

CCI BIOMASS

ALGORITHM DEVELOPMENT PLAN VERSION 6.0 AGB MAPPING

DOCUMENT REF:	CCI_BIOMASS_ADP_V6
DELIVERABLE REF:	D2.4-ADP
VERSION:	6.0
CREATION DATE:	2021-06-16
LAST MODIFIED	2024-12-19

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	2	19.12.2024	

Document Authorship

	NAME	FUNCTION	ORGANISATION	SIGNATURE	DATE
PREPARED	M. Santoro	Project scientist	Gamma Remote Sensing		
PREPARED	O. Cartus	Project scientist	Gamma Remote Sensing		
PREPARED	R. Lucas	Project manager	Aberystwyth University		
PREPARED					
PREPARED					
PREPARED					
PREPARED					
VERIFIED	H. Friendship-Kay	Project Coordinator	Aberystwyth University		
VERIFIED	S. Quegan	Science Leader	Sheffield University		
APPROVED					

Document Distribution

ORGANISATION	NAME	QUANTITY
ESA	Frank Seifert	

Document History

VERSION	DATE	DESCRIPTION	APPROVED
1.0	2019-03-05	ADP of CCI Biomass year 1	
2.0	2020-04-06	ADP of CCI Biomass year 2	
3.0	2021-06-16	ADP of CCI Biomass year 3	
4.0	2023-02-09	ADP of CCI Biomass year 4	
5.0	2023-12-12	ADP of CCI Biomass year 5	
5.0	2024-12-19	ADP of CCI Biomass year 6	

Document Change Record (from Year 1 to Year N)

VERSION	DATE	DESCRIPTION	APPROVED
2.0	2020-04-06	Adapted the ADP to year 2 algorithm. Added the strategy to map AGB changes	
3.0	2021-06-16	Adapted the ADP to year 3 algorithm. Added aspects to be developed in future studies	
4.0	2023-03-30	Adapted the ADP to year 4 algorithm. Added aspects to be developed in future studies	

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	3	19.12.2024	

5.0	2023-12-12	Adapted the ADP to year 5 algorithm. Added aspects developed on biomass change.	
6.0	2024-12-19	Adapted the ADP to year 6 algorithm. This ADP focuses on AGB mapping only.	

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	4	19.12.2024	

TABLE OF CONTENTS

1. INTRODUCTION	8
2. CCI BIOMASS CORE ALGORITHM	10
3. CAVEATS OF THE CORE ALGORITHM	11
4. PROPOSED DEVELOPMENT OF CORE ALGORITHM	14
Consolidate the use of spaceborne LiDAR observations	14
Characterizing the AGB - LiDAR height model	15
Use of vegetation structural information	15
Use of coarse resolution EO data	16
5. CONCLUSIONS	16
6. REFERENCES	16

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	5	19.12.2024	

SYMBOLS AND ACRONYMS

ADP	Algorithm Development Plan
AGB	Above Ground Biomass
AFOLU	Agriculture, Forestry and Other Land Use
ALOS	Advanced Land Observing Satellite
ARD	Analysis Ready Data
ASCAT	Advanced Scatterometer
ATBD	Algorithm Theoretical Basis Document
BCEF	Biomass Conversion & Expansion Factor
CCI	Climate Change Initiative
CCI-Biomass	Climate Change Initiative – Biomass
CTV	Cultivated Terrestrial Vegetation
DARD	Data Access Requirements Document
E3UB	End to End ECV Uncertainty Budget
ECV	Essential Climate Variables
EO	Earth Observation
ESA	European Space Agency
FAO	Food and Agriculture Organisation
GCOS	Global Climate Observing System
GEDI	Global Ecosystem Dynamics Investigation
GEE	Google Earth Engine
GHGs	Greenhouse Gases
GLAS	Geoscience Laser Altimeter System
GSV	Growing Stock Volume
ICESat	Ice, Cloud, and land Elevation Satellite
JAXA	Japan Aerospace Exploration Agency
LCCS	Land Cover Classification System
NDVI	Normalised Difference Vegetation Index
NFI	National Forest Inventory
NTV	Natural/Semi-Natural Terrestrial Vegetation
OED	Overarching Environmental Descriptor
PSD	Product Specification Document
PVASR	Product Validation and Algorithm Selection Report
PVIR	Product Validation and Intercomparison Report
PVP	Product Validation Plan
SAR	Synthetic Aperture Radar

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	6	19.12.2024	

SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture & Ocean Salinity
SRTM	Shuttle Radar Topography Mission
URD	User Requirement Document
WCM	Water Cloud Model

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	7	19.12.2024	

Table 1-1: Reference Documents

ID	TITLE	ISSUE	DATE
RD-1	Users Requirements Document		
RD-2	Product Specification Document		
RD-3	Data Access Requirements Document		
RD-4	Product Validation and Algorithm Selection		
RD-5	Algorithm Theoretical Basis Document		
RD-6	End to End ECV Uncertainty Budget		
RD-7	Product Validation Plan		
RD-8	Algorithm Theoretical Basis Document of GlobBiomass project		
RD-8	Product Validation and Intercomparison Report		

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	8	19.12.2024	

1. Introduction

Above-ground biomass (AGB, units: Mg ha^{-1}) is defined by the Global Carbon Observing System (GCOS) as one of 54 Essential Climate Variables (ECV). For climate science communities, AGB is a pivotal variable of the Earth System, as it impacts the surface energy budget, the land surface water balance, the atmospheric concentration of greenhouse gases (GHGs) and a range of ecosystem services. The GCOS requirement is for AGB to be provided wall-to-wall over the entire globe for all major woody biomes at 500 m to 1 km spatial resolution with a relative error of less than 20% where AGB exceeds 50 Mg ha^{-1} and a fixed error of 10 Mg ha^{-1} where the AGB is below that limit.

One of the objectives of the Climate Change Initiative (CCI) Biomass project is to generate global maps of AGB using a variety of Earth Observation (EO) datasets and state-of-the-art models for several epochs and assess AGB changes over 1-year differences and a 10-year difference. The maps should be thematically consistent with data layers similar to the AGB datasets that are produced in the framework of the CCI Programme (e.g., Fire, Land Cover, Snow etc.).

Algorithms to estimate AGB from EO data are described in the Algorithm Theoretical Basis Document (ATBD) [RD-5] while the End-to-End ECV Uncertainty Budget (E3UB) document [RD-6] describes the precision associated with the estimates of AGB and AGB change. The ATBD and the E3UB documents are live documents, updated annually to provide a thorough description of the algorithms implemented to generate AGB and AGB change maps. The current version of the ATBD and the E3UB documents describe the CORE algorithm used to generate version 6 of the Climate Research Data Package. This consists of global datasets of AGB and related AGB change maps using data representative for the years 2007, 2010 and between 2015 and 2022.

The original CORE algorithm was based on the GlobBiomass global retrieval algorithm [RD-8] (see <http://globbiomass.org/products/global-mapping/>).

For version 2, the CORE algorithm was enhanced by expanding on concepts presented in the first version of this document. Namely, (i) the retrieval models expressed the Synthetic Aperture Radar (SAR) backscatter as a function of forest height and canopy density, (ii) models between canopy density, forest height and AGB were implemented in the retrieval models (iii) the model training accounted for the effect of local topography on the relationship between SAR backscatter and biomass. These advances were possible thanks to an in-depth analysis of the Ice, Cloud and land Elevation Satellite (ICESat) Geoscience Laser Altimeter System (GLAS) observations of canopy density and height (Kay et al., 2021), and the increasing number of publications that focus on the relationship between LiDAR height metrics and AGB. As a consequence, the CORE retrieval algorithm provided estimates of AGB instead of Growing Stock Volume (GSV) so that a Biomass Conversion and Expansion Factors (BCEF) layer becomes unnecessary.

For version 3, the CORE algorithm was consolidated with the addition of recent LiDAR observations by the Global Ecosystem Dynamics Investigation (GEDI) and the ICESat-2 missions. Also, the CORE algorithm implemented measures to avoid unnatural fluctuations of the AGB estimates. These measures, however, could not fully compensate for artefacts because of the different setting of the EO data available in 2010, 2017 and 2018. To quantify biases in each of the three maps, a model-based framework relying on the plot database available to CCI Biomass was implemented with the Plot2Map

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	9	19.12.2024	

tool (Araza et al., 2022) and coarse resolution maps of AGB bias (0.1°) were generated. The bias layers are supposed to build confidence on the reliability of the map rather than to represent a correction factor to be applied straight to the AGB estimates, also because of the much poorer pixel spacing (10,000 ha vs. 1 ha). The AGB change maps derived from the Year 3 dataset were based on AGB differencing rather than signal differencing because of the multi-sensor approach pursued in this project. Given that AGB changes were assessed on maps of different quality and only for three epochs, the approach was preliminary.

For version 4, the estimation of AGB relied on annual multi-temporal observations of L-band SAR backscatter, which replaced the annual mosaics (i.e., a single observation) and on more extensive datasets from spaceborne LiDAR missions. LiDAR data, together with a large database of AGB statistics published by National Forest Inventories (NFIs), allowed a more accurate characterisation of the model that expresses height as a function of AGB. With such a model, systematic retrieval errors, due for example to an incorrect characterisation of the maximum AGB in a region, could be alleviated. Indeed, we identified this parameter as causing significant biases and thus being a major issue in previous versions of the Climate Research Data Package (CRDP). The retrieval models based on the BIOMASAR approach evolved towards a more precise characterisation of the parameters in the Water Cloud Model (WCM) relating AGB to SAR backscatter. The retrieval was also relaxed in regions with sloping terrain because the SAR data had higher radiometric quality than in previous project years. In addition, the merging rules for BIOMASAR-C and -L AGB were revisited to better account for their mutual contribution. The availability of a time series of AGB estimates from each of the approaches allowed for more robust merging rules to be defined.

The estimation of AGB change did not depart from its original formulation, i.e., a map differencing approach. The assessment of AGB change maps based on AGB differences with a time series of maps created with state-of-art retrieval techniques was the overall objective of algorithmic advances in the AGB change mapping for version 4.

For version 5, the estimation of AGB was consolidated. Multi-temporal SAR acquisitions were available throughout the entire interval foreseen for AGB mapping. Also the spaceborne LiDAR dataset was populated with more recent measurements to increase the spatial density. Validation of the CRDP v4 confirmed that the maximum AGB, i.e., the estimates of local maximum canopy height and the height-to-AGB model, had great impact on the level of the AGB estimates. For this version, the fit of this model was revised. We also approached several aspects pointed at in the previous version of this document, namely, the retrieval errors due to banding, topography and stratification by vegetation type. While the first two cannot be corrected for because caused by artefacts in the original SAR datasets, a mangrove-specific model training was introduced for this version. AGB change still relied on the differencing method.

For this version, the algorithmic developments on biomass estimation and biomass change estimation have been split into two documents. This document focuses on the algorithm development of biomass estimation. The CORE retrieval algorithm has reached maturity and changes from v5 were minor. Training of the structural function relating canopy height to AGB and the calibration of the Water Cloud Model have been revised following a review of estimation errors as revealed by the validation of the AGB maps. The merging of the BIOMASAR-C and -L AGB estimates has been consolidated thanks to the larger amount of estimates in time compared to previous versions of the CRDP.

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	10	19.12.2024	

Some ideas to be pursued in future activities are presented in this document. Such ideas involve both the estimation of AGB and the estimation of AGB over time to track changes, as it is believed that a multi-sensor approach to estimating AGB is superior to using a single set of observations. With the multi-sensorial approach, it is not possible to relate a change in AGB to a change in signals.

This document builds on the ATBD and E3UB documents for v6 to identify major elements that require development in future years of the CCI Biomass project. In addition, we consider the review of the CCI BIOMASS data products of v5 reported in the Product Validation Report (PVR) [RD-8]. As for the ATBD and the E3UB documents, this Algorithm Development Plan relies on the most recent versions of the Users Requirements Document (URD) [RD-1] and the Product Specifications Document (PSD) [RD-2].

Section 2 reviews the CCI Biomass CORE algorithm implemented in v6. Section 3 elaborates on the known major weaknesses of the CORE algorithm based on the initial assessment of AGB retrieval reported in the ATBD. Section 4 lists potential solutions to the issues identified in Section 3. Advancing the estimation of AGB change based on the experiences gathered with the AGB data products foreseen by the CRDP of the CCI Biomass project is the topic of Section 5.

2. CCI Biomass CORE algorithm

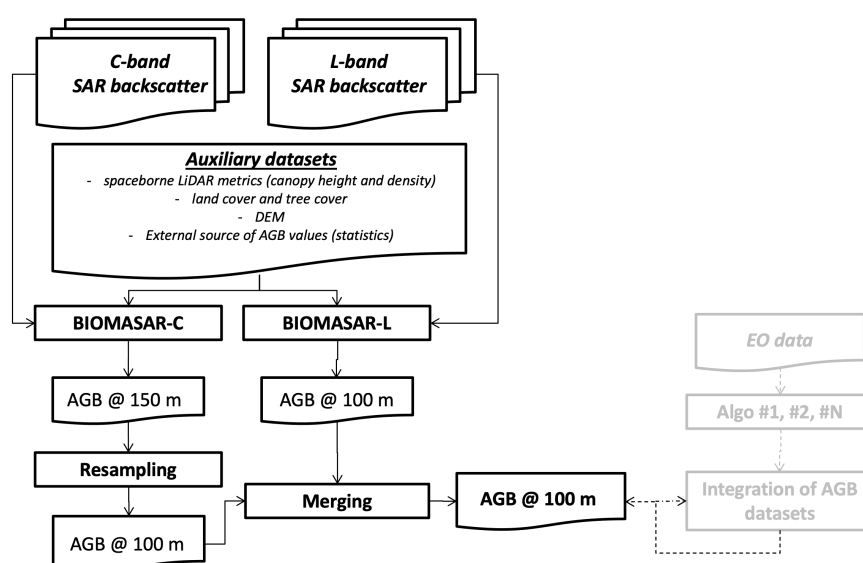


Figure 2-1: Functional dependencies of datasets and approaches forming the CCI Biomass CORE global biomass retrieval algorithm. The shaded part of the flowchart represents potential improvements following the implementation of additional retrieval techniques [RD-3].

Figure 2-1 shows the flowchart of the CORE biomass estimation procedure of the CCI Biomass project to generate annual, global datasets of AGB estimates [RD-5]. The shaded part of the flowchart represents potential improvements following the implementation of additional retrieval techniques. [RD-5].

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	11	19.12.2024	

With the CORE algorithm, two independent estimates of AGB are obtained from the same BIOMASAR algorithms but with different modelling frameworks. The SAR backscatter is related to canopy density and height with a WCM, i.e., a parametric model that simplifies the scattering in the canopy and below the canopy with a few parameters and variables (canopy density and canopy height). A simple model, trained with LiDAR data, is used to relate these variables. A second model, linking height and AGB, is then used to express the SAR backscatter directly as a function of AGB. Linear weighting of AGB estimates obtained from the inversion of the WCM and single backscatter observations is applied to generate a final estimate of AGB.

Starting with v5, the SAR datasets consisted of the best possible setting of (freely available and global) images. The Sentinel-1 dataset was consolidated in the form of monthly averages to speed up computation and reduce redundancies. The Advanced Land Observing Satellites (ALOS) -1 and -2 SAR datasets have been provided by the Japanese Aerospace Exploration Agency (JAXA) in the form of individual strips for the Fine Beam mode and per-cycle mosaics of 46 days for the ScanSAR mode. Each location is characterised by multiple dual-polarised observations as opposed to a single dual-pol observation from the annual mosaics used until v4.

Following the approach that was started in v3, the CORE algorithm makes even more explicit use of laser observations in the retrieval model and follows a consolidated line of research aiming at relating LiDAR-based canopy height metrics to AGB measurements rather than to AGB estimates from maps. Also, the retrieval still accounts for the dependency of the scattering on the look direction of the radar and the ground surface by using experimental relationships between incidence angle and the SAR backscatter. This approach was preferred to the development of models that would have probably failed due to the subtle difference in backscatter as landscape and topography change. Finally, the estimation of the model parameters implements a robust model calibration approach consisting of a blend of self-calibration and least squares regression with respect to a reference dataset of canopy density. Merging of AGB estimates from BIOMASAR-C and BIOMASAR-L exploits the time series of AGB estimates from each approach to construct a set of merging rules of increased robustness with respect to the weights used in previous versions of the CRDP. Quantitative assessment of the results achieved with the CORE algorithm is presented in the PVR.

3. Caveats of the CORE algorithm

The above brief summary of the CCI Biomass CORE algorithm highlights the major elements of the retrieval approach. This may not be the best possible algorithm but rather is a global approach constrained by the available EO data and ground observations. The CCI Biomass CORE algorithm relies on several assumptions that appear viable when comparing large-scale averages of estimated AGB with corresponding values based on inventory information [RD-5] and [RD-7]. Nevertheless, these assumptions, which were made to allow the CORE algorithm to perform globally, also introduce systematic errors into the retrieved AGB, which may become apparent when focusing on particular areas [RD-4], [RD-5] and [RD-7].

Here, we provide a list of caveats and potential areas of improvement of the CORE algorithm. These are then expanded in Section 4 with a proposed development of the CORE algorithm.

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	12	19.12.2024	

- The retrieval of AGB implemented in v1 was found to be rather conservative because it missed the extreme values of AGB. One of the reasons was that the retrieval models were canopy-centric and did not explicitly involve height information. In v2, we exploited height information in the form of models, with interesting preliminary results. The models were based on ICESat GLAS metrics, which did not provide a uniform sampling of all land masses on Earth and required us to be rather generic in the way the models could describe the relationship between canopy density, height and AGB. With the denser coverage of GEDI and ICESat-2, the models between AGB and tree height were further characterized in v3. The impact of the models on the AGB maps was substantial, reducing the overestimation in the low AGB range and underestimation in the high AGB range. Both GEDI and ICESat-2 data products were still under development, which led to moderate usage in v3. In v4, the interaction with the data production teams and progressive ingestion of new data releases improved the models and, thereof, the auxiliary datasets used by the retrieval algorithms (e.g., the maximum AGB). The data were further investigated in v5 to understand the impact of errors on the map products and additional filters were implemented to prevent that macroscopic errors would generate biases. More extensive ICESat-2 datasets and additional filtering options to remove systematic errors allowed in v6 to further characterize the spatial distribution of canopy height, which implied a more accurate set of structural functions and a detailed dataset of maximum AGB. As some of the filters applied to the ICESat-2 data are still a research item, it is foreseen that future versions of the datasets have an impact on the accuracy of the AGB estimated with the CORE retrieval algorithm.
- The AGB retrieval model uses two sets of models to link the SAR backscatter (predictor) to the response variable (AGB). These are under continual development as more data suitable for training the allometric models become available.
 - The model that expresses canopy density as a function of canopy height is based on LiDAR observations. As of v6, the CORE retrieval algorithm still implements the model trained on ICESat GLAS data. The GLAS dataset is strongly filtered to ensure a correct estimation of the parameters of this structural model. The consequence is an uneven characterisation of these parameters because the density of the footprints was highly variable. GEDI data are the only alternative because both canopy density and canopy height are provided as part of the Level 2 datasets, whereas the ATL08 product based on ICESat-2 data only contains canopy height. Investigation of the relationship between canopy height and canopy density observations by GEDI revealed inconsistencies in the observations of canopy cover, which requires further investigations. The major limitation of a model based on GEDI data is the impossibility to characterise it throughout the boreal zone because the coverage of GEDI is limited to latitudes between $\pm 52^\circ$. To overcome this, measures need to be sought that harmonise models from ICESat GLAS and GEDI. Even though the data from the two missions were acquired during two different decades, we assume that the model is time-invariant.
 - The structural function that expresses canopy height as a function of AGB was based on spaceborne LiDAR observations and estimates of AGB from a map until v3. Several measures were implemented to limit the impact of the uncertainties affecting the map-based values of AGB on the model. Still, if any systematic error in the form of a bias affected the AGB estimates, these propagated to the estimates of the coefficients of the allometric function. Starting with v4, we used a more extensive set of LiDAR

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	13	19.12.2024	

observations than in previous versions of the CORE retrieval algorithm and explored a new way to characterise the model by relying on AGB observations rather than on AGB estimates. The AGB observations consisted of average values reported by NFIs at the level of administrative or ecological units and were related to average values of canopy height from spaceborne LiDAR data for the same units. To characterise the model in space, the data were grouped into regions. This approach was found to be promising. While the refinement of the regions improved the overall characterization of the model globally, we still identify several caveats that need to be addressed in future versions of the CORE retrieval approach. The NFI statistics are not harmonised with each other, and the definition of forest land underlying the average values reported by the NFIs and used here to select the LiDAR footprints is not harmonised. The strata used to group observations were based on some macro-ecological patterns, which cancels out small scale variability of the relationship between height and AGB, for example due to spatial variability of wood density or growth factors. In addition, the use of average values instead of the original ones measured at plots and footprints might alter the shape of the model, leading to over- or under-estimates in the retrieved AGB. This aspect is difficult to assess because the NFI data used to generate the AGB statistics are not publicly available.

- The retrieval of AGB is based on some simplifying assumptions that cause the retrieval models to be too general to capture the spatial variability of the relationship between the radar observations and vegetation properties. Vegetation structural information should provide the backbone for a more targeted estimation of model parameters. Unfortunately, most EO-based datasets that could complement a retrieval do not have a full error characterisation so that the impact of a direct implementation in our retrieval schemes may not be controllable.
- Regarding alternative approaches to retrieving AGB from the set of observations currently available from spaceborne sensors, we could not identify yet approaches that may improve our retrievals while at the same time fulfilling the requirements in terms of spatial resolution, temporal coverage and global representation of the CCI biomass maps. Non-parametric approaches based on machine learning or artificial intelligence are not targeted because they would not be supported by a dense and large range of AGB observations.
- A wide range of observations is, in our opinion, fundamental to avoid systematic biases caused by the fact that no remote sensing observation is a direct measure of biomass. One line of research that has been developing quickly in recent years is inversion of coarse-resolution observations from spaceborne microwave radiometers and scatterometers to AGB. Although such observations do not match the requirement on spatial resolution of the CCI Biomass maps, data from radiometer and scatterometer missions cover several decades and have been demonstrated to allow characterisation of biomass dynamics. As such, experiences gathered at coarse resolution may act as guidelines in the process of establishing rules to ensure that the dynamics of AGB obtained from less frequent high-resolution EO data are well captured.
- Finally, regardless of the procedures here developed to estimate AGB, the accuracy of the retrieval depends strongly on the quality of the EO data used as predictors. We have identified a number of systematic issues in the SAR data that prevent us obtaining the highest possible

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	14	19.12.2024	

quality AGB results. It is believed that having the possibility to pre-process the EO data would allow such quality to be attained. Hence continual interaction with data providers is needed.

4. Proposed development of CORE algorithm

Consolidate the use of spaceborne LiDAR observations

Observations that sense forest structure are of major benefit to the estimation of AGB. Unfortunately, the majority of EO data available globally is in the form of energy reflected to the sensor, so that AGB can only be inferred with parametric or non-parametric approaches (Santoro and Cartus, 2018). SAR interferometry and laser scanning instead generate observations that contain information on the vertical and horizontal distribution of vegetation, thus providing a more direct measure of parameters involved in the computation of biomass (canopy height, density of canopy).

The TanDEM-X and SRTM missions were conceived to acquire interferometric datasets that would allow the generation of surface elevation models (Farr et al., 2007; Krieger et al., 2007). Over forested terrain, an estimate of vegetation height can be inferred from the surface elevation if the terrain elevation is known. To obtain the true vegetation height, an additional step that compensates the InSAR-based height of the vegetation for the penetration of microwaves into the canopy is required (Walker et al., 2007). Although high resolution, accurate surface elevation models based on interferometric data exist, there is no global dataset of terrain elevation, which hinders the use of interferometry for a “direct” measure of the vegetation vertical structure. It will not be until the BIOMASS mission is flying that estimates of ground elevation may be possible (Quegan et al., 2019), although the coverage will not be global (Carreiras et al., 2017) and will be at a coarser spatial resolution than the CCI Biomass products (Quegan et al., 2019). To the best of our knowledge, there is no spaceborne mission planned that can provide a global estimate of terrain elevation.

Laser instruments also measure the elevation of the Earth surface and, in the case of vegetation, return a profile of reflection intensity along the vertical direction. The GLAS instrument on-board the ICESat satellite operated between 2003 and 2009 and recorded millions of waveforms along its orbital path. Unlike interferometric datasets, the signal recorded by a laser instrument contains also a ground return, so that an external dataset of terrain elevation is not required to estimate the height of vegetation. Waveform information in the GLA14 product was processed globally in the GlobBiomass project [RD-8] from which canopy density and several height percentiles were computed. A GLAS footprint has an approximately 70 m diameter and footprints were acquired sequentially along an orbit; however, the distance between orbits was around 60 km, leading to a sparse sampling of the Earth’s vegetation. For this reason, it is preferred to use the GLAS datasets of canopy height and canopy density to derive models in support of the retrieval model relating SAR backscatter and AGB rather than as surrogate reference data for model training.

Since 2018, the GEDI and ICESat-2 laser systems have been providing observations with a much denser coverage of the Earth land masses than ICESat GLAS. Because of the global coverage, the ICESat-2 data are used operationally in the CORE retrieval algorithm. However, in spite of the much denser coverage, our retrieval approach does not foresee estimation of AGB based solely on the LiDAR observations as this is already taken care of, for example by the GEDI team. Our understanding is also that retrieval of

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	15	19.12.2024	

AGB should combine multiple observations from spaceborne SAR, optical and laser observations and exploit the information content on AGB in each set of observations.

The data providers warn about the use of some of their measurements (Neuenschwander and Pitts, 2019; Dubayah et al., 2020) in early data versions. With the further advances of processing routines by the data providers, the accuracy of the laser measurements will improve. Another reason for following closely the development of data products by the GEDI and ICESat-2 teams is their interest in releasing global datasets of forest variables, including AGB. Recent estimates of AGB based on GEDI are available either at footprint level or as aggregated values in 1 x 1 km² large grid cells. A pan-boreal dataset of AGB where ICESat-2 observations were used in combination with Landsat images has been recently released. The Biomass Harmonization activity is currently assessing CCI Biomass and GEDI AGB products to create knowledge and allow for improvements of the individual data products.

Characterizing the AGB - LiDAR height model

In the CORE algorithm developed since Year 2, we have introduced models linking AGB with top-of-canopy height in the WCM. The characterisation of this power-law function was based on the ICESat GLAS top-of-canopy height measurements (RH100) and the GlobBiomass AGB dataset. Although the trend between AGB and RH100 was, on average, similar to results based on measurements at local scale, there is substantial work needed to: (i) reduce uncertainties and (ii) improve the spatial characterisation of the model parameters. Studies at local sites allow determination of precise models, but these may not be generalisable to larger areas. Remote sensing maps, in contrast, allow us to obtain a region-wide perspective on how height and AGB are related but these relationships may be locally inaccurate. The availability of dense sets of LiDAR observations of RH100 (and in general, different height metrics) from GEDI and ICESat-2 allowed a more detailed characterisation of AGB-to-height model, which however suffered from the early versioning of the data, implying that some height ranges may exhibit deficiencies. While the accuracy of the ICESat-2 and GEDI datasets will improve, there is a need to understand how well we can characterise the model spatially. Here, we identify local models, such as those developed in the context of CCI Biomass from airborne laser datasets and plot inventory data [RD-5], as a diagnostic tool for the map-based model. However, in regions poorly covered by LiDAR observations, it will still be impossible to quantify the reliability of the map-based model.

Use of vegetation structural information

One of the limitations of the currently implemented BIOMASAR algorithms is the coarse representation of vegetation structure. In Year 1, some of the model parameters were estimated after stratifying the world by the Food and Agriculture Organisation (FAO) ecological zones. In Year 2, we introduced a finer stratification based on subdivisions of 883 ecoregions to characterise the relationship between canopy density and RH100 but still used ecological domains to characterise the relationship between RH100 and AGB. Vegetation structural information developed in the DARD [RD-3] should provide more targeted estimation of model parameters and models.

In the same vein, knowledge gathered by investigating the relationship between EO observables and AGB in specific forest classes should be exploited. When evaluating the GlobBiomass and the CCI Biomass map (Year 1) in mangrove forests, the specific scattering mechanisms occurring at C- and L-band were not correctly accounted for in the retrieval model. The AGB of mangroves was often

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	16	19.12.2024	

underestimated because the absorption of microwaves in the canopy leads to low backscatter. The same reasoning applies to plantation forests. The reliability of the AGB map products is unknown because the validation activities have not covered such vegetation types due to the lack of suitable data available to the validation team.

Use of coarse resolution EO data

From the analyses reported in previous validation reports, estimation of AGB of high AGB forests still needs to be improved. Observations from coarse resolution sensors operating at C- and L-band, such as Advanced Scatterometer (ASCAT), Soil Moisture and Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP), have tremendous potential to improve AGB estimates. However, these datasets have a spatial resolution that ranges between 25 km and 50 km. It is unclear whether estimates at such coarse resolution can be transferred to 1 ha. In this respect, the experiences by the soil moisture community concerning the re-scaling of coarse resolution soil moisture fields to high resolution maps could inform a similar strategy when estimating AGB.

5. Conclusions

The development of the CORE retrieval algorithm of the CCI Biomass project has implemented several aspects presented in the previous versions of this document. The current CORE algorithm has reached maturity, in the sense that it can be applied to generate AGB maps for any year provided that the set of radar backscatter measurements are available. However, this does not imply that the AGB estimates are free from errors, given that the retrieval relies on observations that only see a portion of the forest biomass (above ground) and the inversion models implement several assumptions that tend to generalize the response of the radar backscatter to AGB.

We see two major developments that may further improve the accuracy of the retrieval, beyond the improvements already achieved in the first five years of the CCI Biomass project:

- Consolidation of LiDAR observations in the CORE retrieval algorithm.
- Integration of coarse resolution and high resolution EO datasets

The former will provide a more solid baseline for the models implemented in the retrieval model. The latter will increase the reliability of the AGB estimates in time and improve the accuracy of the AGB estimates in forests with the highest AGB densities ($> 300 \text{ Mg ha}^{-1}$).

Although not directly used in the retrieval algorithms, plot inventory measurements have a fundamental role in characterising spatial errors in AGB estimates by modelling biases. The modelling of biases was prototyped but needs further development.

6. References

Araza, A., de Bruin, S., Herold, M., Quegan, S., Labriere, N., Rodriguez-Veiga, P., Avitabile, V., Santoro, M., Mitchard, E.T.A., Ryan, C.M., Phillips, O.L., Willcock, S., Verbeeck, H., Carreiras, J., Hein, L., Schelhaas, M.-J., Pacheco-Pascagaza, A.M., da Conceição Bispo, P., Laurin, G.V., Vieilledent, G., Slik, F.,

	Ref	CCI Biomass Algorithm Development Plan v6		
	Issue	Page	Date	
	6.0	17	19.12.2024	

Wijaya, A., Lewis, S.L., Morel, A., Liang, J., Sukhdeo, H., Schepaschenko, D., Cavlovic, J., Gilani, H., Lucas, R., 2022. A comprehensive framework for assessing the accuracy and uncertainty of global above-ground biomass maps. *Remote Sensing of Environment* 272, 112917. <https://doi.org/10.1016/j.rse.2022.112917>

Carreiras, J.M.B., Quegan, S., Le Toan, T., Ho Tong Minh, D., Saatchi, S.S., Carvalhais, N., Reichstein, M., Scipal, K., (2017). Coverage of high biomass forests by the ESA BIOMASS mission under defense restrictions. *Remote Sensing of Environment* 196, 154–162. <https://doi.org/10.1016/j.rse.2017.05.003>

Di Gregorio, A. (2005). Land cover classification system: classification concepts and user manual: LCCS (Vol. 2). Food & Agriculture Organization.

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., Qi, W., Silva, C., (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing* 100002. <https://doi.org/10.1016/j.srs.2020.100002>

Farr, T., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. and Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Review of Geophysics*, 45 (RG2004), 1-33.

Kay, H., Santoro, M., Cartus, O., Bunting, P., Lucas, R., (2021). Exploring the relationship between Forest Canopy Height and Canopy Density from spaceborne LiDAR Observations. *Remote Sensing* 13, 4961. <https://doi.org/10.3390/rs13244961>.

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M. and Zink, M. (2007). TanDEM-X: a satellite formation for high-resolution SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 45 (11), 3317-3341.

Neuenschwander, A. and Pitts, K. (2019). The ATL08 land and vegetation product for the ICESat-2 Mission. *Remote Sensing of Environment*, 221, 247-259.

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D.H.T., Lomas, M., D'Alessandro, M.M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Smallman, T.L., Soja, M.J., Tebaldini, S., Ulander, L., Villard, L., Williams, M., (2019). The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment* 227, 44–60. <https://doi.org/10.1016/j.rse.2019.03.032>

Santoro, M. and Cartus, O. (2018). Research pathways of forest above-ground biomass estimation based on SAR backscatter and interferometric SAR observations. *Remote Sensing*, 10 (4), 608, doi:10.3390/rs10040608.

Walker, W. S., Kelndorfer, J. M., LaPoint, E., Hoppus, M. and Westfall, J. (2007). An empirical InSAR-optical fusion approach to mapping vegetation canopy height. *Remote Sensing of Environment*, 109 (4), 482-499.