D2.5 Report on Estimation of biomass volume at low productivity MLs (Executive Summary).

***MAIL**: Identifying Marginal Lands in Europe and strengthening their contribution potentialities in a CO_2 sequestration strategy*

MAIL project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 823805; [H2020 MSCA RISE 2018]

Project title	Identifying Marginal Lands in Europe and strengthening their contribution potentialities in a CO_2 sequestration strategy
Call identifier	H2020 MSCA RISE 2018
Project acronym	MAIL
Starting date	01.01.2019
End date	31.12.2021
Funding scheme	Marie Skłodowska-Curie
Contract no.	823805
Deliverable no.	D2.5
Document name	MAIL_D2.5_Executive Summary.pdf
Deliverable name	Report on Estimation of biomass volume at low productivity m/ms MLs
Work Package	WP2
Nature[1]	R
Dissemination[2]	CO
Editor	IABG
Authors	Jesús Torralba, Simonas Garsva and Rewanth Ravindran
Contributors	Luis A. Ruiz and Lindner Martin
Date	29-10-2020

[1] **R** = Report, **P** = Prototype, **D** = Demonstrator, **O** = Other

[2] **PU** = Public, **PP** = Restricted to other programme participants (including the Commission Services), **RE** = Restricted to a group specified by the consortium (including the Commission Services), **CO** = Confidential, only for members of the consortium (including the Commission Services).

MAIL CONSORTIUM

Aristotle University of Thessaloniki (AUTH) Greece	Industrieanlagen Betriebsgesellschaft MBH (IABG) Germany
Gounaris N. – Kontos K. OE (HOMEOTECH) Greece	Centrum Badan Kosmicznych Polskiej Akademii Nauk (CBK PAN) Poland
Universitat Politècnica de València (UPV) Spain	Fundación Centro De Servicios Y Promoción FOrestral Y de su Industria De Castilla y León (CESEFOR) Spain

ABBREVIATIONS

Term	Explanation
AGB	Above-Ground Biomass
AIC	Akaike Information Criterion
ALS	Airbone Laser Scanning
CLC	Corine Land Cover
DBH	Diameter at Breast Height
DEM	Digital Elevation Model
DSM	Digital Elevation Surface
DTM	Digital Terrain Model
ERDF	European Regional Development Fund
ESA	European Space Agency
EU	European Union
GSV	Growing Stock Volume
ha	Hectares
HRL	High Resolution Layer
IWCM	Interferometric Water Cloud Model
LiDAR	Light Detection and Ranging
Mg	Megagrams
ML	Marginal Lands
rRMSE	Relative Root Mean Square Error
RADAR	Radio Detection and Ranging
R^2	Coefficient of determination
SAR	Synthetic Aperture Radar
SNAP	Sentinel Application Platform
TCD	Tree Cover Density

US	United States
WCM	Water Cloud Model

Contents

EXECUTIVE SUMMARY

Deliverable 2.5 “Report on Estimation of biomass volume at low productivity m/ml MLs” refers to task 2.6 and it was divided into 6 chapters. The main objective of this task described in chapter 1 was to show and validate methods to quantify biomass in MLs with the use of SAR and LiDAR sensors in the reference areas provided by the ***MAIL*** consortium.

To achieve this objective, a literature review of the most frequently used methodologies with RADAR and LiDAR for AGB estimation has been carried out in chapter 2. In this chapter, it was identified that large-scale estimation of biomass and carbon content of vegetation is not simple. The traditional methods for forest biomass estimation, generally manual, are not sufficient to cover the need to have a broad and detailed knowledge of the biomass stored in natural environments (forests, shrublands, ML with vegetation, etc.). RADAR and LiDAR remote sensing sensors have the capacity to record large areas and derive from the data obtained, different forest parameters. These sensors can directly measure parameters such as height or number of individuals in a given area, but they can also indirectly estimate parameters such as wood volume, biomass, and carbon content.

To implement the different methodologies for estimating biomass on marginal lands using RADAR and LiDAR data, it was necessary to define the test areas and training data. Chapter 3 was a description of each of the pilot areas and the ground truth data acquired. This chapter explains that to carry out task 2.6 we do not have biomass data from marginal lands themselves because for biomass estimation with RADAR and LiDAR it is necessary to have biomass data derived from allometric equations using forest parameters with DBH or tree height. For this reason, only 3 forest test sites were used for this task: Espadán (Spain), Nogueruelas (Spain), and Thessaloniki (Greece), as they were the only sites with field data acquired. For the RADAR methodology, the remote sensing data used were free data provided by the ESA Copernicus Sentinel-1 and for the LiDAR methodology, the data used were acquired through a private aerial system and have been provided by the UPV partner through the project [CGL2016-80705-R] project financed by the Spanish Ministry and ERDF (European Regional Development Fund).

The biomass estimation for the test areas had been implemented using 3 different approaches explained in chapter 4.

i) *Water Cloud Method (WCM).* This method used the Sentinel-1 C-band and analyzed over a forest area the relationship between backscatter generated at the top of the forest canopy and the backscatter generated in the soil gaps. The workflow started with radiometric, and geometric correction and speckle filtering with the support of a DEM. Then the non-vegetation areas were masked with the help of the CORINE land cover (CLC). Areas defined as forest were classified into dense forest and soil with the use of the Tree Cover Density (TCD) High-Resolution Layer (HRL) from Copernicus. In these two areas, the parameters of forest backscatter and ground backscatter were estimated. Then the β values and the maximum value of GSV were selected to apply the WCM equation. Finally, the WCM values were transformed to AGB. To reduce noise in backscatter and reduce the error of AGB estimation, different polarizations and the combination of images acquired at different times of the year were analyzed.

ii) *Interferometric Water Cloud Model (IWCM).* In this method, the backscatter was identified in a similar way to the WCM method but generalized to include gaps in the vegetation cover by the introduction of the area fraction covered by vegetation. To perform the interferogram, two images were selected from September 2015, which was the date when the data was measured in the ground. The images were pre-processed to obtain the "backscatter image" and the "coherence image". On the other hand, the height of the vegetation was obtained from the subtraction of the DTM from the DSM from the LiDAR dataset that was used as "phase heights". Finally, the IWCM model was calculated and converted the obtained values to biomass values.

iii) *LiDAR.* The methodology used allows estimating the biomass at plot level from aerial LiDAR data. First, the value of the biomass at the plot level was calculated from the field data. At the same time, a pre-processing of the ALS data was carried out, removing the noise, normalizing the heights, and trimming the clouds according to the size and shape of the plots. Afterward, the height and intensity metrics ALS per plot were obtained. The different ALS metrics were analyzed using the Akaike information criterion (AIC) to select the relevant predictors for biomass adjustment. With ALS metrics as independent variables and ground truth biomass values as dependent variables, multiple linear regression models were generated for each study area and species. Finally, the accuracy of the different models was evaluated with different statistics by leave-one-out cross-validation.

To implement the SAR *WCM* methodology, Sentinel-1 C-band, CORINE land cover layers, and Tree Cover Density (TCD) High-Resolution Layer (HRL) from Copernicus were required, which are free and open access, as well as ESA's SNAP processing software. For the *IWCM* SAR methodology, Sentinel-1 C-band (free and open access), DTM, DSM (depending on the resolution is also free at different scales), and biomass field data (usually involve an acquisition cost) was required. In addition, the free SNAP software and commercial MATLAB software were also required, although it could be programmed in other free languages. To implement the *LiDAR* methodology, field data that generally have a cost, airborne LiDAR data that have an acquisition cost, or low-density point clouds that in some EU countries are free, were required. For processing the LiDAR data, LAStools and the Fusion software that is distributed free of charge by the US Forest Service were used, while the statistical analysis of the data, the free software RStudio was used.

The results for each methodology were analyzed in Chapter 4 and compared in Chapter 5. In summary, the results with the *WCM* method had a low precision, generally around 30-80% of rRMSE, mainly due to an early signal saturation – short C-band wavelength has limited penetration leading to loss of signal sensitivity at higher biomass levels (above 100 Mg/ha) under non-optimal environmental and meteorological conditions at the time of image acquisition. On the other hand, this method is the only one transferable to all of Europe, although it is very sensitive to the weather conditions in which the different images were taken. The *IWCM* method improved the precision with respect to the WCM method, reaching an rRMSE of 36% - 48.2%% for some stands. This method is more complex to implement and its transferability depends on the availability of field data. The *LiDAR* methodology was the one that obtained the better precisions. It also obtained biomass estimation equations with an R^2_{adj} of 0.69 to 0.83 depending on the test site and the dominant tree species. Regardless of the good results of LiDAR methodology, and its evident usefulness in biomass estimation, these results are not transferable to other test sites. To be transferable, LiDAR and biomass data should have been available to fit the equations to each test site and each set of species.

In Chapter 6, a series of recommendations were given for the use of RADAR and LiDAR data as well as the relevance of field data. The methodologies proposed in Task 2.6 for the estimation of biomass on marginal lands, given the spatial and temporal resolution at which these can work, can be considered as a good tool for the future monitoring and evaluation of the forest stands established in the MLs.