Forest height maps obtained with ICESAT-2 data

A Degree Thesis Submitted to the Faculty of the

Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona

Universitat Politècnica de Catalunya

by

Carlos Torra Di Capua

In partial fulfilment of the requirements for the degree in

BACHELOR'S DEGREE IN TELECOMMUNICATIONS TECHNOLOGIES AND SERVICES **ENGINEERING**

Advisor: Mercè Vall-llossera

Co-Advisor: David Chaparro

Barcelona, May 2022

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

telecos BCN

Abstract

Climate change and global warming are currently on everyone's lips. The increase in the planet's average temperature caused by anthropogenic greenhouse gas emissions into the atmosphere is causing alterations in the climate that would not occur naturally.

Forests are main protagonists in this process. Climate change has a large effect on forests, due to the increase in temperatures and the alteration of precipitation levels, causing longer and more intense droughts that lead to forest dieback, increased number of big fires, and reduced capacity of forests to capture CO_2, among other effects. These impacts, in turn, increase CO_2 levels in the atmosphere aggravating the problem of climate change. Forests and tree plantations have the capacity to store carbon dioxide, which is necessary for the construction of their vital structures such as trunks, branches, leaves or fruits. This process allows them to act as the main land regulator of global CO_2 levels, reducing the impact of climate change when they act as carbon sinks (i.e., when they are not seriously affected by dieback, fires, deforestation, etc…).

In view of this situation and considering the fundamental role played by forests in these processes, this project aims to enhance our capacity to monitor a key forest structure variable: vegetation height. This variable is highly informative of forest structure and biomass, as well as of deforestation episodes, and thus contributes on monitoring the capacity of forests to act as a carbon reservoir. In this sense, we will obtain maps of the height of forest areas in Europe and North America using data measured by a LiDAR sensor on board of the ICESat-2 satellite. This has not been achieved before using ICESat-2, and therefore it will contribute to expand the current vegetation height products available. We will also provide a comparison of the resulting product with both precedent canopy height datasets (from ICESat satellite) and in situ information.

The resulting maps are consistent with the expected spatial patterns of forest height distribution, with higher forests in the most humid sites (e.g., northwestern US or central Europe), and lower ones in the drier regions (e.g., the Mediterranean region). The comparison between products and between our maps and the in-situ data needs to be extended with a larger sample of points and a more appropriate in situ canopy height dataset. Nevertheless, we find a good correlation ($r = 0.62$) between our product and the ICESat data for the North American region, which is a promising result.

We would continue the project by obtaining data from a more adequate in situ database for validation, by extending the ICESat-2 – ICESat comparison to a pixel-by-pixel approach, and by using new products (such as GEDI ones), which would allow us to perform a better validation assessment.

Resum

El canvi climàtic i l'escalfament global estan actualment en boca de tots. L'augment de la temperatura mitjana del planeta provocat per les emissions antropogèniques de gasos d'efecte d'hivernacle a l'atmosfera està provocant alteracions en el clima que no es produirien de manera natural.

Els boscos són els principals protagonistes d'aquest procés. El canvi climàtic té un gran efecte sobre els boscos, a causa de l'augment de les temperatures i a l'alteració dels nivells de precipitació, provocant sequeres més llargues i intenses que porten a la mort dels boscos, a l'augment del nombre de grans incendis i a la reducció de la capacitat dels boscos per a capturar CO_2, entre altres efectes. Aquests impactes, al seu torn, augmenten els nivells de CO_2 en l'atmosfera agreujant el problema del canvi climàtic. Els boscos i les plantacions d'arbres tenen la capacitat d'emmagatzemar diòxid de carboni, necessari per a la construcció de les seves estructures vitals com a troncs, branques, fulles o fruits. Aquest procés els permet actuar com el principal regulador terrestre dels nivells globals de CO_2, reduint l'impacte del canvi climàtic quan actuen com a embornals de carboni (és a dir, quan no es veuen greument afectats per la mort, els incendis, la desforestació, etc...).

Davant d'aquesta situació i tenint en compte el paper fonamental que juguen els boscos en aquests processos, aquest projecte pretén millorar la nostra capacitat de seguiment d'una variable clau de l'estructura forestal: l'altura de la vegetació. Aquesta variable és altament informativa de l'estructura i biomassa del bosc, així com dels episodis de desforestació, i per tant contribueix a monitorar la capacitat dels boscos per a actuar com regulador de carboni. En aquest sentit, obtindrem mapes de l'altura dels boscos a Europa i Amèrica del Nord utilitzant dades mesurades per un sensor LiDAR a bord del satèl·lit ICESat-2. Això no s'ha aconseguit fins ara utilitzant ICESat-2, per la qual cosa contribuirà a ampliar els actuals productes d'altura de la vegetació disponibles. També proporcionarem una comparació del producte resultant amb els conjunts de dades d'altura del dosser precedents (del satèl·lit ICESat) i la informació in situ.

Els mapes resultants són consistents amb els patrons espacials esperats de la distribució de l'altura dels boscos, amb boscos més alts en els llocs més humits (per exemple, el nord-oest dels EUA o Europa central), i més baixos a les regions més seques (per exemple, la regió mediterrània). La comparació entre productes i entre els nostres mapes i les dades in situ ha d'ampliar-se amb una mostra més àmplia de punts i un conjunt de dades d'altura de dosser in situ més adequat. No obstant això, trobem una bona correlació (r = 0,42) entre el nostre producte i les dades del ICESat per a la regió d'Europa, la qual cosa és un resultat prometedor.

Continuaríem el projecte obtenint dades d'una base de dades in situ més adequada per a la validació, ampliant la comparació ICESat-2 - ICESat a un enfocament píxel a píxel, i utilitzant nous productes (com els de GEDI), la qual cosa ens permetria realitzar una millor avaluació de la validació.

Resumen

El cambio climático y el calentamiento global están actualmente en boca de todos. El aumento de la temperatura media del planeta provocado por las emisiones antropogénicas de gases de efecto invernadero a la atmósfera está provocando alteraciones en el clima que no se producirían de forma natural.

Los bosques son los principales protagonistas de este proceso. El cambio climático tiene un gran efecto sobre los bosques, debido al aumento de las temperaturas y a la alteración de los niveles de precipitación, provocando sequías más largas e intensas que llevan a la muerte de los bosques, al aumento del número de grandes incendios y a la reducción de la capacidad de los bosques para capturar CO2, entre otros efectos. Estos impactos, a su vez, aumentan los niveles de CO2 en la atmósfera agravando el problema del cambio climático. Los bosques y las plantaciones de árboles tienen la capacidad de almacenar dióxido de carbono, necesario para la construcción de sus estructuras vitales como troncos, ramas, hojas o frutos. Este proceso les permite actuar como el principal regulador terrestre de los niveles globales de CO2, reduciendo el impacto del cambio climático cuando actúan como sumideros de carbono (es decir, cuando no se ven gravemente afectados por la muerte, los incendios, la deforestación, etc.…).

Ante esta situación y teniendo en cuenta el papel fundamental que juegan los bosques en estos procesos, este proyecto pretende mejorar nuestra capacidad de seguimiento de una variable clave de la estructura forestal: la altura de la vegetación. Esta variable es altamente informativa de la estructura y biomasa del bosque, así como de los episodios de deforestación, y por tanto contribuye a monitorizar la capacidad de los bosques para actuar como reservorio de carbono. En este sentido, obtendremos mapas de la altura de los bosques en Europa y América del Norte utilizando datos medidos por un sensor LiDAR a bordo del satélite ICESat-2. Esto no se ha conseguido hasta ahora utilizando ICESat-2, por lo que contribuirá a ampliar los actuales productos de altura de la vegetación disponibles. También proporcionaremos una comparación del producto resultante con los conjuntos de datos de altura del dosel precedentes (del satélite ICESat) y la información in situ.

Los mapas resultantes son consistentes con los patrones espaciales esperados de la distribución de la altura de los bosques, con bosques más altos en los sitios más húmedos (por ejemplo, el noroeste de EE. UU. o Europa central), y más bajos en las regiones más secas (por ejemplo, la región mediterránea). La comparación entre productos y entre nuestros mapas y los datos in situ debe ampliarse con una muestra más amplia de puntos y un conjunto de datos de altura de dosel in situ más adecuado. No obstante, encontramos una buena correlación (r = 0,62) entre nuestro producto y los datos del ICESat para la región de América del Norte, lo cual es un resultado prometedor.

Continuaríamos el proyecto obteniendo datos de una base de datos in situ más adecuada para la validación, ampliando la comparación ICESat-2 - ICESat a un enfoque píxel a píxel, y utilizando nuevos productos (como los de GEDI), lo que nos permitiría realizar una mejor evaluación de la validación.

I dedicate this work to my sisters Alba and Julia and to my mother Rosa and especially to my father Carlos, who from heaven has been encouraging and guiding me during these last years.

Acknowledgements

Personally, I would like to thank my advisor Mercè and my co-advisor David for having helped me throughout the project and for always being available to comment on any doubt, without this, the result of the project would be very different.

I would also like to thank my family, who has always helped me and has encouraged me to go ahead with it and not to give up in case of any problem that I could not solve.

Revision history and approval record

Revision	Date	Purpose
0	01/04/2022	Document creation
1	10/05/2022	Document revision
2	15/05/2022	Document Submission

DOCUMENT DISTRIBUTION LIST

Name	e-mail
Carlos Torra Di Capua	Carlos.torra1998@gmailcom
Mercè Vall-llossera Ferran	merce.vall-llosera@upc.edu
David Chaparro	david.chaparro.danon@gmail.com

Written by:		Reviewed and approved by:	
Date	05/05/2021	Date	10/05/2022
Name	Carlos Torra Di Capua	Name	Mercè Vall-llossera and David Chaparro
Position	Project Author	Position	Project Supervisors

Table of Contents

List of Figures

List of Tables:

1. Introduction

1.1. Statement of purpose

The goal of this project is twofold: (i) to obtain a canopy height map at 1 km resolution from the Ice, Cloud and Elevation Satellite-2 (ICESat-2) and (ii) to provide a comparison between forest canopy height data derived from the Light Detection and Ranging (LiDAR) sensors Ice, Cloud and Elevation Satellites 1 and 2 (ICESat and ICESat-2), as well as with in situ data.

On the one hand, what motivated me the most to do this project was my completely ignorance of the processes involved in data collection of a satellite. This project gave me the opportunity to learn how it works and to be capable of understanding all these data and analysing them to obtain a result.

To achieve that, I had to read, study, and understand bibliography that explained all the methods and technologies used in these past years to collect that data, as well as to obtain derived products, and the derived research and applications of them. Also, I had to develop useful programming skills with Matlab commands that I had not familiar with.

On the other hand, the scientific motivation of the project was based in two arguments:

- The importance of being able to monitor the forest height over the world, as it is a relevant variable to estimate biomass which, in turn, is essential to monitor the amount of carbon stored in forests. In addition, vegetation height is also being explored by researchers at UPC as an auxiliary variable in algorithms to estimate the hydric status of vegetation from microwave radiometry. Therefore, the project is fully in line with the group research activities.
- The fact that global vegetation height maps from previous (ICESat) and new (GEDI) missions are available, but neither global nor continental-scale canopy height maps exist for ICESat-2.

1.2. Requirements and specifications

Before starting the project, we had to define some requirements and propose the different specifications needed to achieve them.

First, we defined the following **requirements**:

- Collect all the ICESat-2 canopy height data that has been provided from the Remote Sensing Laboratory of the CommSensLab UPC group.
- Aggregate the ICESat-2 data collected to a 1-km grid.
- Define and create a Matlab script to process all data and develop the products' comparison.

These requirements allowed us to store all data at the same gridding with 1 km resolution and then, to compare the canopy height from different sensors and extract conclusions from the analysis of this comparison.

Considering all these requirements some **specifications** were stablished to start the project. The main one has been to work in two specific continents (Europe and North America), in order to reduce the size of the file maps and to handle the study in a laptop computer. We selected these zones because they are the ones with more in situ data available for validation. It is difficult to have in-situ data for validation. So, we divided the study in two parts: Europe data and North America data. As the same procedure was applied for both areas, we could work in parallel with both.

1.3. Methods and procedures

This project has applied ICESat-2 data downloaded by Dr. David Chaparro during previous research (Baur et al., 2021). The task here has been to recollect, screen, reprocess, aggregate, and validate these data, as well as to compare the result with another canopy height dataset available.

The main procedures involved in the project are the following:

- Collect all the ICESAT-2 data provided by the research group.
- Develop a Matlab script to process all these data and obtain the parameters that we needed to achieve the objective of the project, mentioned before. These parameters are the next ones:
  - Latitude
  - Longitude
  - Forest height
- Download another global canopy height map available on the literature and based in a previous LiDAR sensor (ICESat) as well as in situ canopy height data for validation and products comparison.
- Develop a Matlab Script to compare the ICESat-2 dataset resulting from this project with the ICESat canopy height map.

1.4. Work plan

The project is divided into a work breakdown structure with different work packages. A Gantt diagram is made having in mind that changes on the time plan depend on different circumstances that I have faced during the realization of the project.

1.4.1. Work Breakdown Structure

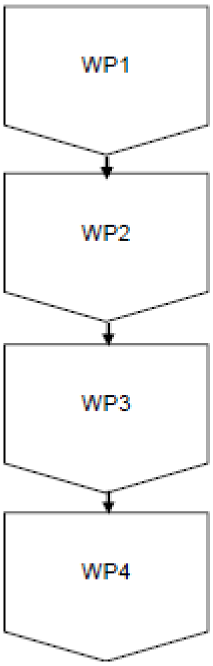

Figure 1. *Work Breakdown Structure*

1.4.2. Milestones

WP#	Tasks#	Short title	Milestone / deliverable	Date (week)
1	1	Theory	DOC about project idea	10/10/2021
2	2	Data processing	Study and Download data	24/12/2022
3	2	Data Structure	Script and program code	23/01/2022
4	1	Results and analyses	Global maps and comparisons	25/03/2022

Table 1. Milestones

1.4.3. Gantt Diagrams

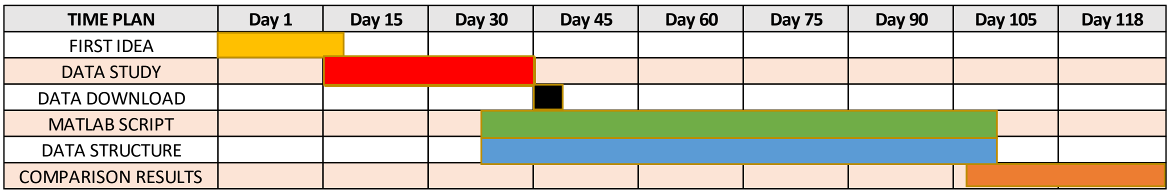

Figure 2. Initial Gantt Diagram.

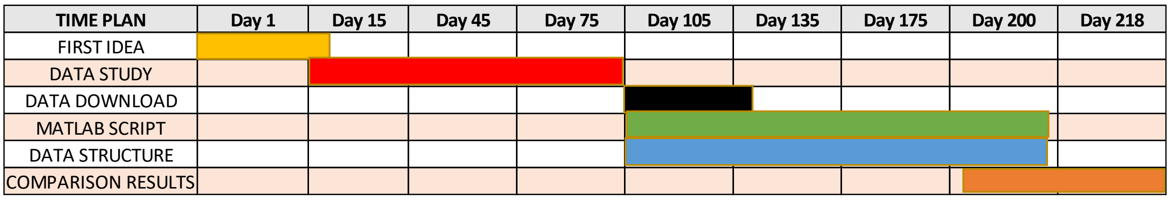

Figure 3. Final Gantt Diagram

1.5. Deviations of the original plan and incidences

The project started at the end of September 2021 and was planned to end in mid-January 2022, but due to some problems and unexpected issues that we will detail below, we decided to extend the duration of the project until mid-May 2022.

The project started very well. We knew what to do each week and we were following the timetable defined and planned at the beginning of the project. Nevertheless, by the end October 2021 we got stuck in the data collection and study. First, there was a delay in the process of providing us with the satellite data stored by our research group. Second, we had some issues with the data provided, for example, we were not familiar with the NetCDF files format and, in general with Earth observation data files. Then, we learned how to open, read and work with these files and data.

An additional problem has been the scarcity of in-situ data., We faced limitations concerning data from the in-situ stations, which should serve for validation of our resulting product. We realized that not enough in situ stations were available (see Section 3), which limited our capacity to provide a good validation process. In addition, these in situ stations are not specifically suited for canopy height measurements (see Section 3), which has finally impacted the validation result. Although alternative datasets were searched both by me and my advisors, we did not find a better alternative. Therefore, we anticipate that the in-situ validation approach is poor due to this limitation. We have included it in the work for academic purposes and have discussed the limitation in this document.

2. State of the art

2.1. Approach to the project

Global change is the greatest environmental challenge faced by our planet. It directly affects forests through global warming and precipitation changes, producing longer and more intense droughts in many regions that can lead to forest loss, worse forest health, and a reduced capacity of land ecosystems to capture CO_2. It also involves human activities leading to deforestation (e.g., substitution of forest by crops and contamination).

Due to their importance, it is necessary to make a regular monitoring of forests conditions, including their structure and their biomass. This monitoring permits to know where forest biomass is gained and where it is reduced. Therefore, we can know which forests are growing (i.e., act as carbon sinks) and we can know which forests are losing biomass (due to fires, droughts, or deforestation, acting as carbon sources). In the context of global change, to monitor the status of forests and their structure will help to know changes on carbon sinks and will help authorities to react and apply policies against climate change effects.

One of the variables which serve to monitor forest conditions is their canopy height. Forest canopy height is used to estimate forest aboveground biomass and timber volume, to monitor the effects of forest degradation, to measure the success of forest restoration, and to model other key ecosystem variables such as primary production and biodiversity. [1] [2]

Traditionally, tree height has been measured in the field and, more recently, using Airborne Laser Scanning (ALS) systems. These lasers, also known as LiDARs (Light Detection and Ranging), quantify the three-dimensional structure of vegetation and of the underlying terrain. They have been used extensively for enhancing geospatial knowledge in the fields of forestry and ecology. In the forestry field it gives structural descriptions of vegetation providing means of estimating a range of ecologically pertinent attributes, such as height, volume, and above-ground biomass. Recently, these sensors have been also used on board of satellites. The focus of this project is on satellite LiDAR sensors.

2.2. Remote sensing with LiDAR

LiDAR, which stands for Light Detection and Ranging, is a remote sensing method that uses light in the form of a pulsed laser that allows measuring the distance between a sensor and an object. These light pulses combined with other data recorded by an airborne or satellite system generate precise, three-dimensional information about the shape of the Earth and its surface characteristics.

A LiDAR instrument principally consists of a laser, a scanner, and a specialized GPS receiver. Airplanes and helicopters are the most used platforms for acquiring LiDAR data over broad areas. There are two types of airborne LiDAR:

- **Topographic LiDAR**. Collects geospatial data of natural and human-made environments to determine Earth's surface models, such as contours, elevation models, etc. This type of data is normally used for urban planning, coastal engineering, emergency responses operations, etc.
- **Bathymetric LiDAR.** The main feature of this kind of LiDAR is its ability to penetrate water. Nonetheless, information can be collected from both land and water.

Nevertheless, during last decades, LiDARs have been also used on board of satellites to obtain information of vegetation and topography of the entire Earth. The main satellites with LiDARs on board used for vegetation monitoring during last years are detailed hereafter.

ICESat (Ice, Cloud, and land Elevation Satellite) was the benchmark Earth Observing System mission for measuring ice sheet mass balance, cloud, and aerosol heights, as well as land topography and vegetation characteristics. ICESat was launched on 13 January 2003. The sensor on board was the Geoscience Laser Altimeter System (GLAS). It was the first laser instrument to be able to carry out continuous monitoring of the Earth. GLAS is a technological component that was designed and developed to measure the topography of ice sheets and to see the effect of temporal changes on them. In addition, it was also able to measure atmospheric and cloud properties to collect data on the height and thickness of these cloud layers. This process was important to make weather and climate predictions in a short period of time. This instrument worked both over land and water.

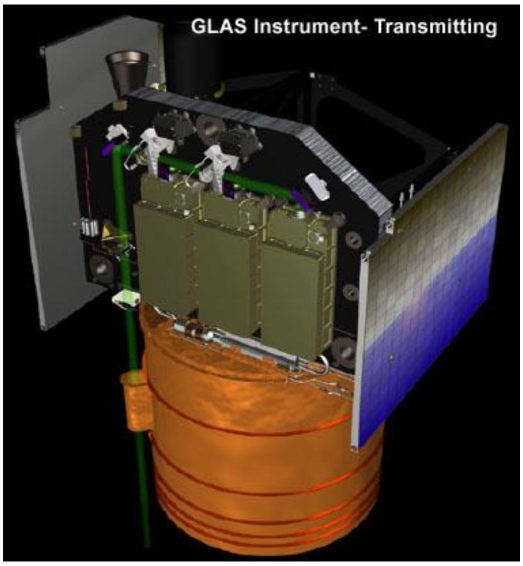

Figure 1. ICESat: GLAS Instrument. (Source: https://icesat.gsfc.nasa.gov/icesat/glas.php)

The laser transmitted short pulses (4 nano seconds) of infrared light (1064 nanometres wavelength) and visible green light (532 nanometres). The telescope (1 meter in diameter) integrated in the GLAS instrument collected the photons that were reflected from the Earth's surface and from the atmosphere. The pulses emitted by the laser traveled at a speed of 40 times per second, illuminating footprints up to 70 meters in diameter.

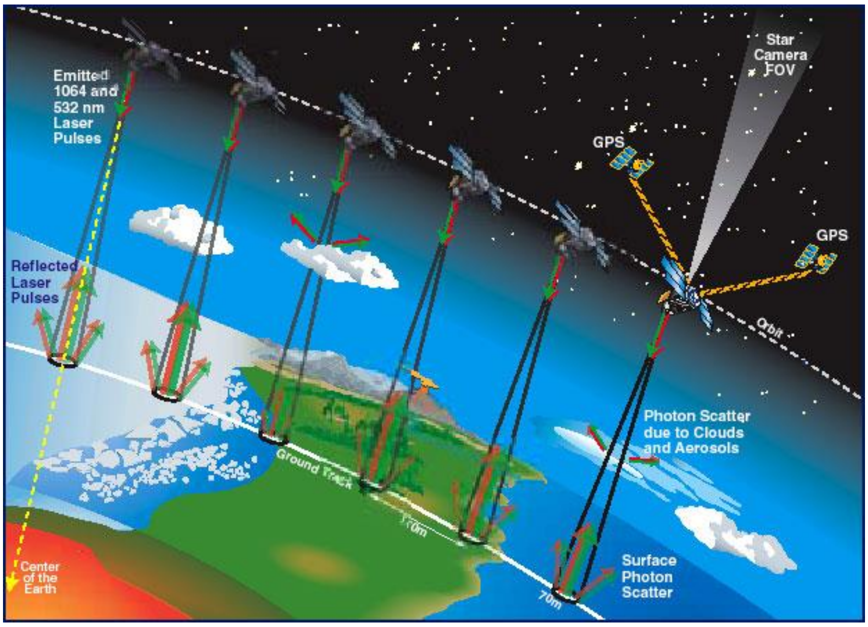

Figure 2. GLAS instrument making measurements from ICESat. It is measuring over land, oceans, and the Atmosphere (Source: https://attic.gsfc.nasa.gov/glas/about.html)

On August 14th of 2010, ICESat stopped working and was successfully decommissioned from operations. Figure 2 summarizes how ICESat instrument measured over land, oceans, and the atmosphere.

ICESat-2 (Ice, Cloud, and land Elevation Satellite-2) is a satellite that was launched on September 15, 2018, with the goal of measuring the elevation of the ice cover and the thickness of the sea ice, as well as the terrestrial topology, the characteristics of the vegetation and clouds.

It uses detection system of single photons. It is called ATLAS (Advanced Topographic Laser Altimeter System). This system is composed of two lasers, one primary and one backup. It can reach a speed of 10,000 pulses per second, it is significantly faster compared to the ICE Sat’s built-in laser, which reaches a speed of 40 pulses per second. Thanks to this speed, this tool can take measurements every 60 centimetres along the satellite's Earth trajectory.

The light pulses emitted travel through lenses and mirrors (see figure 3) within the satellite before being transmitted to the Earth's surface. During this process, the chronometer of the timing mechanism is started, the wavelength of the laser is checked and the size of the footprint on the Earth's surface is stablished. Also, the laser is divided into six beams.

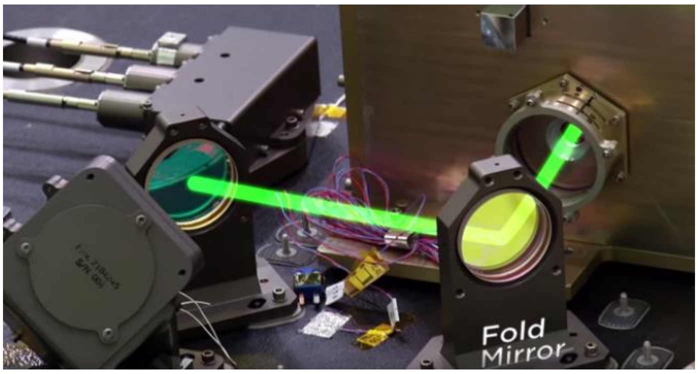

Figure 3. Travel of the pulses of light within the instrument of ICESat -2 (Source: https://icesat-2.gsfc.nasa.gov/mission)

GEDI (the Global Ecosystem Dynamics Investigation mission) is a new spaceborne LiDAR instrument operating on board of the International Space Station (ISS). It has been collecting data since April 18, 2019. While ICESat and ICESat-2 provide global data, the scope of GEDI is quasi-global, focusing between -51.6º and 51.6º latitude.

With the aim of enabling scientists to improve their ability to characterize important carbon and water cycle processes, biodiversity and habitat, this instrument allows measurements of:

- The height of the forest canopy.
- The vertical structure of the canopy.
- The elevation of the surface.

It is an instrument equipped with a system composed of 3 detection and light scattering lasers that produce 8 parallel tracks of observations (see Figure 4). Each of these lasers has a speed of 242 times per second allowing to illuminate 25 meters on the surface where the 3D structure is measured.

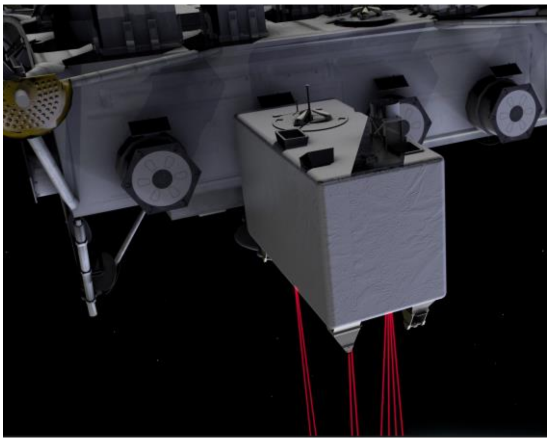

Figure 4. Geodetic-class light built into the satellite. (Source: https://gedi.umd.edu/mission/mission-overview/)

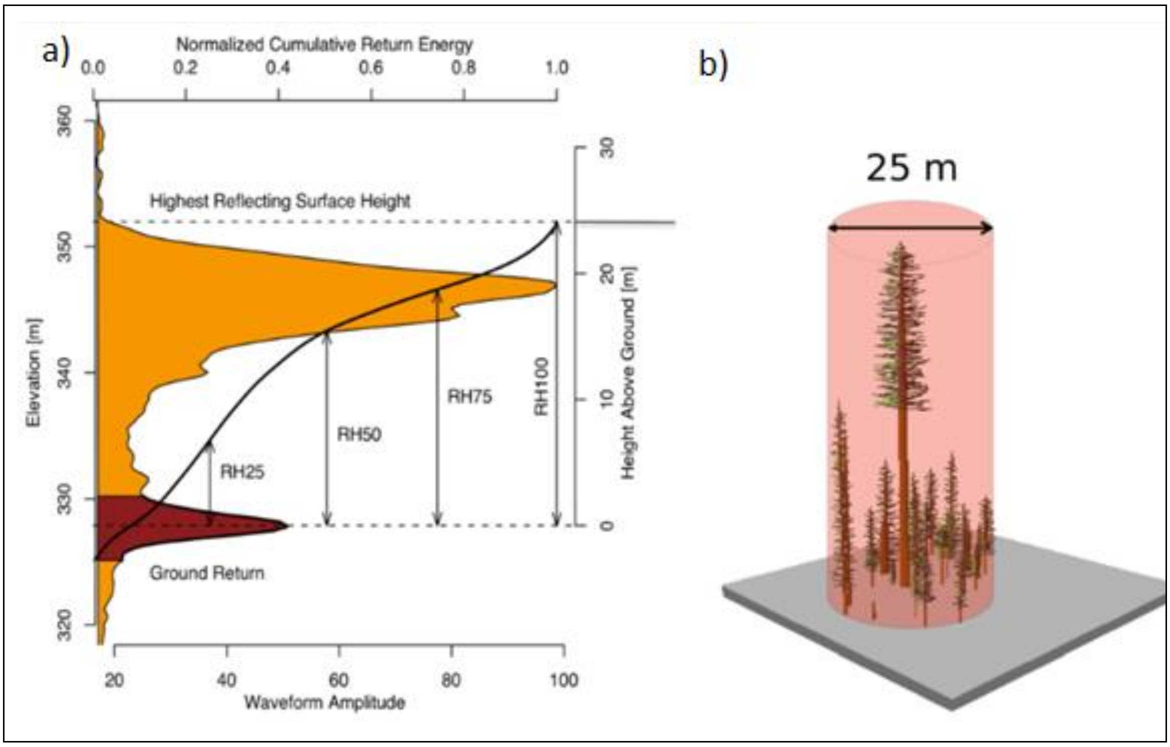

Figure 5. Coverage of the data collected by GEDI. (Source: https://gedi.umd.edu/mission/mission-overview/).

Figure 5a shows a sample example of a GEDI lidar waveform. The light brown area under the curve shows the return of the energy from the canopy, while the dark red area shows the return from the underlying topography. The black line is the cumulative return of the energy, starting from the bottom of the ground return to the top of the canopy. Relative Height (RH) metrics give the height reached at a certain quantile of energy returned relative to the ground. Figure 5b shows the distribution of trees that produced the waveform shown on the left.

To end this section, table 1 presents a comparative summary among the three presented satellite missions using LiDARs.

LIDARS	ICESAT	ICESAT-2	GEDI
Operative Period	2003-2010	2018-Present	2019-Present
Number of lasers	1	2	3
Laser velocity [pulses/s]	40	10.000	242
Footprint covered [m]	70	13	25
Wavelength [nm]	1.064	532	532

Table 1. Lidar comparison.

2.3. Satellite canopy height products

Several canopy height products have been developed from spaceborne LiDAR sensors during last years. Hereafter, we provide details on the main ones:

- The forest height product based on ICESat data consists of a global map of canopy height with a spatial resolution of 1 km [3]. The obtained results were validated with FLUXNET in situ data [4]. A correlation of 0,69 with an error of 4.36 m was found. This good performance has made this product a benchmark for scientific research.
- A forest height product based on GEDI data has been recently developed and consists of a 30 m spatial resolution global forest canopy height map for the year 2019 based on both GEDI and Landsat data. The integration of both datasets opened a new path to explore future satellite synergies for canopy height monitoring. The dataset is available in Google Earth Engine and can be also consulted through *https://gedi.umd.edu/data/download/.*
- A new GEDI product (combined with Sentinel-2 optical data) has been developed with satellite data for 2020 [4]. It presents a new automatic learning supervised approach to interpret the GEDI waveforms and make a global regression of canopy height also using Sentinel-2. It was proposed a new probabilistic deep machine learning approach based on neural networks to avoid explicit modeling of effects such as atmospheric noise, also called static electricity. This model learns to extract robust features that generalize to unseen geographic regions and, in addition, produces reliable estimates of predictive uncertainty. The main advantages of this product are that it provides canopy height data for all land cover types (not only forests) at an unprecedented spatial resolution (10 m.).

Concerning to ICESat-2, no global neither continental product are published, to the author knowledge. The only global dataset mentioned in the literature, to our knowledge, is the one applied as an auxiliary database in Baur et al. (2021), at 9-km gridding, which is used as well in this work for preliminary comparisons. Nevertheless, ICESat-2 has been used for canopy height monitoring in several research works. In that sense, regional studies with ICESat-2 canopy heights have been developed. An example is in dense tropical forests of Mexico, Belize, Guatemala, and Honduras [6] as well as in the United States [7]. Also, novel deep learning models have been applied to generate detailed maps of tree canopy heights, processing multitemporal ICESat-2 data [8].

3. Methodology / project development:

3.1. Study area

According to data collected in 2021 by the FAO [9] Europe's forests extend over 158 million hectares, and the forests of the United States extend over 310 million hectares. We decided to select the regions of Europe (35,818ºN – 56,751ºN; 13.549ºW-28.684ºE) and North America (10.103ºN-64.828ºN; 132.201ºW – 53.602ºE) to carry out this study. These are the regions for which more in situ vegetation height data were available.

3.2. Data sets

3.2.1. SAPFLUXNET

Sapfluxnet [10] is a worldwide database, managed by researchers at CREAF (Center for Ecological Research and Forestry Applications; Cerdanyola del Vallès, Barcelona, Spain), aimed at collecting measurements of sap flow in plants. Its main objective is to gather data on all plant characteristics and traits associated with each area to advance scientific understanding of the ecological factors of vegetation.

Their data on site are divided into different stations for each country. These stations collect data of all the vegetation for a given area, collecting parameters such as latitude, longitude, height, sap flow and hydric potential information, the date of measurements, or the type of plant or forest, among others.

As detailed in section 1, it is important to note that SAPFLUXNET data is not specifically suited for canopy height monitoring (i.e., height is an auxiliary variable). Nevertheless, we chose it for two reasons: (i) previous benchmark literature [11] successfully apply similar datasets [4] and (ii) despite looking for updated in situ height data from FLUXNET and other datasets, they were not longer available. In the latter case, an active search of FLUXNET canopy height data was done, without finding updated canopy height values in it. Therefore, finally the SAPFLUXNET alternative was applied.

To verify that the coordinates indicated in the SAPFLUXNET data corresponded to each station, we collected the longitude and latitude data available for each station in each region and plotted them on a map.

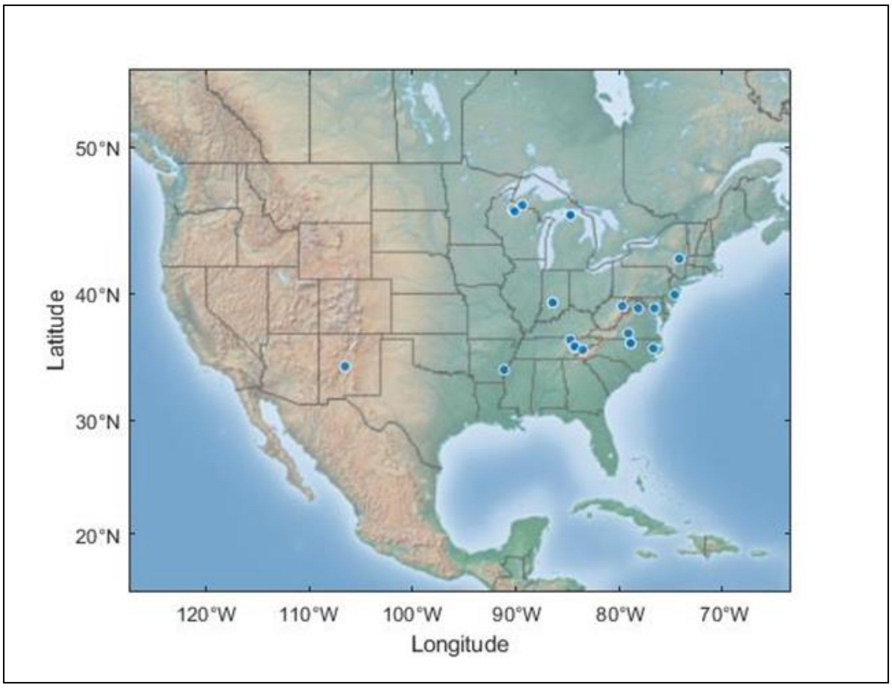

Figure 6. Geographical representation of North America stations according to their coordinates

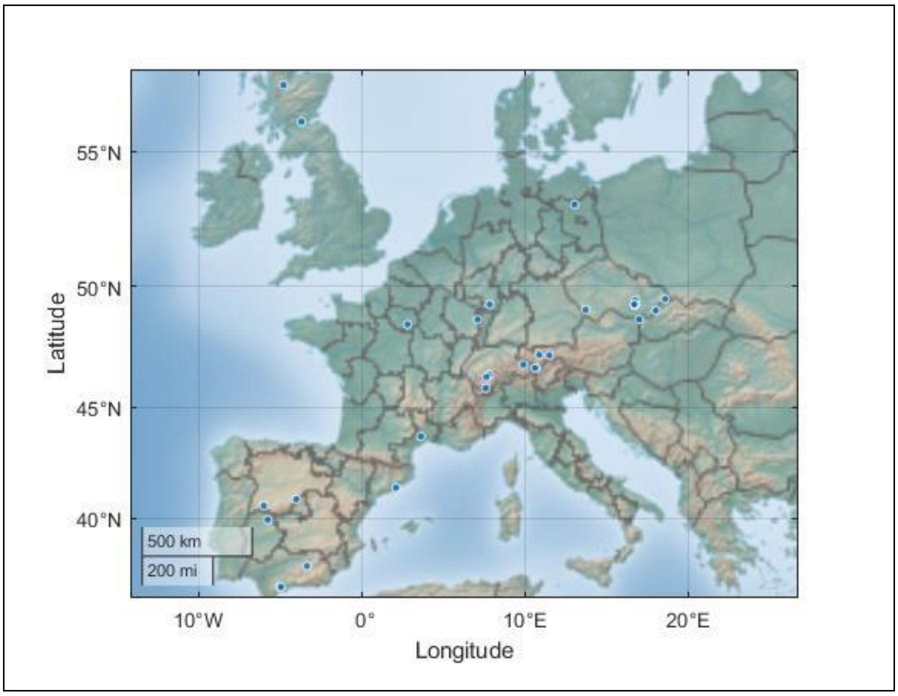

Figure 7. Geographical representation of Europe stations according to their coordinates.

It is interesting to use this type of basemap in the figure shown because then it is possible to difference the mountainous regions in which should not be any station. Nevertheless, in the figure 7 we can see that one station is in that region, that could be caused for an error during the process of data collection. This station will not be considered for the study of the project.

3.2.2. ICESat-2 canopy height data

Data from the ATL08 vegetation height product [12] from ICESat-2 has been provided by the Remote Sensing Laboratory in two different forms:

- A global canopy height product at 9-km gridding. More details of this dataset are found in Baur et al., 2021.
- The raw ATL08 ICESat-2 canopy height data, which contains the mean, median, percentile 98 and maximum of canopy heights for each 100 meters transect. The data period spans between 2016 and 2021.

The 9-km gridding product was used in the first part of this work to make a preliminary analysis of ICESat-2 canopy height data. The raw data (for the period October 2018-September 2019) was then used to produce the final canopy height map at 1 km resolution.

3.2.3. ICESat canopy height data

This is a global 1 km forest canopy height based on the data collected from ICESat. The product has been produced and published by Simard et al. (2011) and validated with field measurements from 66 FLUXNET sites. It has been downloaded from SDAT [13] a suite of Web-based applications that enable users to visualize and download spatial data in user-selected spatial/temporal extents, file formats, and projections.

3.3. Data processing and gridding

The ATL08 canopy height data (i.e., raw ICESat-2 data) was processed, first, by excluding data larger than 110 m (as it represents unrealistic values). Second, it was aggregated at 1 km gridding. The grid of destination was the same as that of ICESat, to make the products comparable.

The ICESat-2 dataset is appropriate only for measuring forest height. Nevertheless, some inaccurate data remained in regions where forests are not dominant. Hence, to get a more accurate map, we decided to apply a forest mask that discard data that do not meet a minimum percentage of forest per pixel. The ESA-CCI land cover map [14]. We decided to stablish a minimum percentage of forest of 60% so only pixels containing ≥60% of forest was considered. This is done to guarantee a highly representative sample of forest data.

Finally, it is worth noting that, as the time applied is limited to one year, the result will show some gaps between ICESat-2 overpasses. To overcome this issue, we have applied a smoothing filter using linear interpolation to fill no-data pixels with the best approximation.

3.3.1. Analysis of SAPFLUXNET Data

Figure 6 provides an example of a particular SAPFLUXNET station. It shows all collected samples and the height of each one within a plot. For each station, the data has been averaged.

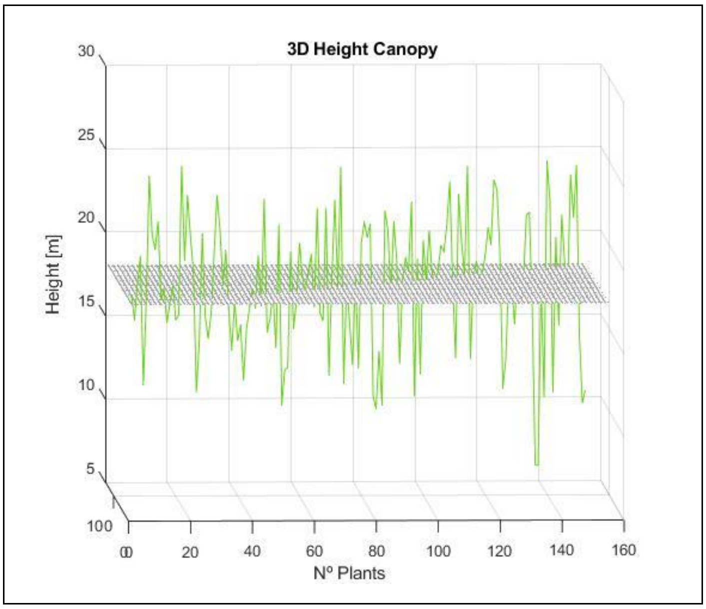

Figure 8. Analysis of SAPFLUXNET data for the station sited in 45.948ºN, -90.26ºW. It shows the height of each plant measured on site (green line) and its average (grey grid).

4. Results

4.1. Maps of 1-km canopy height from ICESat-2

4.1.1. Maps obtained without filtering

Canopy height maps for North-America and Europe are shown in Fig.10. and Fig.11, respectively. These maps provide the result without using the forest mask and without filling gaps between ICESat-2 overpasses.

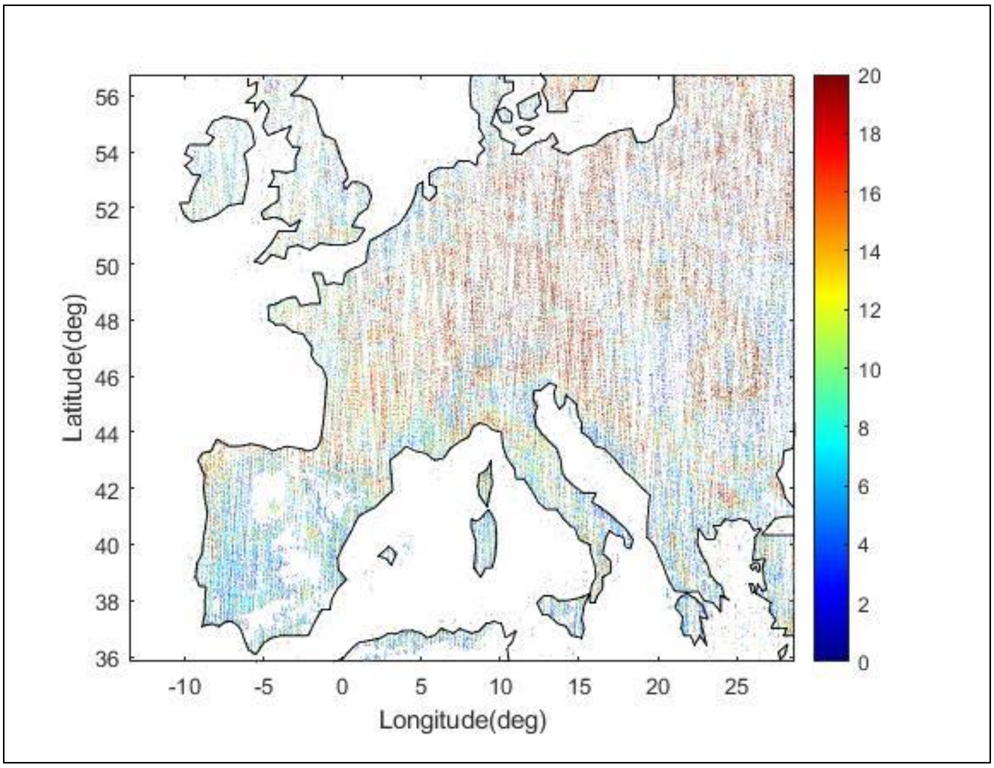

Figure 9. ICESat-2 data 1km map, Europe region.

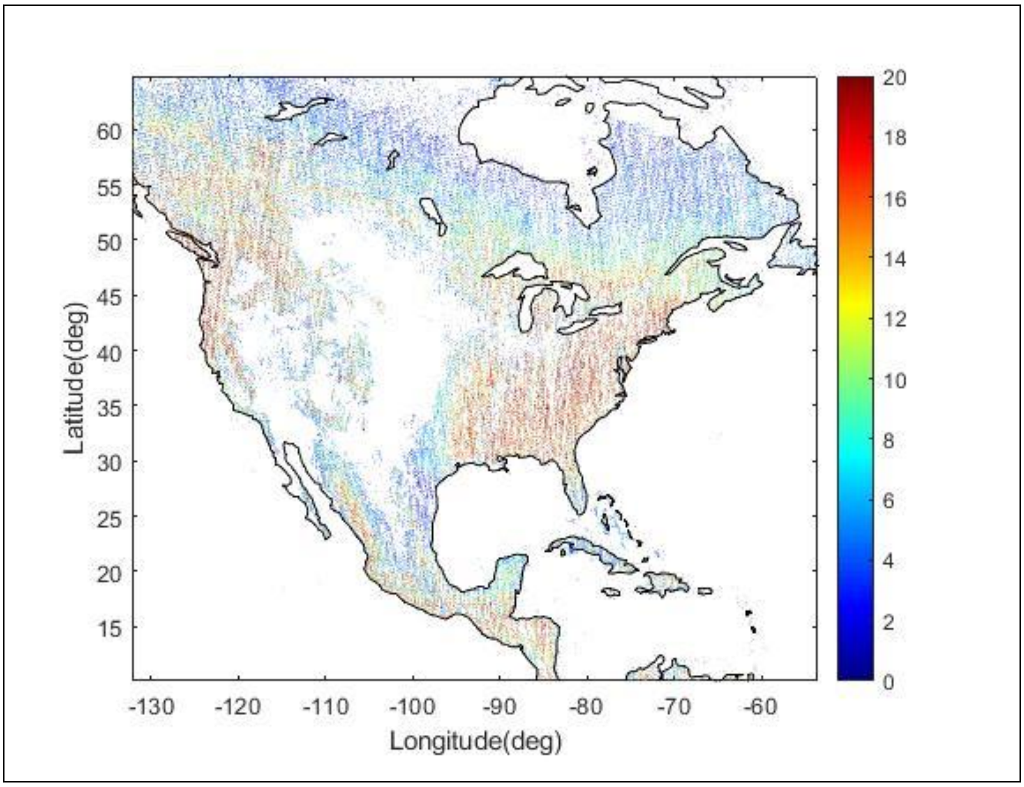

Figure 10. ICESat-2 data 1km map, North America region.

4.1.2. Maps obtained with forest mask and smothing filter

Once both the forest mask and the smoothing filter have been applied, the final product is obtained. These results are shown in Figures 12 and 13 for Europe and North America, respectively.

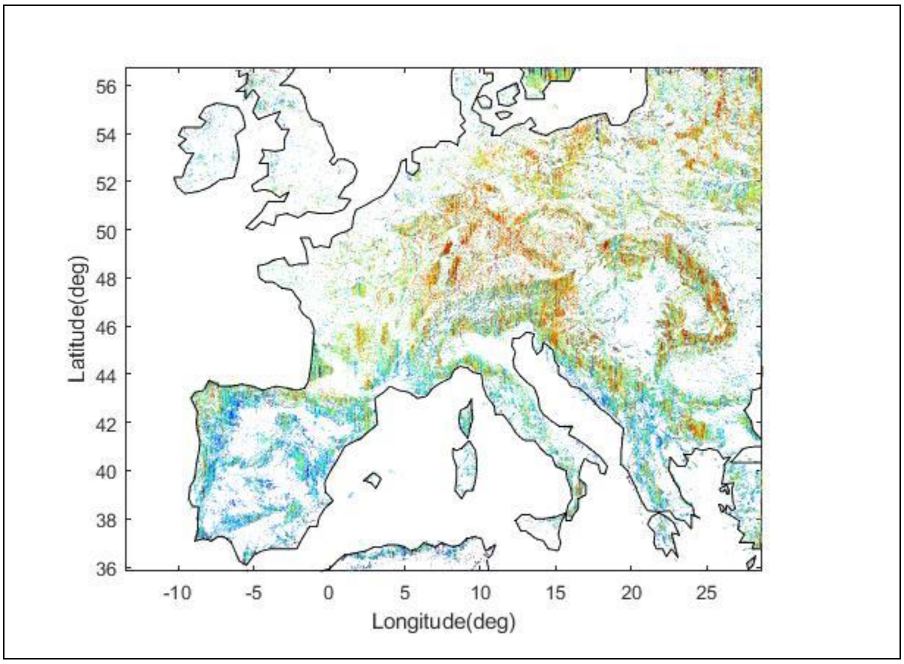

Figure 11. ICESat-2 data 1km map, Europe region.

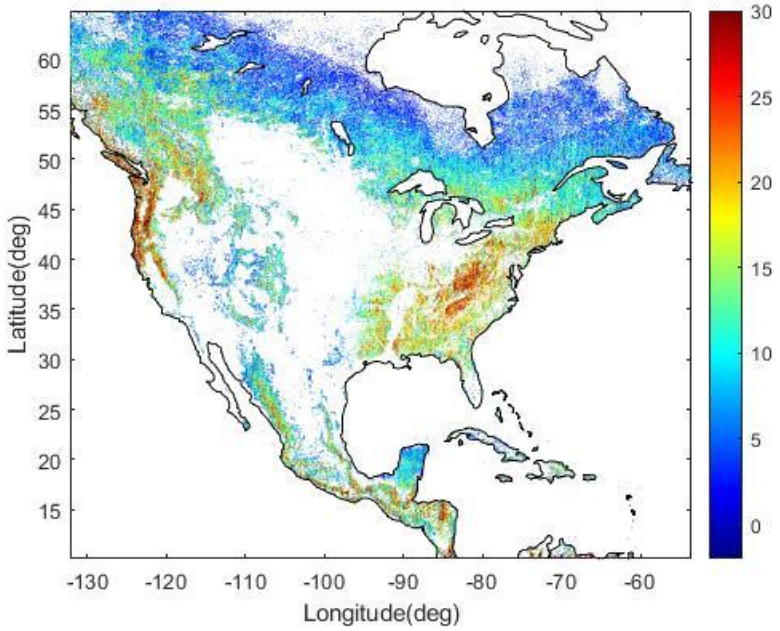

Figure 12. ICESat-2 data 1km map, North America region.

Now the results are presented only for forests and without discontinuities in the map. Analyzing the results obtained in Europe we can see that, on the one hand, the forests of central Europe have taller trees, consistently with the fact that they are placed in humid regions and have large biomass. On the other hand, in southern Europe as well as in northernmost Africa, lower trees are found due to a drier climate. An exception is found in the northwestern Iberian Peninsula, which has as well a humid Atlantic climate with dense and tall forests. Concerning North America, the tallest trees are in the humid forests of the states of Oregon and Washington (northwestern United States), as well as in some regions of California which are the habitat of sequoias' forests, the tallest trees on Earth. Therefore, the resulting maps are coherent from a biogeographical point of view.

4.2. Preliminary comparison between 1-km and 9-km products

To observe the correlation between the data with each mesh applied for ICESat-2 data, we performed a scatter plot comparing the canopy height data collected for both regions studied

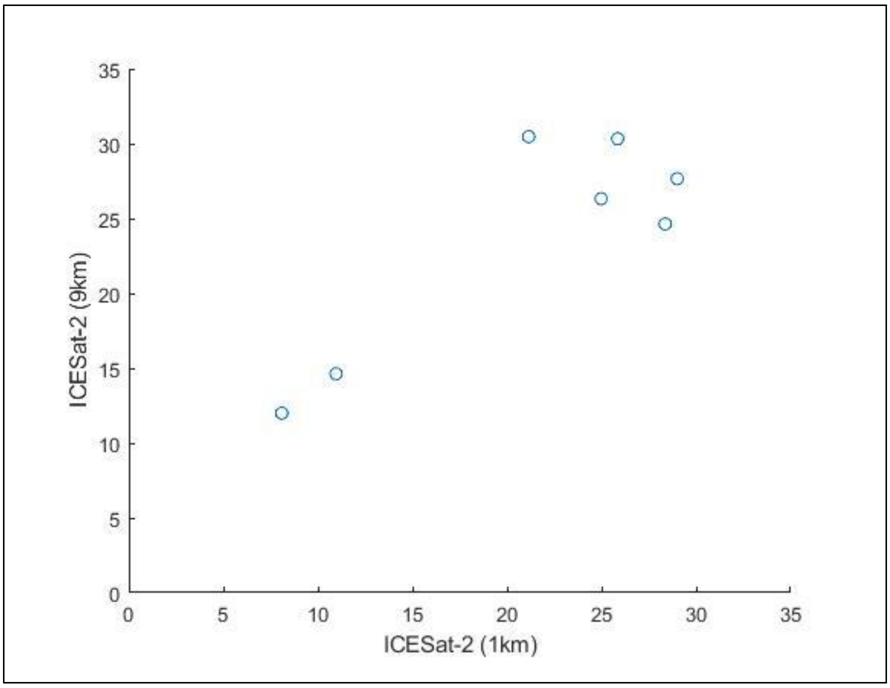

Figure 13. Mesh 1km-9km pixel relationship for Europe region.

As it shows the figure, there is a linear relationship between the two variables obtaining a correlation coefficient of 0,74. The difference with a perfect correlation must be caused by the different resolution of the collected data.

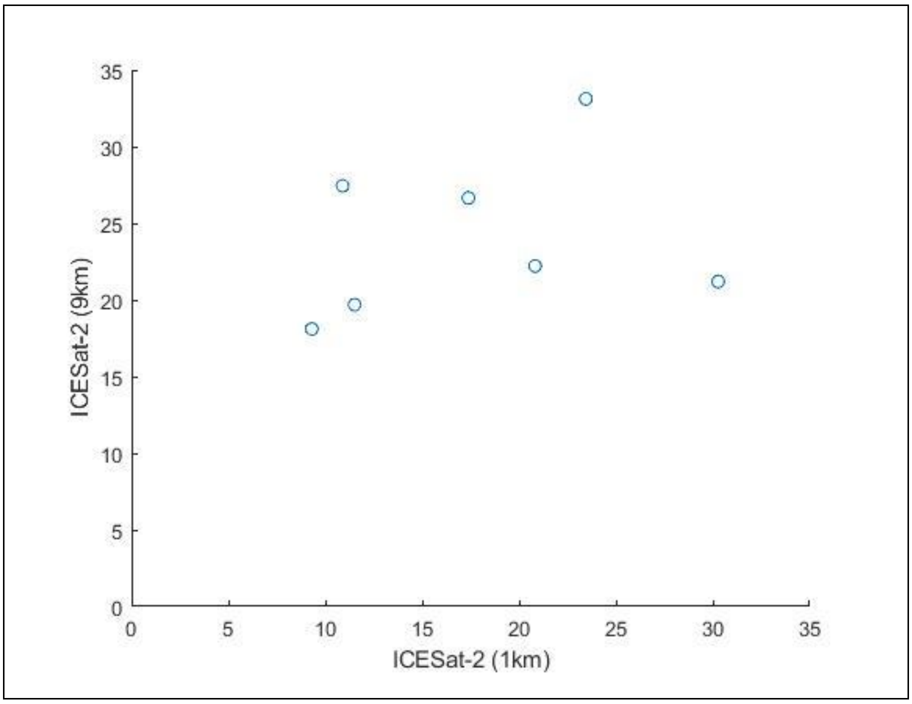

Figure 14. Mesh 1km-9km pixel relationship for North America region

As in the plot shown for Europe region, there is also a positive relationship between the two variables obtaining a correlation coefficient of 0,36. The difference with a perfect correlation must be for the same reason explained for Europe region.

4.3. Product comparison

Once the maps of canopy heights were obtained, a comparison of the new product with benchmark data was carried out. Figures 14 and 15 show the comparison of the ICESat-2, both at 9km and 1km resolution values with respect to in-situ measurements (Sapfluxnet stations) and ICESat data.

On the one hand for the region of Europe, positive trends are observed with Pearson's coefficient of correlation (r) equal to 0,74 for the ICESat-2(1km)-ICESat-2(9km) comparison, and r = 0,42 for the ICESat-ICESat2 (1km) comparison. Note, then, that the ICESat map and our product are well correlated, showing that our product has good performance for Europe according to the fact that the ICESat dataset had been already validated with good results (see Section 2.3) [11].

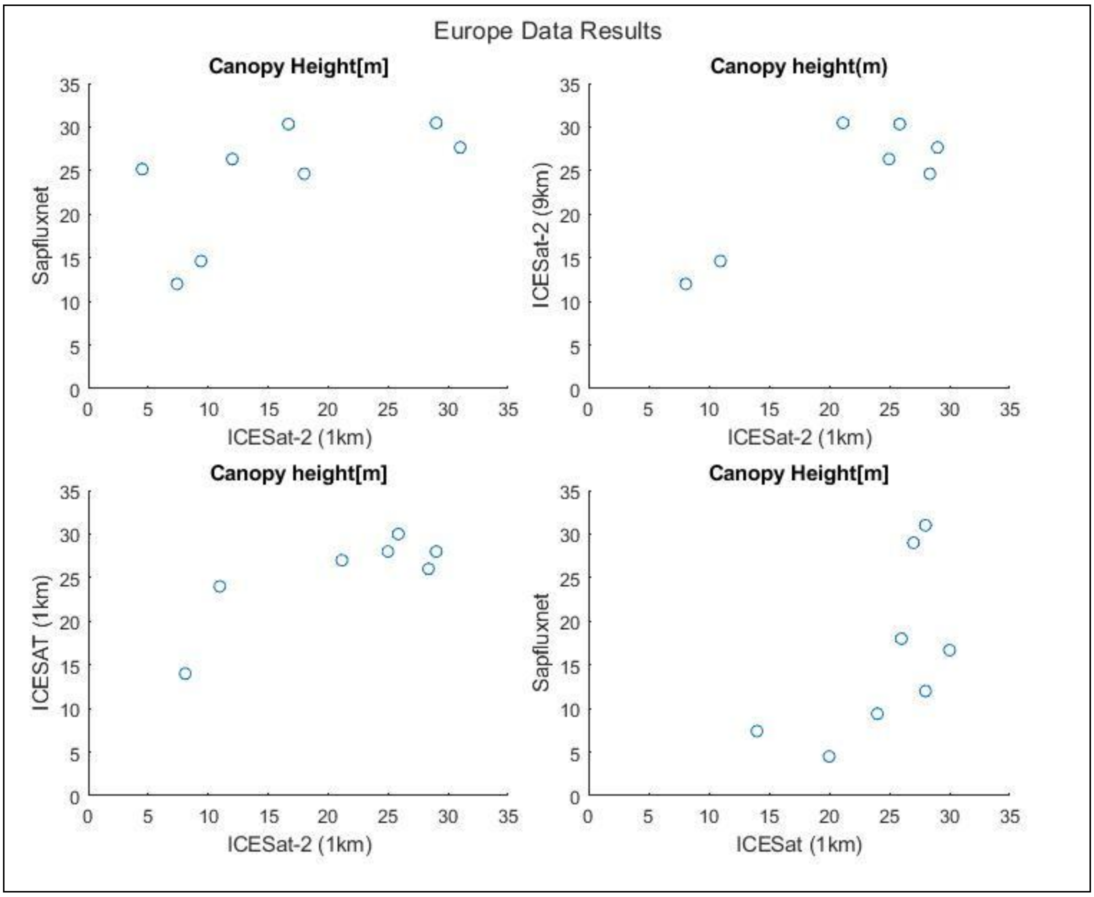

Figure 15. Results obtained for the Europe region

On the other hand, for the Region of North America, as can be seen in the scatter plots represented, there is practically no correlation between the values of the data compared, though visually a positive relationship can be observed, in exception of ICESat-2(1km) - ICESat-2 (9km) shown in the previous section.

In view of this, we suggest that the data obtained from the in-situ stations do not serve to the purpose established for the project. Therefore, the results should be further validated comparing with other in situ and/or airborne products. In addition, one of the major problems for this validation was the low number of in-situ stations devoted to vegetation height measurements.

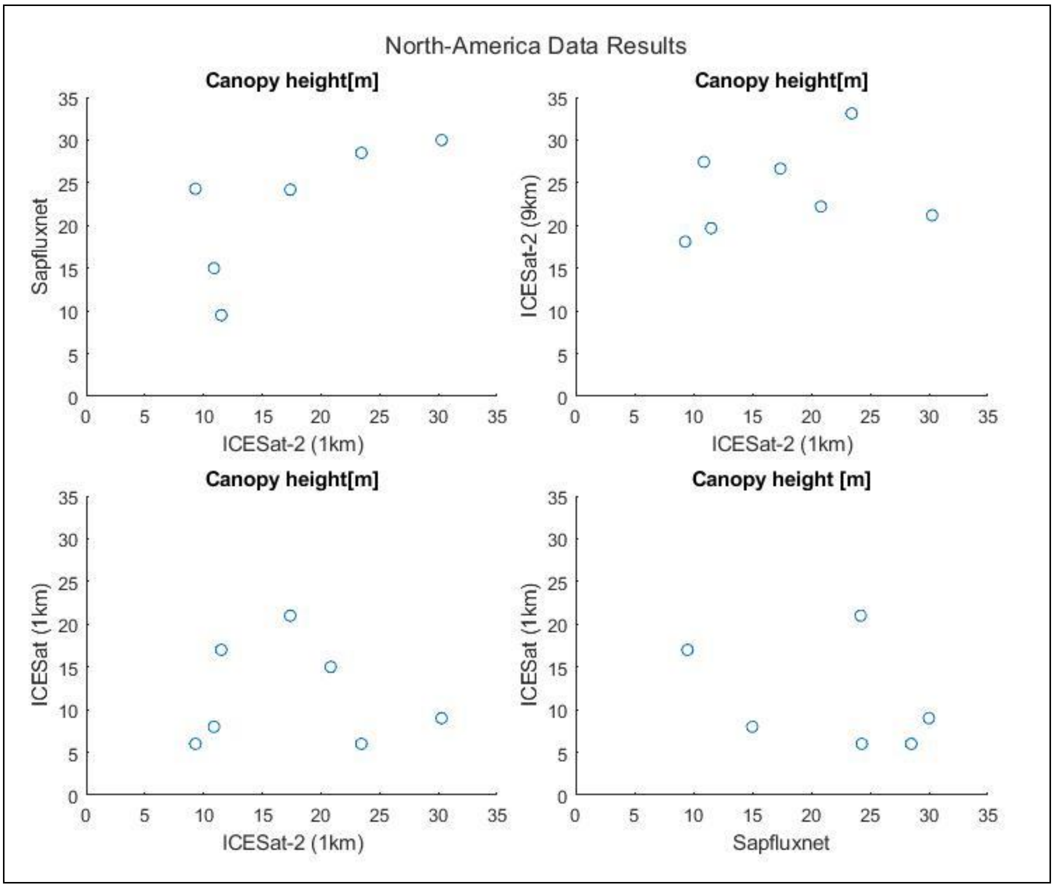

Figure 16. Results obtained for the North America region

5. Budget

The budget of this project is:

- At Hardware level. Computer: An ASUS computer, the original cost of it was 599 €
- Salary: A student of Telecommunications Engineering at the ETSETB is paid 9€/h when he/she is in a company doing a curricular internship. The total hours are 900h.
- At software level. All the software used for this project was the tool Matlab, that with the license provided form the school is free.

Items	Concept	Amount
Item 1	Hardware	599 €
Item 2	Salary	8.100 €
Item 3	Software	0
Total		**8.699 €**

Table 2. Budget.

The total cost is 8.699 €.

6. Conclusions and future work:

With this project we have produced maps of vegetation height in forested areas in Europe and in the United States using ICESat-2 satellite data. This product could complement other products obtained using GEDI or ICESat data. These maps have been obtained at 1km spatial resolution. The results presented are consistent from a biogeographical point of view, showing higher forests in more humid climates and lower trees in drier regions.

In addition, a preliminary validation work has been performed. The obtained maps have been compared with in-situ data from the SAPFLUXNET network and with ICESat canopy height data. The correlation with ICESAT for Europe is the highest obtained, showing good performance 0,74). As ICESat data had obtained good correlation with FLUXNET in-situ measurements [6], we conclude that it is necessary to look for other in-situ stations (and for a larger number of samples) in the validation process.

Future work following on this project would involve, then, an improved validation approach. Also, a pixel-by-pixel comparison between ICESat and ICESat2 data should be explored (i.e., not limiting the products comparison to the location of the in-situ stations). In addition, regarding possible future developments of this project, a monitoring the height of the vegetation studied would allow us to see if the height has increased or decreased over time, and consequently, would provide information to see the impact that deforestation, restoration and/or climate change have had in the vegetation.

Bibliography:

[1] Asner, G.P., Mascaro, J., Muller-Landau, H.C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J.S., Van Breugel, M., 2012. A universal airborne LiDARapproach for tropical forest carbon mapping. Oecologia 168, 1147–1160.

[2] Goetz, S., Dubayah, R., 2011. Advances in remote sensing technology and implicationsfor measuring and monitoring forest carbon stocks and change. Carbon Management 2, 231–244.

[3] Peter Potapov, Xinyuan Li, Andres Hernandez-Serna, Alexandra Tyukavina, Matthew C. Hansen, Anil Kommareddy, Amy Pickens, Svetlana Turubanova, Hao Tang, Carlos Edibaldo Silva, John Armston, Ralph Dubayah, J. Bryan Blair, Michelle Hofton." Mapping global forest canopy height through integration of GEDI and Landsat data". In Remote Sensing of Environment 253 (2021) 112165. Department of Geographical Sciences, University of Maryland, United States. NASA Goddard Space Flight Center, United States

[4] FLUXNET, The Data Portal serving the FLUXNET community. (6 de February 2020). Fluxnet 2015 Dataset. https://fluxnet.org/data/fluxnet2015-dataset/

[5] Nico Lang, Nikolai Kalischek, John Armston, Konrad Schindler, Ralph Dubayah, Jan Dirk Wegner. *"Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles"*. In Remote Sensing of Environment 268 (2022) 112760. EcoVision Lab, Photogrammetry and Remote Sensing, ETH, Zürich, Switzerland, Department of Geographical Sciences, University of Maryland, College Park, MD, USA, Institute for Computational Science, University of Zurich, Switzerland.

[6] Juan Carlos Fernandez-Diaz Member IEEE, Mariya Velikova, Craig L. Glennie, Member, IEEE. *"Validation of ICESat-2 ATL08 Terrain and Canopy Height Retrievals in Tropical Mesoamerican Forests".*

[7] Lonesome Malambo Sorin C.Popescu. Remote Sensing of Environment. *"Assessing the agreement of ICESat-2 terrain and canopy height with airborne lidar over US ecozones"*. Volume 266, 1 December 2021, 112711.

[8] Stefan Oehmcke, Thomas Nyegaard-Signori, Kenneth Grogan, Fabian Gieseke. 2021 IEEE International Conference on Big Data (Big data). *"Estimating Forest Canopy Height with Multi-Spectral and Multi-Temporal Imagery Using Deep Learning",* 15-28 December,2021, Orlando, FL, USA.

[9] FAO, Food and Agriculture Organization of the United Nations. Evaluación de los recursos forestales mundiales en 2020. https://www.fao.org/forest-resources-assessment/2020/es.

[10] SAPFLUXNET, SAPFLUXNET Project. A global database of sap flow measurements (1 of September 2020). https://sapfluxnet.creaf.cat/.

[11] Marc Simard, Naiara Pinto, Joshua B. Fisher, and Alessandro Baccini.*" Mapping forest canopy height globally with spaceborne lidar"* JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 116, G04021, received 17 March 2011; revised 19 August 2011; accepted 27 August 2011; published 19 November 2011.

[12] NSIDC, National Snow & Ice Data Center. ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 5. Coverage (14 October 2018-Present). https://nsidc.org/data/atl08

[13] SDAT. Spatial Data Access Tool. Geospatial Data Visualization/Download based on OGC Standards. https://webmap.ornl.gov/ogc

[14] ESA, European Space Agency. Land cover maps (1 of October 2014). Three global LC maps for the 2000, 2005 and 2010 epochs. Land cover maps | ESA CCI Land cover website (esa-landcover-cci.org)