Technical Note : AGB and TH estimation from SMOS LVOD

Project code SO-TN-CB-GS-0108

Version 3.0

Date 2023/11/02

	Role	***Name***	***Date and signatures***
Written by :	Project engineer Project engineer Project engineer Project engineer Project engineer Lead investigator	S. Boitard A. Mialon C. Salazar N. Roriguez-Fernandez P. Richaume Y.H. Kerr	
Approved by :			
Approved by :			

DOCUMENT STATUS SHEET

Version /Rev .	Date	Pages	Changes	Visa
3.0	Nov. 2023	All pages	• last version • Added the description of the Tree Height product generation • Updated Table 2 to detail version numbers	
3.0	May. 2023	13-16	• Updated Figures 8, 9 and 10 with the v4 of CCI 2018 • Updated Table 2 to detail version numbers • Updated section 3.2: yearly averages are filtered for outliers • Updated section 2.2.1: Use of ESA CCI v4 released in 2023	
1.0	Feb. 2023	18	• added product DOI and modified FTP address	
	Jan. 2023		• First version	
0.0	Dec. 2022		• Starting doc	

DISTRIBUTION LIST

ESTEC/EEM:	

ESRIN:	
ESTEC:	

CNES:	
SAG:	
OTHER:	

ACRONYMS

- AGB	:	Above Ground Biomass
- ATBD	:	Algorithm Theoretical Basis Document
- CCI	:	Climate Change Initiative
- CATDS	:	Centre Aval de Traitement des données SMOS
- CNES	:	Centre National d'Etudes Spatiales
- EASE	:	Equal-Area Scalable Earth
- ESA	:	European Space Agency
- ISEA	:	Icosahedral Snyder Equal Area
- L2	:	Level-2
- LMEB	:	L-band Microwave Emission of the Biosphere
- RFI	:	Radio-Frequency Interference
- SMOS	:	Soil Moisture and Ocean Salinity
- STD	:	Standard Deviation
- TB	:	Brightness Temperatures
- TH	:	Tree Height
- VOD	:	Vegetation Optical Depth

Contents

1 Introduction

Monitoring Above Ground Biomass (AGB) and Tree Height (TH) is of prime importance to track the vegetation health and estimate the carbon stocks, among other applications. The Vegetation Optical Depth (VOD) derived from the multi-angular Brightness Temperatures (TB) measured by the Soil Moisture and Ocean Salinity (SMOS) mission can be processed in order to get global AGB and TH estimations [1].

This technical note describes the processing of the SMOS VOD in L-Band (1.41 GHz) to generate yearly and global AGB and TH maps. The methodology initiated by [1] and further exploited by [2] is applied to the LVOD in the SMOS L2 v700 products coupled with various reference AGB and TH maps. Ultimately, AGB and TH time series covering the life span of SMOS (2010- still ongoing) are created. The estimated global and annual parameter (AGB or TH) maps and their uncertainties are freely provided as NetCDF files. Each file holds the parameter and associated std for all years covered by SMOS for a particular reference map. They are freely distributed by the CATDS - Centre Aval de Traitement des Données SMOS so that users can benefit from them in their studies on e.g. vegetation state monitoring, carbon stock estimation, land cover and land use or any other relevant topics. A paper ([3]) is in preparation and will describe the process in further details.

The purpose of this technical note is to describe this new dataset, its format file content and to give information about the processing. In the following sections, the input data to produce the NetCDF products are described. Then details about the methodology used are given. Finally, the last section describes the structure and content of the files.

The authors would be glad to receive any feedback or suggestions to help improve the products. Do not hesitate to reach out to Arnaud Mialon, Nemesio Rodriguez-Fernandez or Simon Boitard.

2 Input Data

Estimating AGB or TH from SMOS LVOD requires 2 datasets. The first one is the SMOS LVOD and the second one is a reference map of the parameter of interest.

2.1 SMOS LVOD

SMOS is a satellite mission launched in November 2009. The European Space Agency (ESA) and the Centre National d'Etudes Spatiales (CNES) are responsible for its operation. The payload is a passive L-band 2-D interferometer that measures the Earth radiometric emission at L-band. Over continental surfaces, the main objective of SMOS is to derive the soil moisture on a daily basis with a 3-day revisiting time. For a given point, the instrument measures brightness temperatures (TB) at various incident angles. These measurements given as input to the L-MEB radiative transfert model (*L-band Microwave Emission of the Biosphere*) allow the retrieval of both the soil moisture and the VOD. The latter represents the vegetation contribution to the measured TB. Indeed, the vegetation not only emits its own radiation but also scatters the radiation emitted by the underlying surface.

As input to the estimation processing chain, the LVOD from the SMOS Level-2 (L2) MIR_SMUDP2 v700 products were used. Each SMOS L2 MIR_SMUDP2 binary product contains the LVOD (amongst other variables including latitude, longitude, soil moisture, probability of radio frequency interference, probability distribution of Chi2 and quality flags) for a half orbit (ascending or descending) on the ISEA grid. Figure 1 gives an example of the VOD in an input L2 product.

The L2 MIR_SMUDP2 v700 products can be freely downloaded from the ESA SMOS Online Dissemination Service (*https://smos-diss.eo.esa.int/oads/access/*). The SMOS retrieval algorithm is described in [4]. For a complete description of the L2 processor, users are advised to refer to [5]. The product specifications can be found in [6].

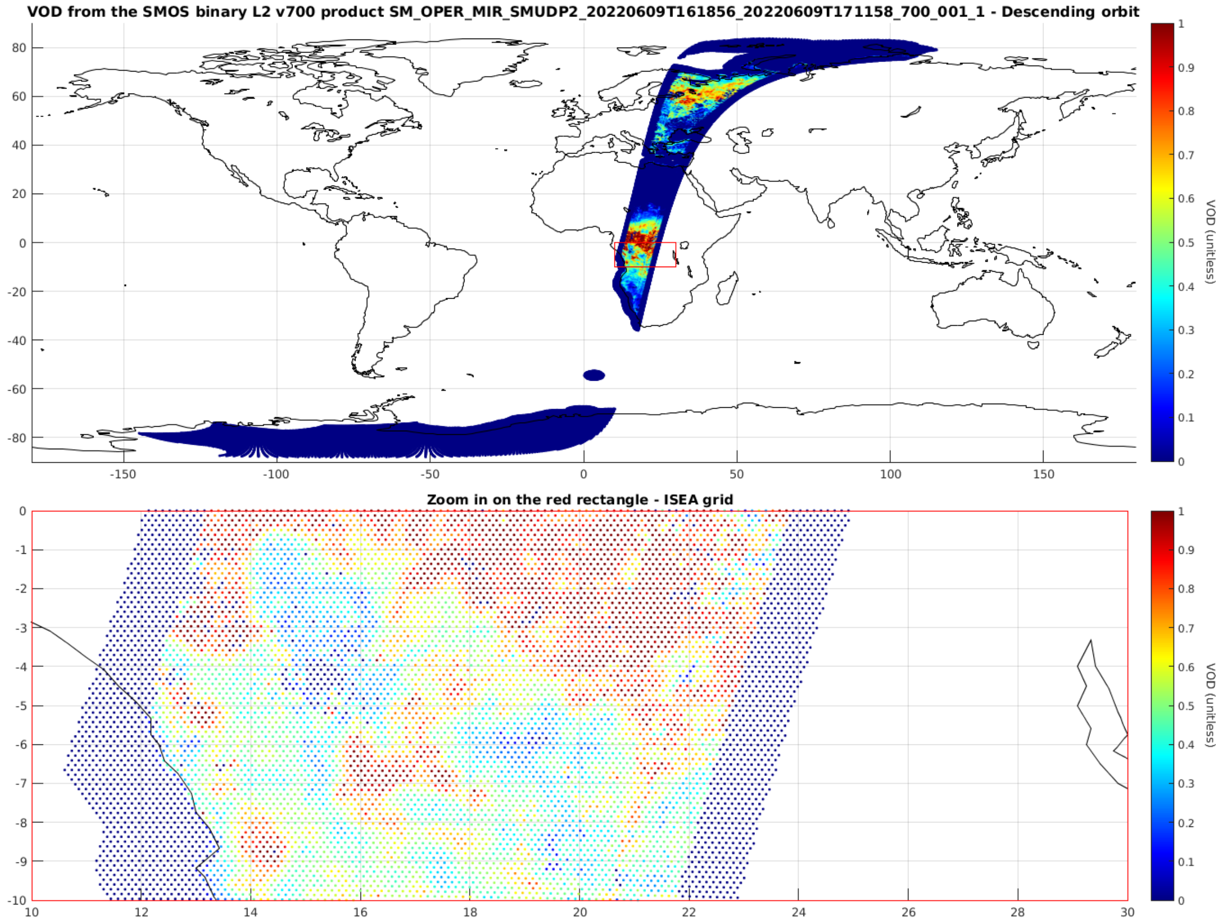

Figure 1: Example of the VOD on a descending pass from one L2 SMOS MIR_SMUDP2 v700 product on the ISEA grid - top: full half orbit, bottom: zoom on the grid over Central Africa.

2.2 Auxiliary data reference maps

2.2.1 *ESA Biomass CCI 2018 AGB map*

The ESA Biomass Climate Change Initiative (CCI) 2018 AGB map ([7]) is an output of the ESA CCI Biomass project. For each 100mx100m pixel, this product gives an estimation of the AGB in $Mg.ha^{-1}$. To reach this high resolution, SAR, LiDAR and optic remote sensing data were combined. Figure 2, extracted from the CCI AGB map product user guide ([8]), shows the AGB map for the year 2017. Readers will find more details about the CCI AGB maps in the above-mentioned product user guide. A first AGB estimation product was generated with version 3. This product was then updated version 4 of the map upon its release in 2023 (see also Table 2).

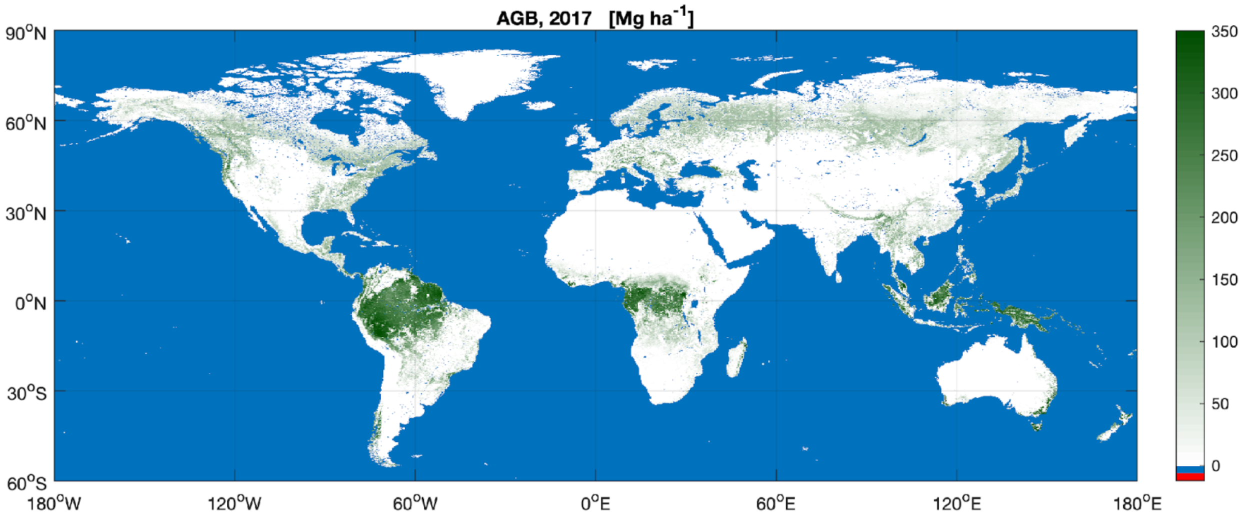

Figure 2: Global AGB estimates for the year 2017. Spatial resolution: 100 m. Figure extracted from [8].

2.2.2 *Avitabile pantropical AGB map*

This pantropical map ([9]) gives an AGB estimate for each 1kmx1km pixel. It was produced by merging two AGB datasets through machine learning techniques. This map is representative of the years 2011 and 2012. Figure 3, extracted from [9], shows the AGB map.

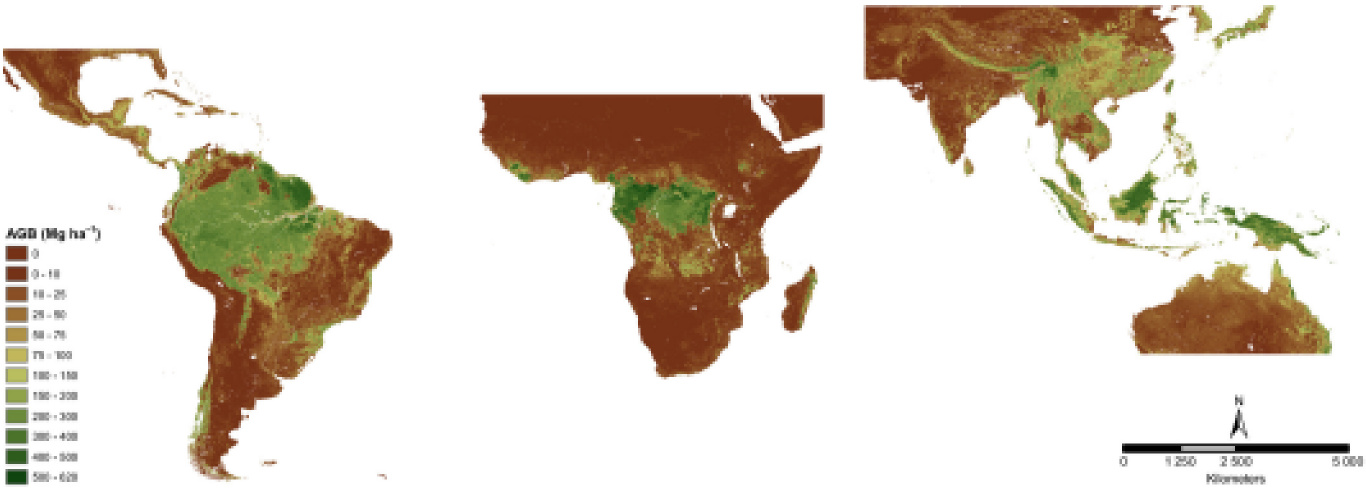

Figure 3: Avitabile AGB pantropical map at 1-km resolution. Figure extracted from [9].

2.2.3 *Lang et al TH map*

This canopy height map was produced thanks to a probabilistic deep learning model which combines sparse and accurate measures from the Global Ecosystem Dynamics Investigation (GEDI) space-borne LiDAR mission with Sentinel-2 optical data. The model infers canopy top height from Sentinel-2 images on land and computes the uncertainty associated to these estimate. This map is representative of the year 2020. More information about the reference map and its development is available in [10]. Figure **??** shows the map projected onto EASE2 grid.

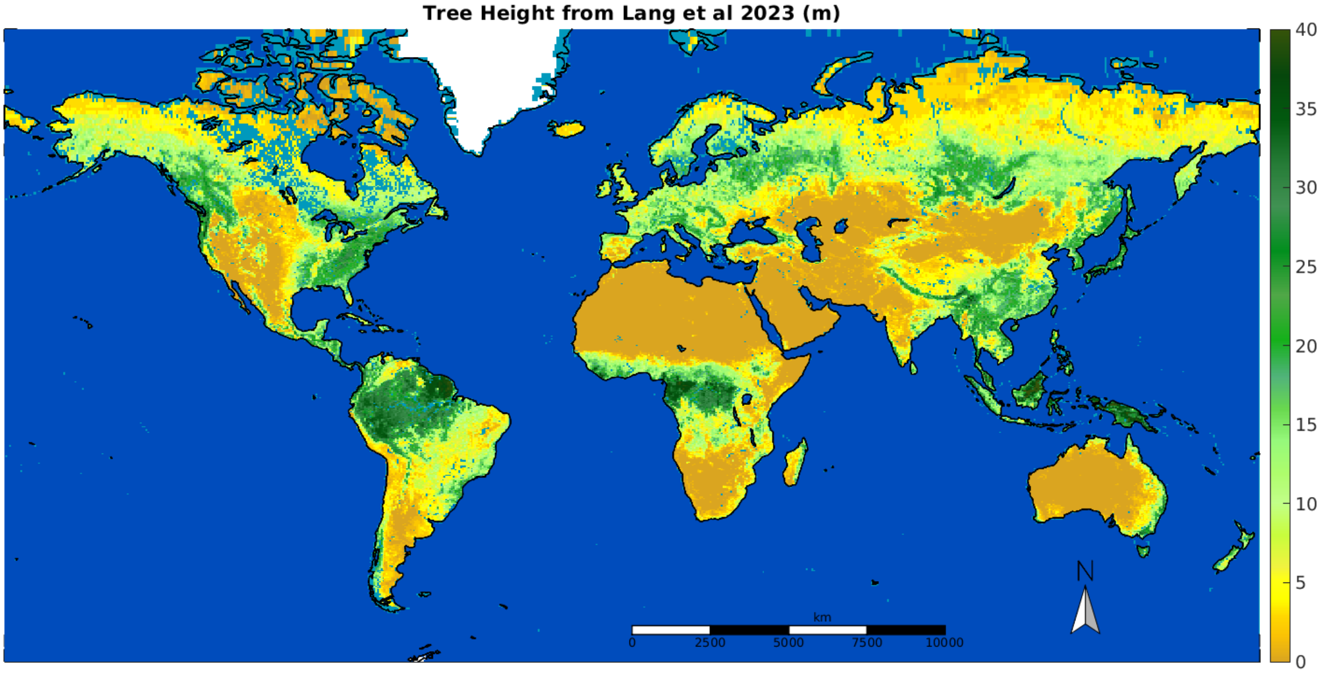

Figure 4: Tree Height from [10] reprojected on the EASE2 grid.

2.2.4 *Future additional reference maps*

In future releases, it is planned to deliver global and yearly AGB or TH estimates produced from additional reference maps. This technical note will be updated accordingly.

3 Methodology

The workflow is divided into three main steps. First, a pre-processing of the input data (white filled and green framed box in figure 5), Second the actual processing of the SMOS LVOD products (white filled and red framed box in figure 5). Third, each chosen reference map associated to a reference year is compared against the annual SMOS LVOD map for the same year (for example SMOS LVOD in 2018 against ESA Biomass CCI AGB map for the year 2018) in order to estimate a logistic relationship between the two quantities (yellow framed scatter plot in figure 5). This relationship is then propagated to other years to get the AGB or TH time series estimation for all years from a reference map (maps at the bottom of figure 5. Figure 5 gives an overview of the processing which is further detailed in the following subsections.

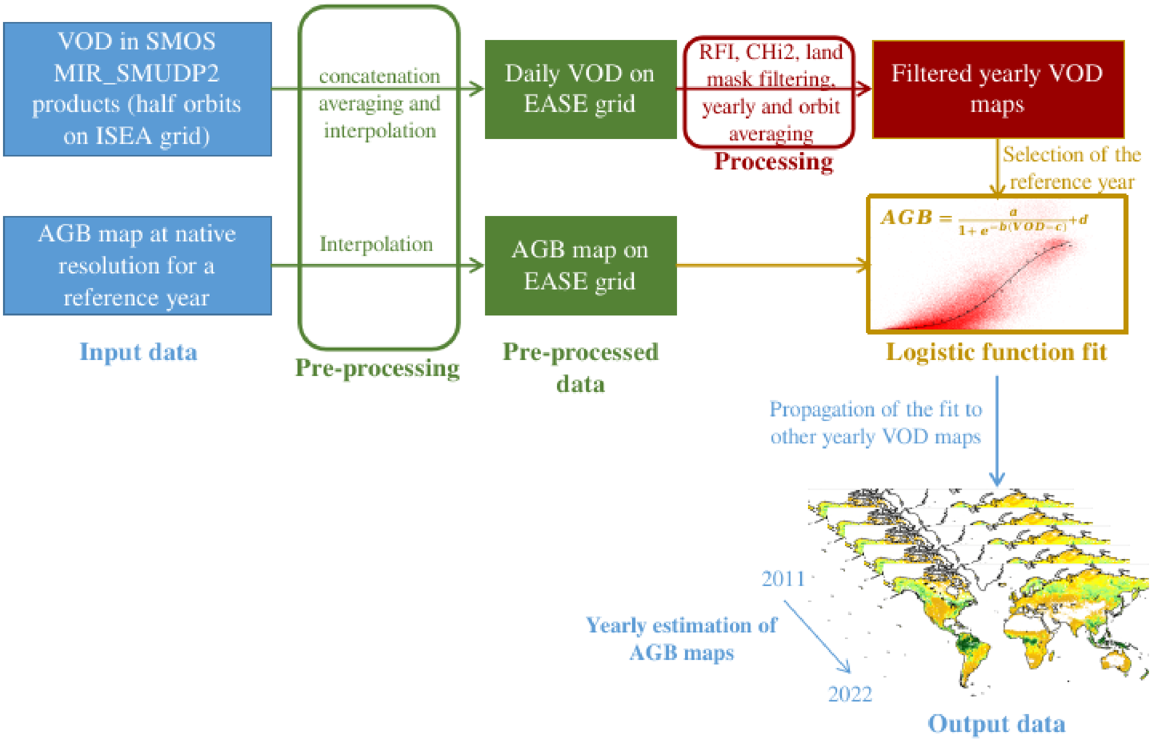

Figure 5: General methodology of yearly AGB map estimation from SMOS LVOD - Methodology adapted from [1].

3.1 Pre-processing

3.1.1 *SMOS LVOD*

As mentioned previously, the SMOS L2 products are distributed as binary files gathering all samples for a half orbit on the ISEA grid. The pre-processing starts with the creation of daily ascending and descending maps of SMOS L2 measurements by concatenating all products sensed the same day. On the DGG nodes where multiple measurements were acquired the same day, the measurement of highest quality and closest to the subtrack is selected. These cases mainly happen at high latitudes. A land mask is also applied on the daily maps. Finally, the variables are interpolated to the EASE grid 2.0 Global, equal area projection (EPSG: 6933), [11, 12]. An example of a preprocessed VOD for one day and for the descending orbits is shown on figure 6.

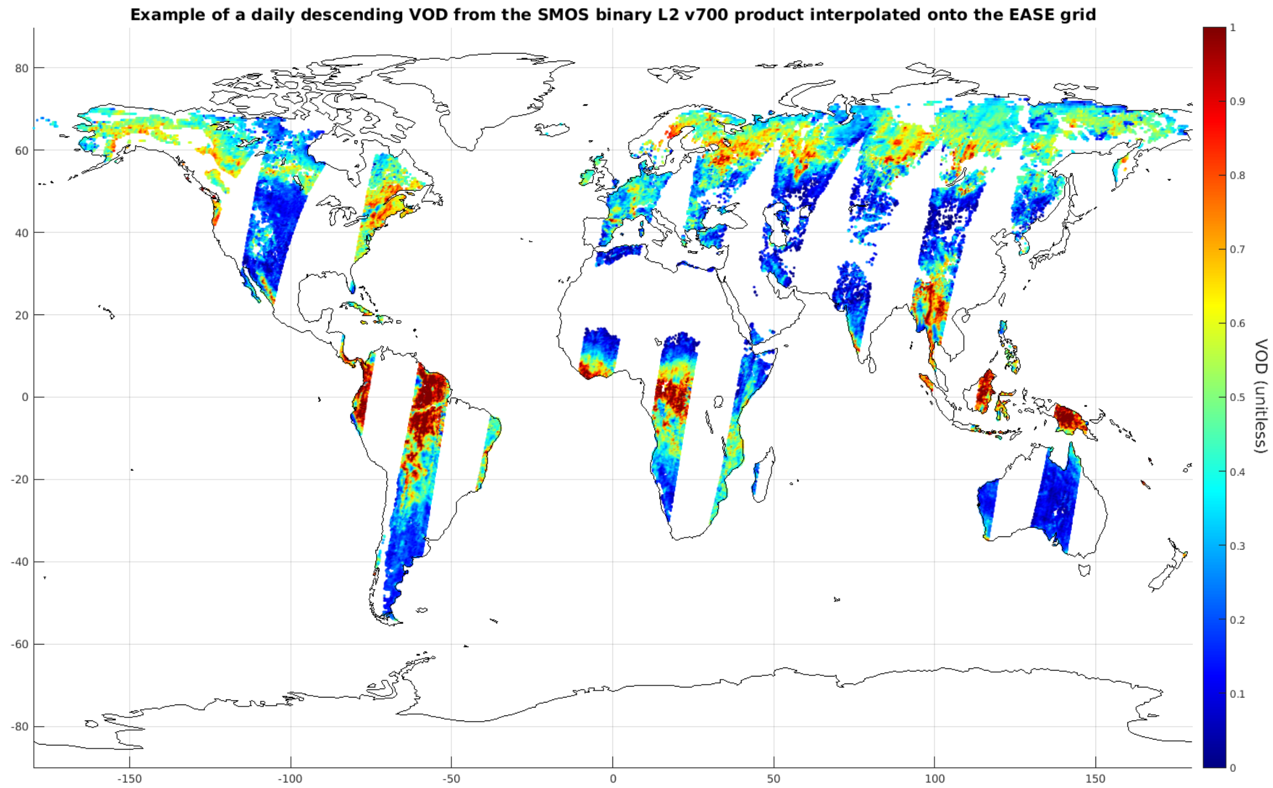

Figure 6: Example of the result of the pre-processing applied on the VOD: from half orbits on the ISEA grid (as shown on Figure 1) to a daily product on the EASE grid.

3.1.2 *ESA Biomass CCI 2018 AGB map*

All tiles of the dataset at 100m were downloaded from [7] and merged together. Then the map was resampled from its native 100m resolution to the same EASE grid as the SMOS L2 products by a simple average.

3.1.3 *Avitabile pantropical AGB map*

As for the ESA Biomass CCI 2018 AGB dataset, the map was resampled from its native 1km resolution to the same EASE grid as the SMOS L2 products by a simple average.

3.1.4 *Lang TH map*

As for the ESA Biomass CCI 2018 AGB dataset and Avitabile pantropical map, the TH map was resampled from its native 10m resolution to the same EASE grid as the SMOS L2 products by a simple average.

3.2 Processing of SMOS LVOD

Once the daily maps for the ascending and descending orbits are created, the data is filtered based on the level of Radio-Frequency Interference (RFI) and the quality of the data retrieval quantified by the probability of Chi2 (Chi2_P field). The SMOS data affected by more than 20% of RFI or where the probability of the Chi2 is below 5% are filtered out. This allows to keep only the best quality retrievals. Outliers that are not within an interval of two sigmas around the average of the VOD temporal series are also discarded. Then, the filtered daily LVOD over land are averaged on a yearly basis, merging ascending and descending orbits altogether. Averaging over a year irons out the effects on L-VOD due to diurnal and seasonal variations of the vegetation water content. An example of the retrieved L-VOD map in 2021 is shown in figure 7.

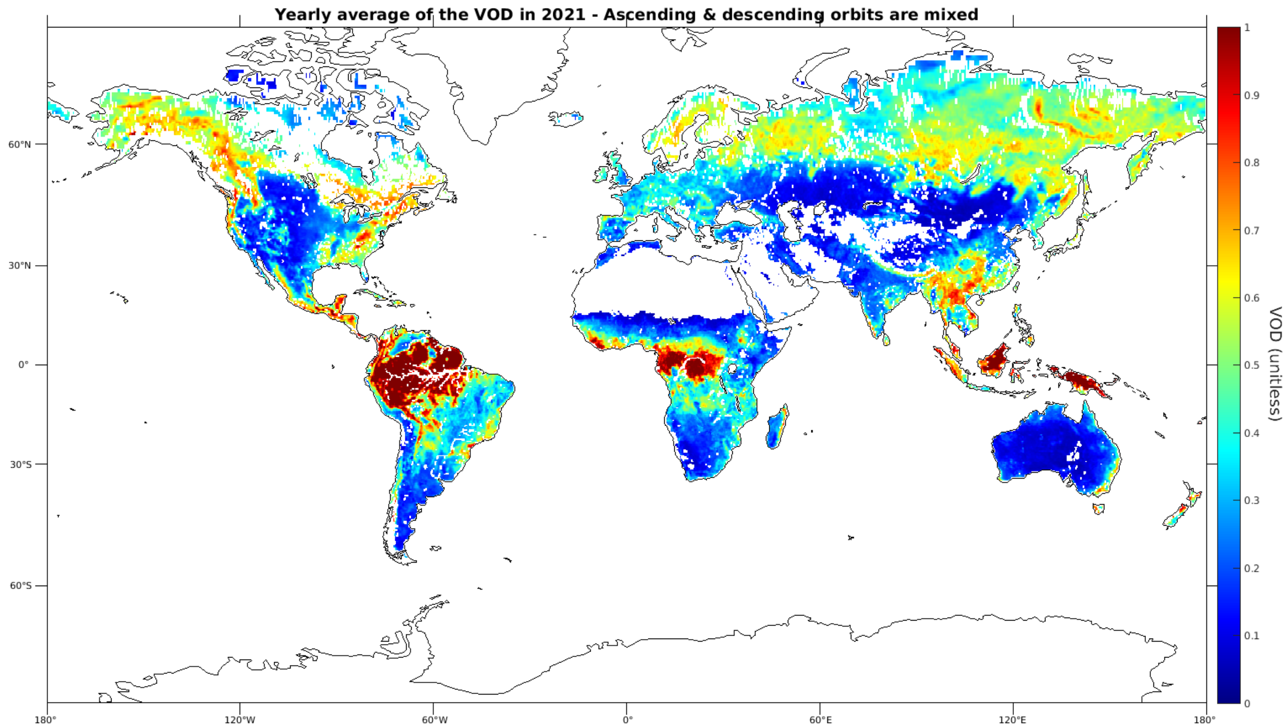

Figure 7: A year of SMOS LVOD averaged over 2021 on the EASE grid - mixed ascending and descending orbits.

3.3 Estimation of the VOD to AGB/TH law

After applying the steps described in the previous parts, there are the global and yearly SMOS LVOD maps on one hand and the reference map for a particular year on the other hand. Both quantities are on the same grid, with the same land mask. A link between them now needs to be established.

First, a scatter plot of the AGB or TH and L-VOD for each land EASE grid pixel helps to check how both quantities are related (red points on figure 8). Second, following the methodology described in [1], the L-VOD is bined into 0.05-width bin. For each L-VOD bin, the mean AGB or TH from the reference map is computed (black points on figure 8). Third, for the AGB products, the set of parameters of the logistic function that best fits the mean AGB/L-VOD distribution is estimated. This logistic function is defined in eq. 1:

$$AGB = \frac{a}{1+e^{-b(VOD-c)}} + d \tag{1}$$

a, b, c and d are the free parameters to obtain the logistic function that best fit the mean AGB distribution in the L-VOD bins (black curve on figure 8). In equation 1, AGB is in $Mg.h^{-1}$ and the $L-VOD$ is dimensionless. Hence a and d are in $Mg.h^{-1}$ and b and c have no dimension.

For the TH product, the method is identical to the difference that rather than fitting a logistic function, a polynomial fit of order 3 between the TH reference map and the L-VOD is performed. The polynomial fit is defined in eq. 2:

$$TH = aVOD^3 + bVOD^2 + cVOD + d \tag{2}$$

Figure 8 resumes the different elements of the method described above for the AGB.

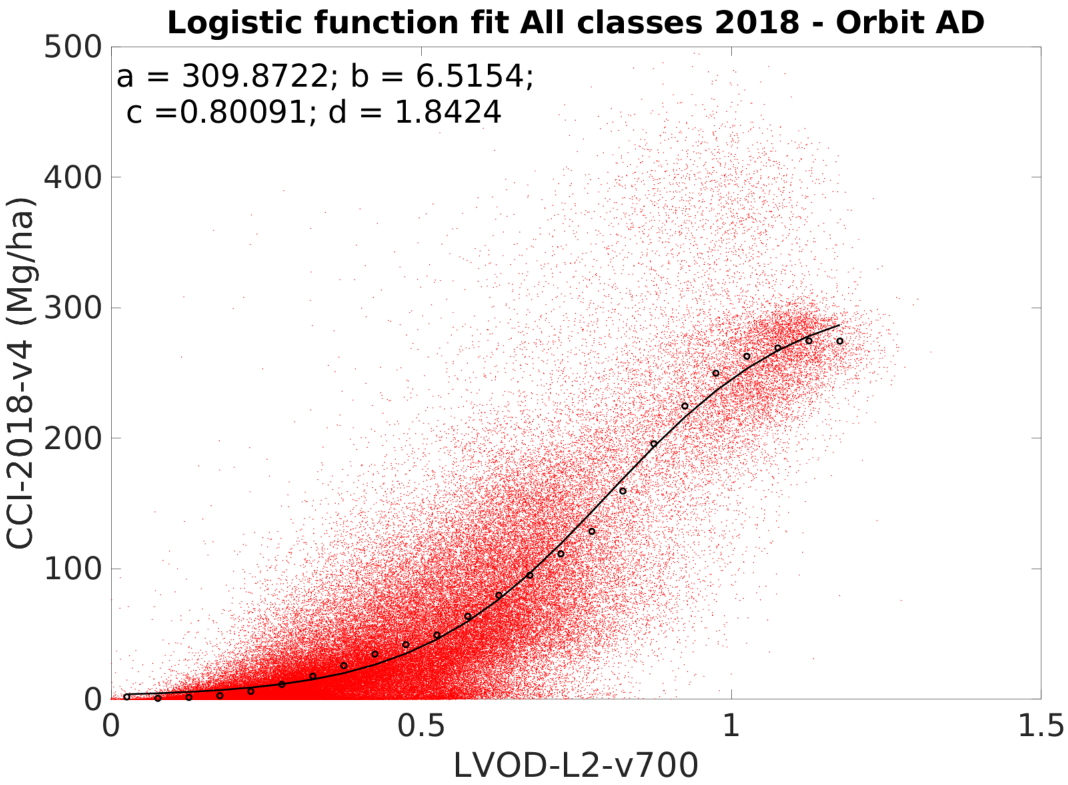

Figure 8: Scatter plot of the averaged SMOS LVOD for 2018 and the AGB from the CCI Biomass project for year 2018. The black dots represent the mean AGB in the 0.05-width LVOD bins and the black curve represents the logistic function that best fits the mean AGB distribution.

Then, the fitted curve with the optimized set of parameters is applied on the LVOD for the reference year. The derived estimated parameter (AGB or TH) is then compared to the reference parameter in order to quantify the differences as on figure 9

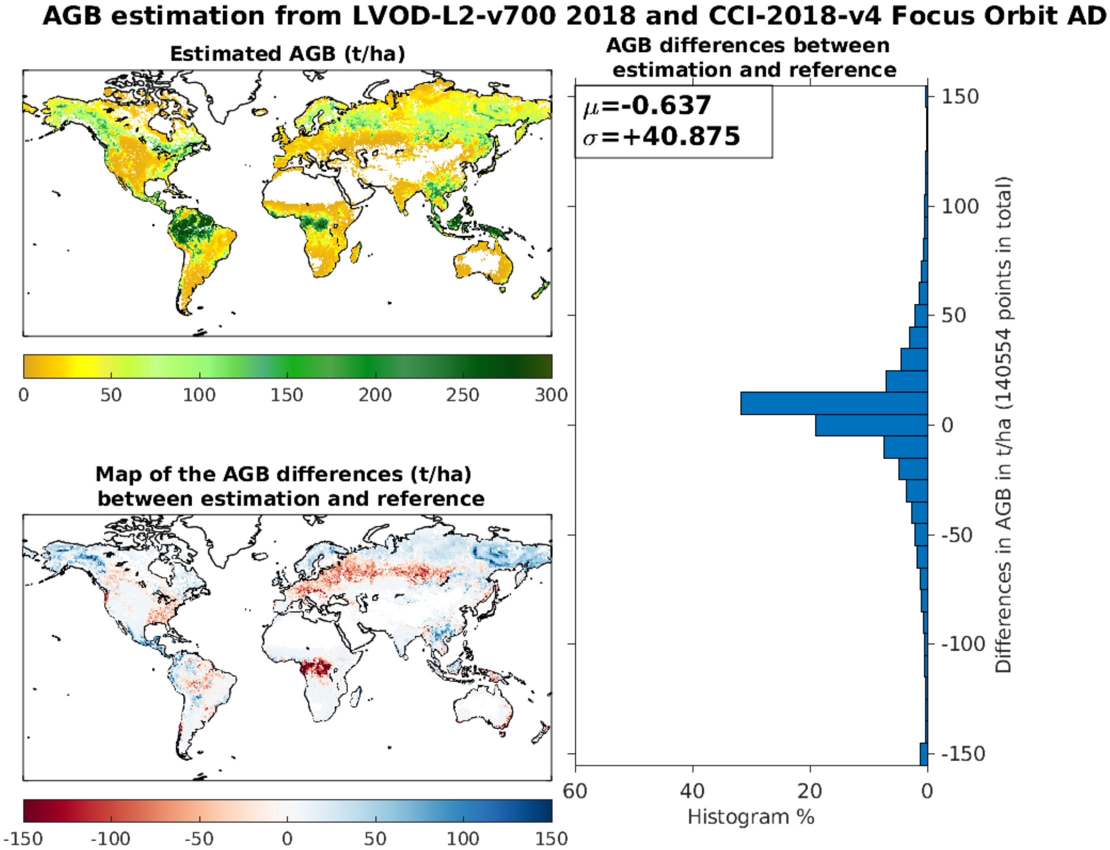

Figure 9: Example of the comparison of the AGB estimated from the LVOD and the optimized logistic function with the reference AGB. Top-left: AGB estimated from the SMOS LVOD and the logistic function; Bottom-left: difference between the estimated AGB and the reference (CCI 2018); bottom-right: histogram of the differences with the mean (μ) and std (σ) of the differences.

Finally, the optimized function is applied on all SMOS LVOD yearly averages. The result is a global and yearly parameter time series on the EASE grid written in a NetCDF file. There is one file per reference map.

3.4 Estimation of the standard deviation (std)

The standard deviation of the estimated parameter for the reference year is derived against the reference values. First, the estimated parameter for the reference year is binned into 10Mg.h^{-1} bins for the AGB and 1m bins for the TH. The mean of the reference parameter values is computed within each bin (blue dots on figure 10). Second, the std of the differences between the reference and estimated parameter is computed (blue vertical bars on figure 10). The result is a discrete std distribution of approximately 30 values. Third, this std distribution is propagated to other years. For each year, the std map is built by affecting to all pixels the reference std value of the bin into which their estimated values fall. Figure 10 gives an example of the reference std estimation. This figure also emphasizes that reference high parameter values may be under estimated when computing AGB or TH from the LVOD and the optimized logistic/polynomial function. Indeed, LVOD tends to saturate over densely vegetated areas even if it does not saturate as much as optical indices and VOD in C and X bands.

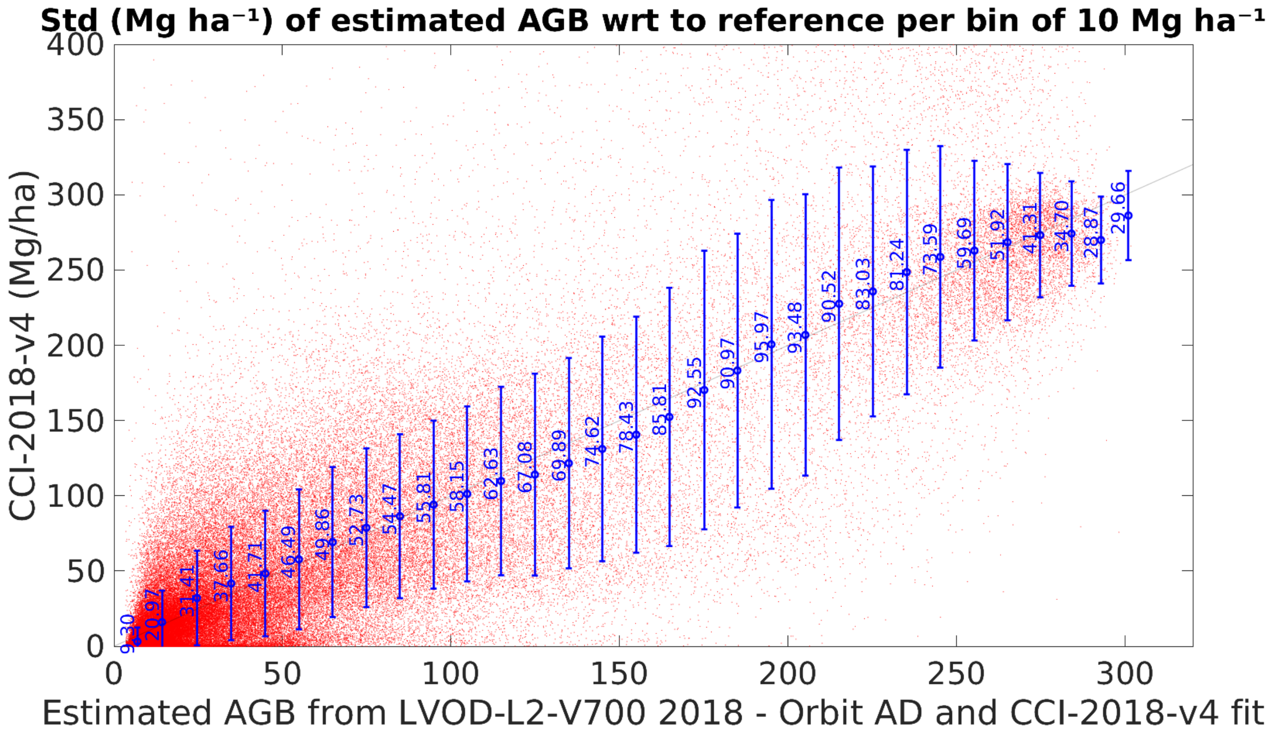

Figure 10: Computation of the std (vertical bars) of the differences between estimated AGB and reference AGB in 10Mg.ha^{-1} bins (blue dots).

4 Description of the NetCDF file

The data are distributed in the NetCDF file format. The NetCDF format is extremely convenient to store and distribute multidimensional data, called variables. A NetCDF file is self describing which means that it contains information around the raw data such as when data elements were captured, the units etc.
A NetCDF file is approximately 60 Mo and contains the different element that are described in the sections below.

4.1 Naming convention

All the files are named as follows:
"SM_SCIE_MIR_L4PPPX_yyyymmddThhmmss_YYYYMMDDTHHMMSS_vvv_ccc_n", where the conventions are very close to those of SMOS Level 2.

Table 1: Naming Description of the L4 AGB Product

SM	in this specific case, it stands for the SMOS mission
SCIE	file class: indicates that this is a scientific product
MIR	file category: MIRAS, as the name of the instrument
L4PPPX	L = for Land 4 = Level 4 product PPP = AGB for Above Ground Biomass ETH for Estimated Tree Height X = counter associated to the AGB reference map (from A to Z, A for CCI-2018, see Table 2)
yyyymmddThhmmss	sensing start time for the data contained in the product With yyyy year mm month dd day of the month hh hour mm minutes ss seconds in this specific case 20110101T000000 as 2011 is the first full operationnal year of SMOS
YYYYMMDDTHHMMSS	sensing stop time for the data contained in the product With YYY year MM month DD day of the month HH hour MM minutes SS seconds in this specific case the last current year of operationnal SMOS measurements (20211231T235959)
vvv	version number of the processor generating the product
ccc	file counter, used to make distinction among products having all other filename identifiers identical: the higher the file counter, the more recent the product
n	processing site (C-PDC=7, C-EC SM=8, C-EC OS=9)

Table 2: Counter description for the reference AGB maps

A	003	corresponds to the CCI-2018 v4 AGB reference map
B	002	corresponds to the Avitabile AGB reference map
C	001	corresponds to the Lang TH map

4.2 Global structure

The distributed NetCDF files follow the common NetCDF conventions. It contains the variables of interest (AGB and associated std) with three dimensions (latitude, longitude and year). The global file structure is displayed on figure 11

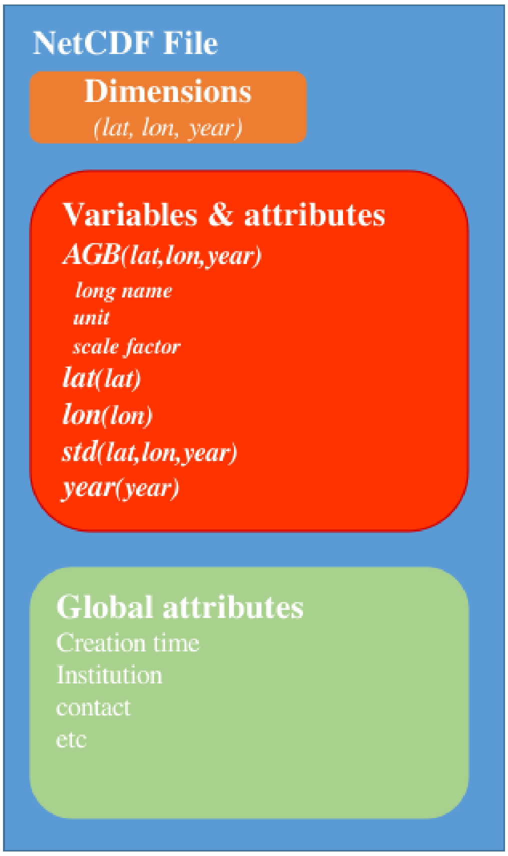

Figure 11: Global structure of the AGB estimates NetCDF files.

4.3 Dimensions

There are three dimensions in the distributed files:

- **latitude**: number of rows in the EASE grid (584 with the chosen resolution of 25km^2)
- **longitude**: number of columns in the EASE grid (1388 with the chosen resolution of 25km^2)
- **year**: number of years for which there is an AGB estimate

4.4 Variables

There are two variables and three coordinate variables in the AGB NetCDF files. A coordinate variable is a unidimensional variable with the same name as a dimension. It is associated to one or several variables dimension and usually defines a physical coordinate corresponding to this dimension.

The two variables are:

- **AGB or TH**: The parameter estimated through the methodology described in this technical note. Its dimensions are the latitude, longitude and year dimensions. Its variable attributes are the *long_name (Above Ground Biomass or Tree Height), units (t/ha or m) and the scale_factor (0.1)*. It is stored as an array of shorts and has a precision of 0.1 t/ha.
- **std**: The standard deviation associated to the estimated AGB. Its dimensions are the latitude, longitude and year dimensions. Its variable attributes are the *long_name (standard deviation of AGB/TH wrt to the reference map), units (t/ha or m) and the scale_factor (0.1)*. It is stored as an array of shorts and has a precision of 0.1 t/ha (or m).

The three coordinates variables are:

- **latitude**: the 584 latitude values (in degrees north) in the EASE grid stored with a float precision;
- **longitude**: the 1388 longitude values (in degrees east) in the EASE grid stored with a float precision;
- **year**: the year values for which there is an parameter estimate.

4.5 Attributes

Several global attributes are included in the distributed NetCDF products such as the creation time, how to cite the product, the reference to this technical note, the contact to gather user feedback, etc. The dataset DOI is *https://doi.org/10.12770/95f76ff0-5d89-430d-80db-95fbdd77f543*. For an exhaustive list of the global attributes, the reader is advised to check them out directly in a product.

5 Data availability

The data can be freely downloaded from the CATDS ftp:

- *https://data.catds.fr/cecsm/Land_products/L4_Above_Ground_Biomass/*

REFERENCES

[1] N. J. Rodríguez-Fernández, A. Mialon, S. Mermoz, A. Bouvet, P. Richaume, A. Al Bitar, A. Al-Yaari, M. Brandt, T. Kaminski, T. Le Toan, Y. H. Kerr, and J.-P. Wigneron, "An evaluation of smos l-band vegetation optical depth (l-vod) data sets: high sensitivity of l-vod to above-ground biomass in africa," Biogeosciences, vol. 15, no. 14, pp. 4627–4645, 2018. [Online]. Available: https://www.biogeosciences.net/15/4627/2018/

[2] A. Mialon, N. J. Rodríguez-Fernández, M. Santoro, S. Saatchi, S. Mermoz, E. Bousquet, and Y. H. Kerr, "Evaluation of the sensitivity of smos l-vod to forest above-ground biomass at global scale," Remote Sensing, vol. 12, no. 9, p. 1450, 2020.

[3] S. Boitard, A. Mialon, N. Rodriguez-Fernandez, C. Salazar, P. Richaume, and Y. Kerr, "Agb etimates from smos vod," –, vol. –, no. –, pp. –, 2023, in prep.

[4] Y. H. Kerr, P. Waldteufel, P. Richaume, J. P. Wigneron, P. Ferrazzoli, A. Mahmoodi, A. Al Bitar, F. Cabot, C. Gruhier, S. E. Juglea, D. Leroux, A. Mialon, and S. Delwart, "The SMOS Soil Moisture Retrieval Algorithm," Geoscience and Remote Sensing, IEEE Transactions on, vol. 50, no. 5, pp. 1384 –1403, 2012.

[5] Y. Kerr, P. Richaume, P. Waldteufel, P. Ferrazzoli, J. Wigneron, M. Schwank, and K. Rautiainen, "Algorithm theoretical basis document (ATBD) for the smos level 2 soil moisture processor," Technical Report TN-ESL-SM-GS-0001-4b SM-ESL (CBSA), p. 145p, 2020. [Online]. Available: https://earth.esa.int/eogateway/documents/20142/37627/SMOS-L2-SM-ATBD.pdf

[6] B. Bengoa, M. Zapata, J. Ortega, and M. Rodriguez, "Smos level 2 and auxiliary data products specifications," Technical Report TSO-TN-IDR-GS-0006 version 8.6, p. 427p, 2020. [Online]. Available: https://earth.esa.int/eogateway/documents/20142/0/SMOS-L2-Aux-Data-Product-Specification.pdf

[7] M. Santoro and O. Cartus, "Esa biomass climate change initiative (biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2017, 2018, 2019 and 2020 v4," NERC EDS Centre for Environmental Data Analysis, 2023. [Online]. Available: https://dx.doi.org/10.5285/af60720c1e404a9e9d2c145d2b2ead4e

[8] M. Santoro, R. Lucas, H. Kay, and S. Quegan, "Cci biomass product user guide v3," CCI_BIOMASS_PUG_V3, p. 47p, 2021. [Online]. Available: https://climate.esa.int/media/documents/D4.3_CCI_PUG_V3.0_20210707.pdf

[9] V. Avitabile, M. Herold, G. Heuvelink, S. L. Lewis, O. L. Phillips, G. P. Asner, J. Armston, P. S. Ashton, L. Banin, N. Bayol, and N. J. Berry, "An integrated pan-tropical biomass map using multiple reference datasets," Global Change Biology, vol. 22, pp. 1406–1420, 2016.

[10] N. Lang, W. Jetz, K. Schindler, and J. D. Wegner, "A high-resolution canopy height model of the earth," Nature Ecology 'I&' Evolution, pp. 1–12, 2023.

[11] M. J. Brodzik, B. Billingsley, T. Haran, B. Raup, and M. H. Savoie, "EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets," ISPRS International Journal of Geo-Information, vol. 1, no. 1, pp. 32–45, 2012, iSSN 2220-9964. [Online]. Available: http://www.mdpi.com/2220-9964/1/1/32

[12] ——, "Correction: Brodzik, M.J., et al. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. ISPRS International J. of Geo-Information 2012, 1, 32 45," ISPRS International Journal of Geo-Information, vol. 3, no. 3, pp. 1154–1156, 2014, iSSN 2220-9964. [Online]. Available: http://www.mdpi.com/2220-9964/3/3/1154