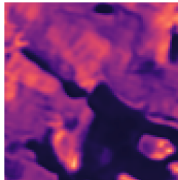
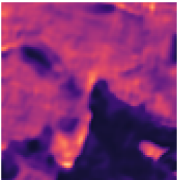
relatively good when using only one satellite source for model training which is particularly interesting for extending the study to other regions of the globe with a higher cloud cover or to use other satellite sources like Planet that have only four spectral bands but a higher spatial resolution. The map we produced showed improved results in comparison to other existing canopy height models over this region. The Landes forest is mainly composed of even-aged forest parcels of maritime pines. This very specific situation may have been “learned” by the deep learning model which tends to smooth the height within one forest stand and its performance on other types of forest structures is probably different. However, the method we developed still appears to be promising to retrieve canopy height on other forest structures such as broadleaved forests. For instance, we applied this methodology to another French “sylvoecoregion” called Sologne, mostly composed of broadleaved forest and obtained a similar accuracy with external validation from NFI plots (MAE = 2.62 m, see Appendix 6). Additionally, the area of interest is mostly flat. Sentinel-1 images are very sensitive to steep terrain and GEDI waveforms are also affected by higher slopes that tend to overestimate tree heights (Kutchartt et al., 2022). Further studies in more mountainous regions, e.g., by including a digital surface model in the deep learning process, could be potentially interesting to address these issues. Considering the availability of the Sentinel-1 and -2 observations, the canopy height map we retrieved could be used to monitor tree height with a yearly time frequency. Furthermore, due to the significant correlation between canopy height and biomass (Saatchi et al., 2011), the canopy height map could be used in a subsequent step to monitor forest biomass at the forest stand level and with a good temporal repetition (at least yearly), thus following the guidance of the Global Forest Observation Initiative (GFOI) to integrate earth observation data into national forest monitoring systems.

CRedit authorship contribution statement

**Martin Schwartz:** Data curation, Writing - original draft, writing - review & editing, Visualization, Investigation, Validation, Formal analysis, Methodology. **Philippe Ciais:** Writing - review & editing, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Catherine Ottlé:** Writing - review & editing, Supervision, Methodology, Conceptualization. **Aurelien De Truchis:** Writing - review & editing, Methodology. **Cedric Vega:** Writing - review & editing, Validation. **Ibrahim Fayad:** Writing - review & editing, Methodology, Data curation. **Martin Brandt:** Writing - review & editing, Conceptualization. **Rasmus Fensholt:** Writing - review & editing, Conceptualization. **Nicolas Baghdadi:** Writing - review & editing, Data curation. **François Morneau:** Writing - review & editing, Validation, Data curation. **David Morin:** Writing - review & editing, Data curation. **Dominique Guyon:** Writing - review & editing, Validation, Data curation. **Sylvia Dayau:** Writing - review & editing, Validation, Data curation. **Jean-Pierre Wigneron:** Writing - review & editing, Supervision, Methodology, Investigation, Conceptualization.

Table 1

Evolution of the MAE on the test dataset for 4 different FCN models trained with 100% (131.633 footprints), 10%, 1% and 0.1% of the original train dataset size. All the models correspond to scenario 1 (All S1 and S2 bands). The second row shows an example of the height prediction, with the corresponding S2 image in the first column.

Size of the training sample (% of the original training sample size)	100 %	10 %	1 %	0,1%
MAE on test dataset	2.02 m	2.43 m	3.02 m	3.94 m
				

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jag.2024.103711>.

## References

- Adam, M., Urbazev, M., Dubois, C., Schullius, C., 2020. Accuracy assessment of GEDI terrain elevation and canopy height estimates in European temperate forests: influence of environmental and acquisition parameters. *Remote Sens.* 12, 3948.
- Ball, J.E., Anderson, D.T., Sr, C.S.C., 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *J. Appl. Remote Sens.* 11, 042609. [10.1117/1.JRS.11.042609](https://doi.org/10.1117/1.JRS.11.042609).
- Beck, J., Wirt, B., Armston, J., Hofton, M., Luthcke, S., Tang, H., 2020. GLOBAL Ecosystem Dynamics Investigation (GEDI) Level 2 User Guide.
- Berger, A., Gschwantner, T., McRoberts, R.E., Schadauer, K., 2014. Effects of measurement errors on individual tree stem volume estimates for the Austrian national forest inventory. *For. Sci.* 60, 14–24. <https://doi.org/10.5849/forsci.12-164>.
- Copernicus Land Monitoring Service, 2018. Tree Cover Density 2018 [WWW Document]. URL <https://land.copernicus.eu/pan-european/high-resolution-layers/forests/tree-cover-density/status-maps/tree-cover-density-2018> (accessed 1.12.22).
- Dalagnol, R., Wagner, F.H., Emilio, T., Streher, A.S., Galvão, L.S., Ometto, J.P.H.B., Aragão, L.E.O.C., 2022. Canopy palm cover across the Brazilian Amazon forests mapped with airborne LiDAR data and deep learning. *Remote Sens. Ecol. Conserv.* 8, 601–614. <https://doi.org/10.1002/rse2.264>.
- Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. *Sci. Remote Sens.* 1, 100002 <https://doi.org/10.1016/j.srs.2020.100002>.
- Dubayah, R.O., Hofton, M., Blair, J.B., Armston, J., Tang, H., Luthcke, S., 2021. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V002. 10.5067/GEDI/GEDI02.A.002.
- Duncanson, L., Kellner, J.R., Armston, J., Dubayah, R., Minor, D.M., Hancock, S., Healey, S.P., Patterson, P.L., Saarela, S., Marselis, S., Silva, C.E., Bruening, J., Goetz, S.J., Tang, H., Hofton, M., Blair, B., Luthcke, S., Fatoyinbo, L., Abernethy, K., Alonso, A., Andersen, H.-E., Aplin, P., Baker, T.R., Barbier, N., Bastin, J.F., Biber, P., Boeckx, P., Bogaert, J., Boschetti, L., Boucher, P.B., Boyd, D.S., Burslem, D.F.R.P., Calvo-Rodriguez, S., Chave, J., Chazdon, R.L., Clark, D.B., Clark, D.A., Cohen, W.B., Coomes, D.A., Corona, P., Cushman, K.C., Cutler, M.E.J., Dalling, J.W., Dalponte, M., Dash, J., de-Miguel, S., Deng, S., Ellis, P.W., Erasmus, B., Fekety, P.A., Fernandez-Landa, A., Ferraz, A., Fischer, R., Fisher, A.G., García-Abril, A., Gobakken, T., Hacker, J.M., Heinrich, M., Hill, R.A., Hopkinson, C., Huang, H., Hubbell, S.P., Hudak, A.T., Huth, A., Imbach, B., Jeffery, K.J., Katoh, M., Kearsley, E., Kenfack, D., Kljun, N., Knapp, N., Král, K., Krůček, M., Labrière, N., Lewis, S.L., Longo, M., Lucas, R.M., Main, R., Manzanera, J.A., Martínez, R.V., Mathieu, R., Memiaghe, H., Meyer, V., Mendoza, A.M., Moneris, A., Montesano, P., Morsdorf, F., Næsset, E., Naidoo, L., Nilus, R., O'Brien, M., Orwig, D.A., Papanthassiou, K., Parker, G., Philipson, C., Phillips, O.L., Pisek, J., Poulsen, J.R., Pretzsch, H., Rüdiger, C., Saatchi, S., Sanchez-Azofeifa, A., Sanchez-Lopez, N., Scholes, R., Silva, C.A., Simard, M., Skidmore, A., Stereńczak, K., Tanase, M., Torresan, C., Valbuena, R., Verbeeck, H., Vrska, T., Wessels, K., White, J.C., White, L.J.T., Zahabu, E., Zraggen, C., 2022. Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. *Remote Sens. Environ.* 270, 112845. [10.1016/j.rse.2021.112845](https://doi.org/10.1016/j.rse.2021.112845).
- El Hajj, M., Baghdadi, N., Wigner, J.-P., Zribi, M., Albergel, C., Calvet, J.-C., Fayad, I., 2019. First vegetation optical depth mapping from sentinel-1 C-band SAR data over crop fields. *Remote Sens.* 11, 2769. <https://doi.org/10.3390/rs11232769>.
- Fayad, I., Baghdadi, N., Bailly, J.-S., Barbier, N., Gond, V., Hajj, M.E., Fabre, F., Bourguine, B., 2014. Canopy height estimation in french guiana with LiDAR ICESat/GLAS data using principal component analysis and random forest regressions. *Remote Sens.* 6, 11883–11914. <https://doi.org/10.3390/rs61211883>.
- Fayad, I., Baghdadi, N.N., Alvares, C.A., Stape, J.L., Bailly, J.S., Scolforo, H.F., Zribi, M., Maire, G.L., 2021a. Assessment of GEDI's LiDAR data for the estimation of canopy heights and wood volume of eucalyptus plantations in Brazil. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 7095–7110. <https://doi.org/10.1109/JSTARS.2021.3092836>.
- Fayad, I., Ienco, D., Baghdadi, N., Gaetano, R., Alvares, C.A., Stape, J.L., Ferraco Scolfaro, H., Le Maire, G., 2021b. A CNN-based approach for the estimation of canopy heights and wood volume from GEDI waveforms. *Remote Sens. Environ.* 265, 112652 <https://doi.org/10.1016/j.rse.2021.112652>.
- Fayad, I., Baghdadi, N., Alcarde Alvares, C., Stape, J.L., Bailly, J.S., Scolforo, H.F., Cegatta, I.R., Zribi, M., Le Maire, G., 2021c. Terrain slope effect on forest height and wood volume estimation from GEDI data. *Remote Sens.* 13, 2136. <https://doi.org/10.3390/rs13112136>.
- Griscom, B.W., Adams, J., Ellis, P.W., Houghton, R.A., Lomax, G., Miteva, D.A., Schlesinger, W.H., Shoch, D., Siikamäki, J.V., Smith, P., Woodbury, P., Zganjar, C., Blackman, A., Campari, J., Conant, R.T., Delgado, C., Elias, P., Gopalakrishna, T., Hamsik, M.R., Herrero, M., Kiesecker, J., Landis, E., Laestadius, L., Leavitt, S.M., Minnemeyer, S., Polasky, S., Potapov, P., Putz, F.E., Sanderman, J., Silvius, M., Wollenberg, E., Fargione, J., 2017. Natural climate solutions. *Proc. Natl. Acad. Sci.* 114, 11645–11650. <https://doi.org/10.1073/pnas.1710465114>.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342, 850–853. <https://doi.org/10.1126/science.1244693>.
- IGN What is the NFI? [WWW Document] <https://inventaire-forestier.ign.fr/spip.php?rubrique74> 2022 accessed 1.12.22.
- Illarionova, S., Shadrin, D., Ignatiev, V., Shayakhmetov, S., Trekin, A., Oseledets, I., 2022. Estimation of the canopy height model from multispectral satellite imagery with convolutional neural networks. *IEEE Access* 10, 34116–34132.
- IPCC, 2019. Special Report on Climate Change and Land.
- Jurjević, L., Liang, X., Gašparović, M., Balenović, I., 2020. Is field-measured tree height as reliable as believed – Part II, A comparison study of tree height estimates from conventional field measurement and low-cost close-range remote sensing in a deciduous forest. *ISPRS J. Photogramm. Remote Sens.* 169, 227–241. <https://doi.org/10.1016/j.isprsjprs.2020.09.014>.
- Kitahara, F., Mizoue, N., Yoshida, S., 2010. Effects of training for inexperienced surveyors on data quality of tree diameter and height measurements. *Silva Fenn.* 44 <https://doi.org/10.14214/sf.133>.
- Kobayashi, K., Salam, M.U., 2000. Comparing simulated and measured values using mean squared deviation and its components. *Agron. J.* 92, 8.
- Kutchartt, E., Pedron, M., Pirotti, F., 2022. Assessment of canopy and ground height accuracy from gedi lidar over steep mountain areas. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 3, 431–438. <https://doi.org/10.5194/isprs-annals-v3-2022-431-2022>.
- Lang, N., Schindler, K., Wegner, J.D., 2019. Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sens. Environ.* 233, 113347 <https://doi.org/10.1016/j.rse.2019.111347>.
- Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., Wegner, J.D., 2022. Global canopy height regression and uncertainty estimation from GEDI LiDAR waveforms with deep ensembles. *Remote Sens. Environ.* 268, 112760 <https://doi.org/10.1016/j.rse.2021.112760>.
- Lang, N., Jetz, W., Schindler, K., Wegner, J.D., 2023. A high-resolution canopy height model of the Earth. *Nat. Ecol. Evol.* 1–12. <https://doi.org/10.1038/s41559-023-02206-6>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>.
- Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., Chen, H., 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. *Int. J. Appl. Earth Obs. Geoinformation* 92, 102163. <https://doi.org/10.1016/j.jag.2020.102163>.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation. pp. 3431–3440.
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., Gascon, F., 2017. Sen2Cor for Sentinel-2, in: *Image and Signal Processing for Remote Sensing XXIII*. SPIE, pp. 37–48. [10.1117/12.2278218](https://doi.org/10.1117/12.2278218).
- Milesi, A., 2022. U-Net: Semantic segmentation with PyTorch.
- Morin, D., Planells, M., Guyon, D., Villard, L., Mermoz, S., Bouvet, A., Thevenon, H., Dejoux, J.-F., Le Toan, T., Dedieu, G., 2019. Estimation and mapping of forest structure parameters from open access satellite images: development of a generic method with a study case on coniferous plantation. *Remote Sens.* 11, 1275. <https://doi.org/10.3390/rs11111275>.
- Morin, D., Planells, M., Baghdadi, N., Bouvet, A., Fayad, I., Le Toan, T., Mermoz, S., Villard, L., 2022. Improving heterogeneous forest height maps by integrating GEDI-based forest height information in a multi-sensor mapping process. *Remote Sens.* 14, 2079. <https://doi.org/10.3390/rs14092079>.
- Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. *Science* 333, 988–993. <https://doi.org/10.1126/science.1201609>.
- Pereira-Pires, J.E., Mora, A., Aubard, V., Silva, J.M.N., Fonseca, J.M., 2021. Assessment of Sentinel-2 Spectral Features to Estimate Forest Height with the New GEDI Data. In: *Camarinha-Matos, L.M., Ferreira, P., Brito, G. (Eds.), Technological Innovation for Applied AI Systems, IFIP Advances in Information and Communication Technology*. Springer International Publishing, Cham, pp. 123–131. [https://doi.org/10.1007/978-3-030-78288-7\\_12](https://doi.org/10.1007/978-3-030-78288-7_12).
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J.B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* 253, 112165 <https://doi.org/10.1016/j.rse.2020.112165>.

- Pourrahmati, M.R., Baghdadi, N., Darvishsefat, A.A., Namiranian, M., Fayad, I., Bailly, J. S., Gond, V., 2015. Capability of GLAS/ICESat data to estimate forest canopy height and volume in mountainous forests of Iran. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* <https://doi.org/10.1109/JSTARS.2015.2478478>.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 234–241. [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc. Natl. Acad. Sci.* 108, 9899–9904. <https://doi.org/10.1073/pnas.1019576108>.
- Saunier, S., Done, F., Kocaman, S., Albinet, C., 2021. Technical Note on Quality Assessment for SkySat.
- Smith, L.N., 2017. Cyclical Learning Rates for training Neural Networks, in: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, Santa Rosa, CA, USA, pp. 464–472. [10.1109/WACV.2017.58](https://doi.org/10.1109/WACV.2017.58).
- Tommaso, S.D., Wang, S., Lobell, D.B., 2021. Combining GEDI and Sentinel-2 for wall-to-wall mapping of tall and short crops. *Environ. Res. Lett.* 16, 125002 <https://doi.org/10.1088/1748-9326/ac358c>.
- UNFCCC, 2015. Adoption of the Paris agreement.
- Wilkes, P., Jones, S.D., Suárez, L., Mellor, A., Woodgate, W., Soto-Berelov, M., Haywood, A., Skidmore, A., 2015. Mapping forest canopy height across large areas by upscaling ALS estimates with freely available satellite data. *Remote Sens.* <https://doi.org/10.3390/RS70912563>.
- Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., Yan, G., 2016. An easy-to-use airborne LiDAR Data filtering method based on cloth simulation. *Remote Sens.* 8, 501. <https://doi.org/10.3390/rs8060501>.
- Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* 5, 8–36. <https://doi.org/10.1109/MGRS.2017.2762307>.
- Zribi, M., Guyon, D., Motte, E., Dayau, S., Wigneron, J.-P., Baghdadi, N., Pierdicca, N., 2019. Performance of GNSS-R GLORI data for biomass estimation over the Landes forest. *Int J Appl Earth Obs Geoinformation* 74, 150–158. <https://doi.org/10.1016/j.jag.2018.09.010>, 2019.