HYDROCOASTAL

SAR/SARin Radar Altimetry for Coastal Zone and Inland Water Level

Product Specification Document
Deliverable D2.3

Sentinel-3 and Cryosat SAR/SARin Radar Altimetry for Coastal Zone and Inland Water

ESA Contract 4000*129872*/20/I-DT

Project reference: HYDROCOASTAL_ESA_PSD_D2.3
Issue: 2.1

29 November 2023

This page has been intentionally left blank

Change Record

Date	Issue	Section	Page	Comment
25/06/20	1.0	all	all	Initial version of the document to distribute and share among the different partners
08/10/220	1.1	all	all	General update according to ESA review and recent work.
20/06/23	2.0	all	all	Minor refinements to reflect the final format of the L2E products. Updated content on L3 products
29/11/23	2.1	all	all	Updates following ESA review

Control Document

Process	Name	Date
Written by:	Albert Garcia-Mondéjar, Ferran Gibert, Ester Vendrel, Stéphanie Urien	20/06/23
Checked by	David Cotton	29/11/23
Approved by:	David Cotton	29/11/23

Subject	Radar Altimetry for Coastal Zone and Inland Water Level	Project	HydroCoastal
Author	**Organisation**	**Internal references**	
Albert Garcia-Mondejar, Ferran Gibert, Ester Vendrell, Stéphanie Urien	isardSAT	HYDROCOASTAL_ESA_PSD	
Joana Fernandes	UPorto		
Michele Scagliola	Aresys		
Ole Andersen, Karina Nielsen, Heidi Ranndal	DTU Space		
Pierre Fabry	Along Track		

Luciana Fenoglio-Marc	U Bonn	
Marcello Passaro	TUM	
Nicolas Bercher	AltHydroLab	
Angelica Tarpanelli	CNR-IRPI	
Elena Zakharova	NUIM	

	Signature	Date
For HydroCoastal team		29/11/23
For ESA	Date: 2023.12.05 '13:29:46 +01'00 Digitally signed by Jérôme Benveniste	

Table of Contents

1 Introduction

1.1 The HYDROCOASTAL Project

The HYDROCOASTAL project is a project funded under the ESA EO Science for Society Programme, and aims to maximise the exploitation of SAR and SARin altimeter measurements in the coastal zone and inland waters, by evaluating and implementing new approaches to process SAR and SARin data from CryoSat-2, and SAR altimeter data from Sentinel-3A and Sentinel-3B.

One of the key objectives is to link together and better understand the interactions processes between river discharge and coastal sea level. Key outputs are global coastal zone and river discharge data sets, and assessments of these products in terms of their scientific impact.

1.2 Scope of this Document

The scope of this document is to identify and specify the format of the data products generated within the HYDROCOASTAL project. It covers both the first test data set and the final product.

1.3 Document Organisation

This document is organised as follows:

- Section 1 (this section) is a short introduction defining the scope of this report.
- Section 2 includes the general definition of products, variables and file formats.
- Section 3 includes L2 products format specifications
- Section 4 includes L3 products format specifications
- Section 5 includes L4 products format specifications
- Last sections include references and the list of acronyms.

1.4 Reference Documents

RD-1 HydroCoastal Proposal: SAR/SARin Radar Altimetry for Coastal Zone and Inland Water Level. Proposal, January 2020.

RD-2 ESA. “Sentinel-3 Product Data Format Specification - SRAL/MWR Level 2 products”, ref. S3IPF.PDS.003.2, issue 2.12, 05 March 2018.

RD-3 Advanced Computer Systems (ACS). CryoSat Ice netCDF L2 Product Format Specification [PFS-I-L2], C2-RS-ACS-ESL-5265 Issue: 1.6, 2019

RD-4 HYDROCOASTAL Deliverable 1.3 ATBD (Algorithm Theoretical Basis Document). V2.0 23/06/2023, isardSAT and HYDROCOASTAL team.

RD-5 HYDROCOASTAL Deliverable 2.1 IODD (Input Output Data Definitions). V1.1 08/10/2020, isardSAT and HYDROCOASTAL team.

2 General definitions

2.1 Product definitions

A block diagram with the main input and output products is shown in Figure 2.1.

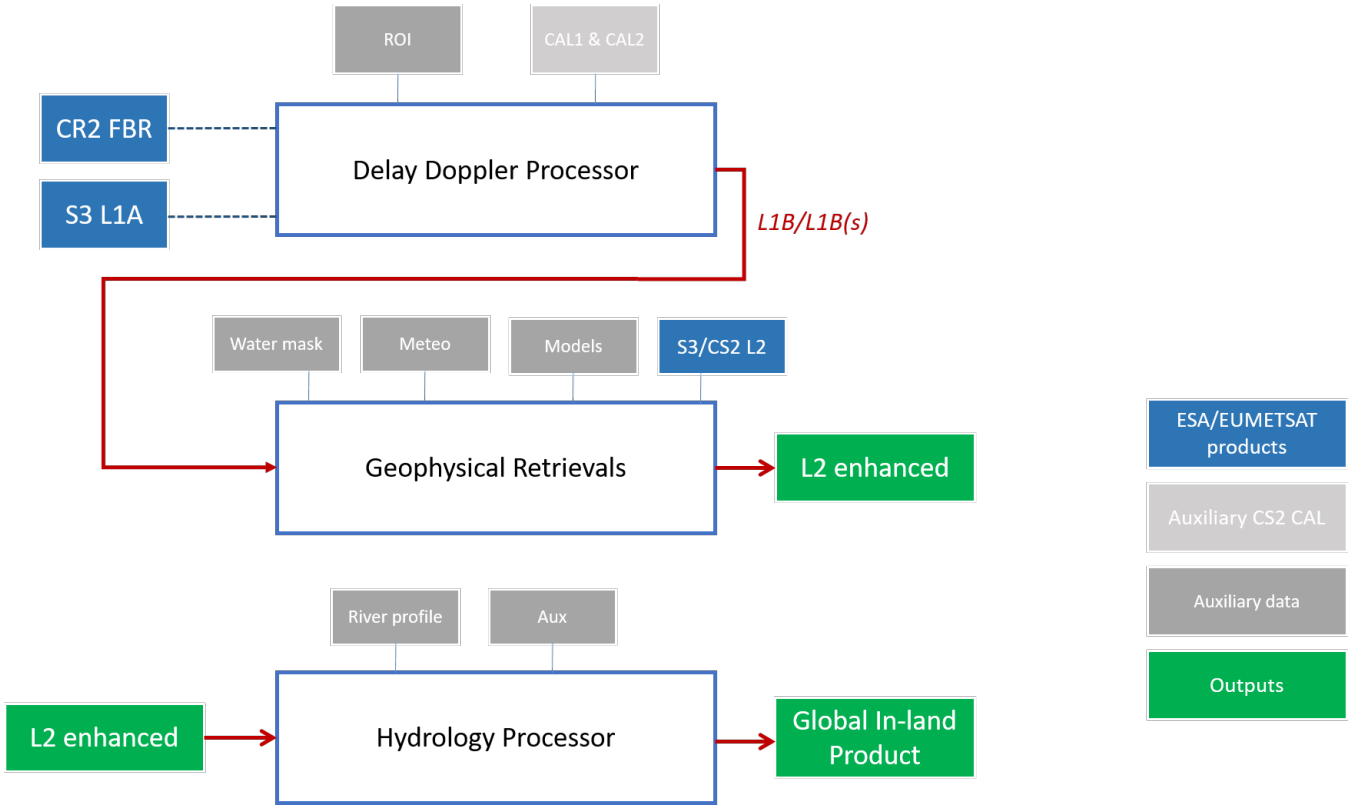

Figure 2.1. Block diagram of the whole process.

2.1.1 Master L2 enhanced product

The master output product is the main product where all the partners will be contributing with the outputs of their algorithms. It will initially be defined as an *empty* product, including essentially the L1B/L1B(s) waveforms and some parameters of the official L2 products. After integrating the contribution from each L2 partner and some geophysical corrections, the final product will be labeled L2 Master enhanced *full* product.
For the test data set, depending on the ROI, the final Master L2 enhanced will include outputs from either Coastal re-trackers or Inland re-trackers, as shown in Figure 2.2.
For the final product, the coastal regions will include L2 output from both the DTU (MWaPP) re-tracker and the UBonn STARS re-tracker, as well as the original L2 output extracted from the standard ESA/EUMETSAT L2 products. The inland water regions will only contain L2 output from the DTU (MWaPP)

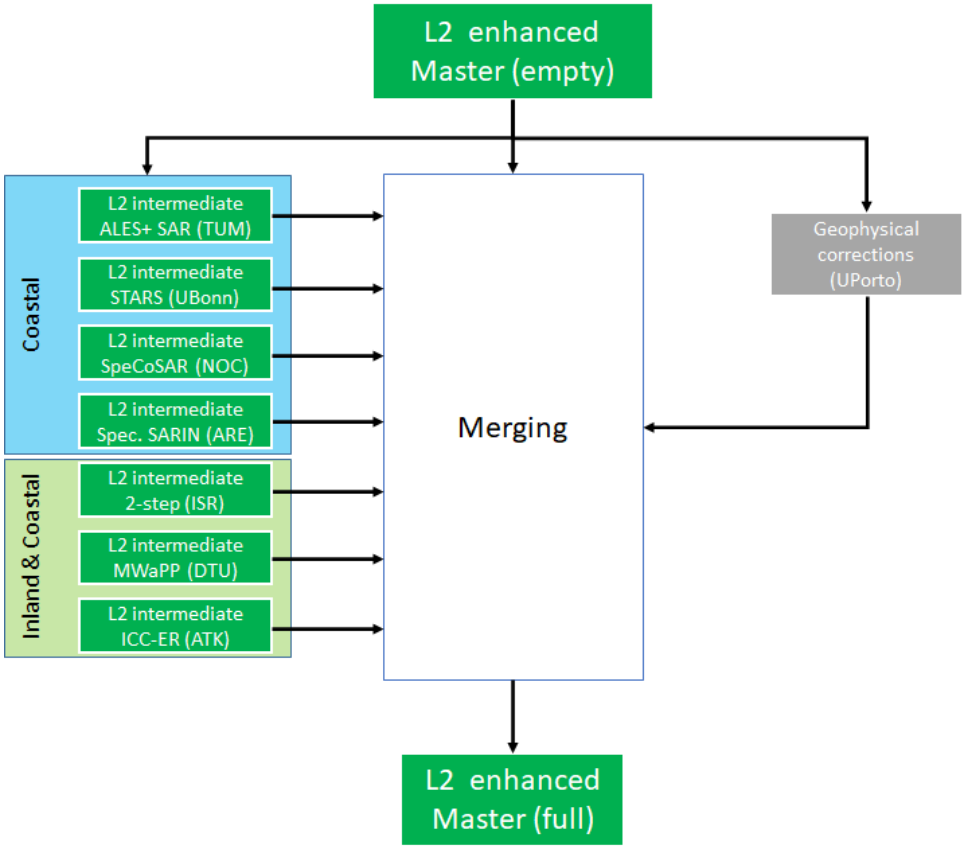

Figure 2.2. Block diagram of the L2 data flow.

2.1.2 Intermediate L2 enhanced product

The products that each of the partners involved in the Retracking needs to generate. The list of parameters will be different depending on the contribution that each of them will need to do. A block diagram with the interaction of the Intermediate L2 enhanced products and the Master L2 enhanced product is found in Figure 2.2.

2.1.3 Auxiliary products

2.1.3.1 Sentinel 3 L2 standard

It will be used as an input to get the geophysical corrections that are not covered by the team and to get the outputs from the ocean re-tracker. The specific geophysical corrections are specified in Table 3.2. Its product format can be checked in the following link: Sentinel 3 L2 Product Format Specification [RD-2].

2.1.3.2 CryoSat-2 L2

It will be used as an input to get the geophysical corrections that are not covered by the HydroCoastal team and to get the outputs from the ocean retracker. The specific geophysical corrections are specified in Table 3.2. Its product format can be checked in the following link: CryoSat Ice netCDF L2 Product Format Specification [RD-3].

2.1.3.3 Other products that may be used (on the global production)

Global River Widths from Landsat (GRWL) Database (Allen et al., 2018). A line layer shapefile with river centerlines and information regarding the width of the river.

2.1.3.3.1 Water Masks

Available global water mask products:

The occurrence product from the Global Surface Water Explorer https://global-surface-water.appspot.com/download (Pekel et al., 2016). This product is based on Landsat imagery and contains the likelihood of water as a percentage with a pixel resolution of 30 m.

HydroLakes https://hydrosheds.org/page/hydrolakes (Lehner et al., 2008). A polygon layer shapefile mask of 1.4 million global distributed lakes.

2.2 Variable types

Variable type	Description	Range
bl	boolean	True, False
str	string of characters	--
uc	8-bit unsigned integer (ubyte)	0 to 255
sc	8-bit signed integer (byte)	-128 to 127
us	16-bit unsigned short integer	0 to 65535
ss	16-bit signed short integer	-32768 to 32767
ul	32-bit unsigned long integer	0 to 4294967295
sl	32-bit signed long integer	-2147483648 to 2147483647
sll	64-bit signed “long long” integer	-9223372036854775808 to 9223372036854775807

fl	32-bit single precision floating point	1.17549e-38 (min) 3.4028e+38(max)
do	64-bit double precision floating point	2.22e-308(min) 1.79e+308(max)

2.3 NetCDF format file

The NetCDF format has been lately widely used to provide remote sensing data, especially in the oceanographic framework. The main advantages of such encapsulating data format are its flexibility in the definition/creation/access of data, its transversal capability to share machine/platform-independent data and their self-describing characteristics. Thanks to such potentialities and ease in data sharing such format has been selected to provide the official distributed L1A, L1BS, L1B, and L2 for the Sentinel 3 and Sentinel 6 missions and it has been incorporated in the CryoSat-2 Baseline D production.

A NetCDF file is composed by the following elements:

- Dimensions: used to represent a real physical dimension (e.g., time, latitude, longitude, or height) or to index other quantities (e.g., number of waveforms or samples). A dimension has both a name and a length.
- Variables: used to store the data in a NetCDF file. A variable corresponds to an array of values of the same type. Each variable is completely defined by its name, data type and shape (described by the list of dimensions). A scalar value is defined as a 0-dimensional array. A variable can also contain related attributes, which can be added, deleted or modified once the variable has been created.
- Attributes: used to keep information of the data (metadata). Generally, they provide information about a specific variable. These are identified by the name of the variable, jointly with the name of the attribute (e.g., units, scale factor, or offset to be added).
- General attributes: used to provide a global description of the dataset as a whole.

2.4 Files alignment

L2 master and intermediate product files from the processor are synchronized/aligned in (1) files number and (2) number of records per files (i.e., variables share dim 0 of same length). Time, location and record number variables are used to link records from these product files.

3 L2 enhanced product Format Specifications

3.1 L2 enhanced NetCDF format

L2 products are compliant with the NetCDF-4 format, following similar variable convention names as the ones provided in Sentinel-3 format product specifications [RD-2]. A NetCDF file contains dimensions, variables, attributes, and global attributes as described in Section 2.3. The global attributes description can be found in Section 3.3.

Table 3.1 Dimensions for Level 2 enhanced product

Dimension name	Description	Value
time	Number of along track records (L2 measurements) in the file. Record frequency is 20Hz.	# of Ku records
Ns	Number of samples in a waveform. It is 128*zp for SAR and 512*zp for SARIn (SARIn always has zp=2).	128*zp (SAR) 1024 (SARIn)
Nl	Maximum number of looks per stack	256 (SAR) 64 (SARIn)

3.2 L2 enhanced product variables

A description of the L2 Master Enhanced product variables is found in Table 3.2. Variables are thematically grouped. For the group called "*Common parameters for each retrackers*", each variable suffix "<RET>" is to be replicated for each retracker developer according to the following notation:

- **ISR** for isardSAT 2-step analytical retracker
- **ARE** for Aresys, for Specialised SARin retracker
- **DTU** for DTU's MWaPP retracker
- **ATK** for Along-Track's ICC-ER empirical retracker
- **UBO** for UBonn's Statistical STARS retracker
- **TUM** for TUM's ALES+ for SAR retracker
- **ESA** for the ocean retracker extracted from official products.

The final L2E products contain the variables from the retracker selected as the best performing algorithms tested earlier in the project, i.e. UBO over coastal areas and DTU over inland waters and the corresponding retracker suffix "<RET>" is :

- **dtu** for DTU's MWaPP retracker
- **ubo** for UBonn's Statistical STARS retracker

- **ESA** for the ocean retracker extracted from official products.

It is important to remark the differences between the different L2 Enhanced Master file versions:

1. The *empty* version of the product contains only the variables tagged as *DDP*, *ESA* and *GeoCor* in Table 3.2. This is the version provided as input to each L2 retracker partner.
2. Each L2 Enhanced Master *intermediate* file is the result of adding the output parameters of a specific L2 retracker to an *empty* L2 Enhanced Master file. The number of L2 *intermediate* files depends on the number of L2 retrackers: each L2 tracker produces its own L2 *intermediate* file. The list of variables inside a specific L2 *intermediate* file is the number of variables in the *empty* version plus the variables added by the specific retracker as identified in the column "Filled by" in Table 3.2. Each L2 *intermediate* file name must include the L2 retracker identifier, as defined in HYDROCOASTAL Deliverable 2.1 IODD [RD-5] Section 4.3.
3. The *full* version of the product integrates all the L2 *intermediate* files generated from a specific L2 *empty* file and incorporates the Geophysical corrections from UPorto.
4. The *final* version of the product is a subset of the full version which integrates only the selected retracker and doesn't include the waveform related variables. In Table 3.2 the last column "Final L2E" indicates whether the variable is included in the final product (Y) or not (N).

For the geophysical corrections module, the value "Uporto" in the "Filled by" column corresponds to the variables generated by the University of Porto while "GeoCor" means that the variables are extracted from the L2 official products. Details of the L2 variables selected and the process applied to them can be found in the HYDROCOASTAL Deliverable 2.1 IODD [RD-5] and in the HYDROCOASTAL Deliverable 1.3 ATBD respectively [RD-4].

Table 3.2 L2 netCDF product variables. The colors are indicative of the variable provider.

Var name	Description	units	Type	Dims	scale_factor	add_offset	Filled by	Final L2E
time	Time at surface of the along track measurement (at the reference point of the tracking window position).	seconds	do	time			DDP	Y
seq_count	Record sequence counter	count	us	time			DDP	Y
lat	Latitude of measurement at nadir [-90, +90]: Positive at North, Negative at South	degrees	sl	time	1.e-06	0	DDP	Y
lon	Longitude of	degrees	sl	time	1.e-06	0	DDP	Y

	measurement at nadir [-180, +180]: Positive at East, Negative at West							
alt	Altitude of the satellite Centre of Mass	meters	sl	time	0.0001	700000	DDP	Y
range	Reference tracker range corrected for USO frequency drift and internal path correction	meters	sl	time	0.0001	700000	DDP	Y
roll	Roll mispointing measured by STRs and post-processed by AOCS or by ground facility.	degrees	sl	time	0.0001	0	DDP	Y
pitch	Pitch mispointing measured by STRs and post-processed by AOCS or by ground facility.	degrees	sl	time	0.0001	0	DDP	Y
yaw	Yaw mispointing measured by STRs and post-processed by AOCS or by ground facility.	degrees	sl	time	0.0001	0	DDP	Y
scale_factor	Scaling factor in order to retrieve sigma-0. It includes antenna gains and geometry satellite - surface. It is not applied to waveforms	dB	do	time			DDP	Y
doppler_freq_sar	Doppler frequencies used to calculate the Doppler beams	degrees	sl	time* Nl			DDP	N
velocity_x	Satellite velocity vector: x component	m/s	do	time			DDP	Y
velocity_y	Satellite velocity vector: y component	m/s	do	time			DDP	Y
velocity_z	Satellite velocity vector: z component	m/s	do	time			DDP	Y

range_migration_corr	Slant range correction for each beam in the stack	m	do	time* Nl			DDP	N
altitude_rate	Altitude rate at 20Hz	m/s	sl	time	0.01	0	DDP	Y
RIP_sar	Integrated stack in fast time dimension	Watt	fl	time* Nl			DDP	N
look_angle_start	Look angle of the first contributing look/beam (Look angle is defined as angle between nadir satellite and the given surface for that beam or look)	radians	ss	time			DDP	N
look_angle_stop	Look angle of the last contributing look/beam (Look angle is defined as angle between nadir satellite and the given surface for that beam or look)	radians	ss	time			DDP	N
burst_nb_start	Burst index of the first contributing look	count	ss	time			DDP	N
burst_nb_stop	Burst index of the last contributing look	count	ss	time			DDP	N
looks_i	I component for the looks of the stack	count	sc	time * Ns*Nl			DDP	N
looks_q	Q component for the looks of the stack	count	sc	time * Ns*Nl			DDP	N
i_scale_factor	The i-samples scaling factor, computed in order to best fit the i-samples within 1 byte. The scaling, needed to convert the looks_i into sqrt(watt), is applied as follows: looks_i_sqr_watt (ku_rec,Nl, Ns) = looks_i (ku_rec,Nl, Ns) * i_scale_factor(ku_rec	sqrt Watt / count	fl	time *Nl			DDP	N

	,Nl)							
q_scale_factor	The q-samples scaling factor, computed in order to best fit the q-samples within 1 byte. The scaling, needed to convert the looks_q into sqrt(watt), is applied as follows: looks_q_sqr_watt (ku_rec,Nl, Ns) = looks_q (ku_rec,Nl, Ns) * q_scale_factor(ku_rec,Nl)	sqrt Watt / count	fl	time *Nl			DDP	N
looks_i_masked	I component for the looks of the stack after masking.	count	sl	time * Ns*Nl			DDP	N
looks_q_masked	Q component for the looks of the stack after masking.	count	sl	time * Ns*Nl			DDP	N
waveform_i2q2	The fully calibrated, multi-looked power waveform. The final scaling is given in the variable "waveform_scale_factor", in order to best fit the waveform into 2 bytes	count	sl	time * Ns			DDP	N
waveform_scale_factor	The waveform scaling factor, computed in order to best fit each waveform within 2 bytes. The scaling, needed to convert the waveform into Watt, is applied as follows: waveform_i2q2_watt(ku_rec, Ns) = waveform_i2q2 (ku_rec, Ns) * waveform_scale_factor (ku_rec)	Watt / count	fl	time			DDP	N

Geophysical Corrections and metadata								
h_surf	Altitude above sea level at which the DTC (upt_dry_tropo) and WTC (gpd_wet_tropo) have been computed	m	fl	time			UPorto	Y
upt_dry_tropo	Dry Tropospheric Correction from UPorto	m	fl	time			UPorto	Y
gpd_wet_tropo	Wet Tropospheric Correction from UPorto	m	fl	time			UPorto	Y
gpd_wet_tropo_flag	Data source flag for the Wet Tropospheric Correction from UPorto		uc	time			UPorto	Y
geoid_EIGEN_6C4	Geoid heights from EIGEN-6C4 model	m	fl	time			UPorto	Y
mod_dry_tropo_cor	Dry tropospheric correction **(Only for CS2)**	m	do	time			GeoCor.	Y
mod_wet_tropo_cor	Wet tropospheric correction **(Only for CS2)**	m	do	time			GeoCor.	Y
mod_dry_tropo_cor_zero_altitude	Model dry tropospheric correction at zero altitude **(Only for S3)**	m	do	time			GeoCor.	Y
mod_dry_tropo_cor_meas_altitude	Model dry tropospheric correction at measurement altitude **(Only for S3)**	m	do	time			GeoCor.	Y
mod_wet_tropo_cor_zero_altitude	Model wet tropospheric correction at zero altitude **(Only for S3)**	m	do	time			GeoCor.	Y
mod_wet_tropo_cor_mea	Model wet tropospheric	m	do	time			GeoCor.	Y

s_altitude	correction at measurement altitude **(Only for S3)**							
rad_wet_tropo_cor	Radiometer wet tropospheric correction **(Only for S3)**	m	do	time			GeoCor.	Y
rad_wet_tropo_cor_sst_gam	Radiometer wet tropospheric correction using SST and Gamma **(Only for S3)**	m	do	time			GeoCor.	Y
tb_238	Channel 1 main beam BT **(Only for S3)**	K	do	time			GeoCor.	Y
tb_365	Channel 2 main beam BT **(Only for S3)**	K	do	time			GeoCor.	Y
sig0_ocean	Corrected « ocean » backscatter coefficient **(Only for S3)**	dB	do	time			GeoCor.	Y
rad_surf_type	Radiometer-derived surface type (S3)		do	time			GeoCor.	Y
rain_flag	Altimeter rain flag (S3)		do	time			GeoCor.	Y
open_sea_ice_flag	Ocean/sea-ice flag **(Only for S3)**		do	time			GeoCor.	Y
rad_along_track_avg_flag	Radiometer along-track averaging flag **(Only for S3)**		do	time			GeoCor.	Y
tb_238_quality_flag	Quality flag for Channel 1 main beam BT **(Only for S3)**		do	time			GeoCor.	Y
tb_365_quality_flag	Quality flag for Channel 2 main beam BT **(Only for S3)**		do	time			GeoCor.	Y
sig0_ocean_qual	« Ocean » backscatter coefficient validity		do	time			GeoCor.	Y

	flag **(Only for S3)**							
inverse_baro	Inverse Barometric Correction	m	do	time			GeoCor.	Y
atm_cor_sig0	Atmospheric Attenuation Correction **(Only for S3)**	dB	do	time			GeoCor.	Y
GIM_iono	GIM Ionospheric Correction	m	do	time			GeoCor.	Y
altimeter_iono	Ionospheric Correction from the altimeter **(Only for S3)**	m	do	time			GeoCor.	Y
ocean_tide_got	Ocean Tide: Includes the corresponding the short-period part of the loading tide and equilibrium long-period ocean tide height (only for S3) **(Only for S3)**	m	do	time			GeoCor.	Y
ocean_tide_fes	Ocean Tide: Includes the corresponding the short-period part of the loading tide and equilibrium long-period ocean tide height (only for S3) **(Only for S3)**	m	do	time			GeoCor.	Y
load_tide_got	Load tide height for geocentric ocean tide (GOT solution) **(Only for S3)**	m	do	time			GeoCor.	Y
load_tide_fes	Load tide height for geocentric ocean tide (FES solution)	m	do	time			GeoCor.	Y
ocean_tide_eq	Equilibrium long-period ocean tide	m	do	time			GeoCor.	Y

	height							
ocean_tide_non_eq	Non-equilibrium long-period ocean tide height	m	do	time			GeoCor.	Y
solid_earth_tide	Solid Earth Tide	m	do	time			GeoCor.	Y
geocentric_polar_tide	Geocentric Polar Tide	m	do	time			GeoCor.	Y
hf_fluct_cor	High frequency fluctuations of the sea surface topography	m	do	time			GeoCor.	Y
sea_state_bias	Sea state bias correction	m	do	time			GeoCor.	Y
geoid	Geoid Height	m	do	time			GeoCor.	Y
MSS	Mean Sea Surface	m	do	time			GeoCor	Y
surf_type	0 open_ocean or semi-enclosed_seas, 1 enclosed_seas or lakes, 2 continental_ice, 3 land		uc	time			GeoCor.	Y
Common parameters for each retracker								
retracked_epoch_<RET>	Estimated epoch in meters w.r.t center of the window (tracker range is given to the center of the window) using the <RET> retracker. (no geophysical corrections applied)	m	do	time			Retrack	Y (ubo, dtu)
retracked_range_<RET>	Corrected range by the retracker offset (using <RET> retracker), the reference range includes already the USO frequency drift and the	m	do	time			Retrack	Y (ubo, dtu)

	internal/instrument corrections. (no geophysical corrections applied)							
retracked_Pu_<RET>	Retrieved power using the <RET> retracker.	dB	do	time			Retrack	Y (ubo, dtu)
retracked_sig0_<RET>	Backscattering coefficient computed from the retracked power once corrected by the sigma0 scaling factor (scale_factor)	dB	do	time			Retrack	Y (ubo, dtu)
flags_<RET>	Flag indicating successful or failed retracking		do	time			Retrack	Y (ubo, dtu)
Specific parameters for the retrackers								
swh_ESA	Significant wave height	m	do	time			Retr. ESA	Y
ssha_ESA	Sea surface height anomaly	m	do	time			Retr. ESA	Y
swh_isr	Significant wave height	m	do	time			Retr. ISR	N
misfit_analytical_isr	Correlation between the real waveform and the fitted one. It is the Pearson correlation coefficient expressed as percentage. Waveforms with values below 90% should be discarded.	percent	do	time			Retr. ISR	N
swh_are	Significant Wave Height	m	sl	time			Retr. ARE	N
misfit_are	Misfit computed according to the following formula: sqrt(sum((L1B_wave-fitted_wave).^2))/sqrt(sum((L1B_wave).^2))	unitless	sl	time			Retr. ARE	N

num_peaks_dtu	Number of peaks in waveform	count	sl	time			Retr. DTU	Y
dif_height_dtu	Difference from average peak height	m	uc	time			Retr. DTU	Y
num_peaks_atk	Number of processed peaks in the waveform (math notation : M_{PWF}) ; this number being 0 or 1 for LRM and SARM.	count	uc	time			Retr. ATK	N
water_fraction_atk	Fraction of water within the Doppler footprint of the altimetric measurement. This is the ratio, within the Doppler footprint, of the water surface vs. the surface of the Doppler footprint itself. It is computed by intersecting the ground projected Doppler footprint and the water mask.	--	ui	time			Retr. ATK	N
ind_pk_start_atk	first peak start index.	count	uc	time			Retr. ATK	N
ind_pk_stop_atk	first peak stop index.	count	uc	time			Retr. ATK	N
elv_rtk_atk	elevation of the retracked point above the reference ellipsoid (for SARIn mode only)	metres	ss	time			Retr. ATK	N
lat_rtk_atk	geodetic latitude of the retracked point	degrees	sl	time			Retr. ATK	N
lon_rtk_atk	longitude of the retracked point	degrees	sl	time			Retr. ATK	N

swh_ubo	Significant Wave Height	m	do	time			Retr. UBO	Y
ssb_ales_20_ku_tum	Sea State Bias correction specific to the ALES+ SAR Range output	m	fl	time			Retr. TUM	N

3.3 L2 global attributes

Table 3.3: L2 enhanced product Global attributes

Attribute name	Description	Format
project_name	Name of the project ()	string
netcdf_version	NetCDF version	string
product_name	Name of the product	string
mission_name	Name of the mission	string
operation_mode	Name of the altimeter mode (SAR)	string
retrackers	*Empty*	string
altimeter_sensor_name	Name of the altimeter sensor	string
gnss_sensor_name	Name of the GNSS sensor	string
doris_sensor_name	Name of the DORIS sensor	string
creation_time	UTC Date of the creation time of the data set (YYYY-MM-DD HH:MM:SS.mmmmmm)	string
first_meas_time	UTC Date of the first measurement of the data set (YYYY-MM-DD HH:MM:SS.mmmmmm)	string
last_meas_time	UTC Date of the first measurement of the data set (YYYY-MM-DD HH:MM:SS.mmmmmm)	string
xref_altimeter_level0	Name of the altimeter level 0 used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_navatt_level0	Name of the navatt level 0 used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_altimeter_orbit	Name of the orbit file(s) used on input to the L1A / FBR processing, copied from L1A / FBR product	string

xref_altimeter_ltm_lrm_cal1	Name of the LTM LRM CAL1 file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_altimeter_ltm_sar_cal1	Name of the LTM SAR CAL1 file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_altimeter_ltm_ku_cal2	Name of the LTM Ku CAL2 file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_altimeter_ltm_c_cal2	Name of the LTM Ku CAL2 file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_altimeter_characterisation	Name of the altimeter characterisation file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_doris_uso	Name of the DORIS USO file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
xref_time_correlation	Name of the time correlation file used on input to the L1A / FBR processing, copied from L1A / FBR product	string
semi_major_ellipsoid_axis	Semi-major axis of the reference ellipsoid (meters)	do
ellipsoid_flattening	Flattening coefficient of the reference ellipsoid	do
absolute_orbit	Absolute orbit number (set to +000 if not used)	ss
phase_number	Phase number. If not used set to +00000	sl
reference_tracking_gate	Index of the gate where the tracking window is referenced. It is 43, starting from 0, in for the products from Sentinel 3, 63 for the ones coming from SAR CryoSat-2 and 255 for SARIn. No zero padded waveforms considered.	us
zero_padding_factor	Zero padding factor used during the delay doppler processing	sl
hamming_flag	Indicates if the hamming has been applied during the delay doppler processing (set to 1 if Hamming was applied)	sl
ddp_processor	Indicates the DDP processor used	string
configuration	DDP processor configuration	string
history	Provides an audit trail for modifications to the original data	string
NCO	NCO version used for NetCDF product upgrade, when any	string

4 L3 Inland water Format Specifications

L3 products are Water Level data derived from L2 measurements confined by the auxiliary data sets such as global lake and river water masks. Each L3 record provides an estimation of river/lake water level for a given overflight over the waterbody as made by the satellite.
L3 products are processed using two different L3 processors: one from AltiHydroLab.fr (AHL) and another one from DTU Space.

For the Sentinel-3 repeat orbit missions, this leads to time series of water level with outlier rejection routines applied.
For the CryoSat-2 non-repeat mission this leads to space-time L2 data of water level, with not any outlier rejection routines applied.
CryoSat-2 and Sentinel-3 data are treated differently, because the Sentinel-3 data are on repeat tracks, and so outliers in a time series can be identified
L3 product files are organised per ROI and/or basin.

4.1 L3 files format

L3 products format is HDF5, natively compatible with standard netCDF4 libraries.

4.2 AHL L2WM Products description (CryoSat-2 only)

The CryoSat-2 L2WM products from AHL are organised as concatenated subsets of L2E products coordinates (time, lon, lat) for the L2E measurement falling into the Water Mask.
There are as many L2WM files as L2E files, except for locations where the L2E files do not intersect the HydroBASINS and/or the SWORD databases.
Inside a L2WM file, each block of data, called an “overflight”, is assigned an “overflight identifier” and metadata from various databases.

The files have the following naming convention:
“level”_shrink_”L2EProductName”_”producedBy”_”version”.nc

e.g. for one of the Rhine Products:
l2wm_shrink_HCA_L2E_CS_LTA__SIR1SAR_FR_20100722T123818_20100722T124001_D001_ahl_v0_19.13.nc

They are held in sub-directories under the river name (e.g. “Rhine”).

The products contain:

- Per L2E measurement:
  - Subset of L2E time, lon, lat coordinate variables.
  - Record number of the L2E meas. in the original L2E file
  - GSW/Occurrence values (one per L2E meas.)

- Per Overflight group:
  - ID

- Population (nb of L2E meas. in the overflight)
- time, lon and lat of the central meas. of the overflight
- Pfafstetter basin identifier (extracted from HydroBASINS database)
- SWORD nodes & reaches metadata (from SWORD v15)

The AHL L2WM Products contents are organised in hierarchical groups as described below:

- group: “<posting_freq>” (here “22hz”)
  - group: “ovf”
    - group: sword
      - group: nodes
      - group: reaches

4.2.1 (L2E native) Posting frequency data group

This group contains the time and lon, lat coordinates of the original L2E measurements falling into the water mask. There are also additional variables such as the overflight identifier (that groups L2E meas. per overflights) and variables extracted from the water mask database.

Var name	Description	units	Type	Dims
time	Time at surface of the along track measurement (at the reference point of the tracking window position).	seconds	do	time
lon	Longitude of measurement at nadir [-180, +180]: Positive at East, Negative at West	degrees	sl	time
lat	Latitude of measurement at nadir [-90, +90]: Positive at North, Negative at South	degrees	sl	time
iswater	Water Masking flag from GSW/Occ datasets. True: measurement coordinates falls into the Water Mask (%s) ; False: measurement is not over Water Mask. [Always True is the case of this product]	n.a.	bool	time
occ_values	GSW/Occ Water Occurrence rate values read from raster data for (lon,	1e2 %	ubyte	time

	lat) coordinates. Water Mask used is GSW/Occ, version 2019v2, type raster. Threshold value applied to Occurrence data is (>=) 10.0.			
ovf_id	Identifier of the water body overflight to which the record belongs. Each time the satellite overflights a waterbody (=record coordinates falling into the water mask), this variable takes a new value. Values: 0: land ; >0: waterbody overflight ID. Can be used to link together records from L<n> altimetry data (n<3) to L3/Alti-Hydrology data. Derived from Water Mask Data Base GSW/Occ. Padding of missing records enabled.	n.a.	do	time

4.2.2 Overflights data group

This group contains the information related to each of the “overflight groups” (groups of L2E meas. over uninterrupted sections of a waterbody). There is one record per overflight.

Var name	Description	units	Type	Dims
time	Time of the central L2E meas. of the overflight	seconds	do	time
lon	Longitude of the central L2E meas. of the overflight	degrees	sl	time
lat	Latitude of the central L2E meas. of the overflight	degrees	sl	time
id	Identifier of the water body overflight to which the record belongs	n.a.	do	time

4.2.3 SWORD/Nodes data group

This group contains the SWORD/Nodes metadata related to each of the “overflight groups” (groups of L2E meas. over uninterrupted sections of a waterbody). There is one record per overflight.

Var name	Description	units	Type	Dims

time	Time of the central L2E meas. of the overflight	seconds	do	time
lon	Longitude of the SWORD Node (known as ‘x’ in the SWORD/Nodes database)	degrees	sl	time
lat	Latitude of the SWORD Node (known as ‘y’ in the SWORD/Nodes database)	degrees	sl	time
distance_score	ECEF distance in between the SWORD Node and the overflight (lon, lat, alt=0) point over WGS84. This score is used to select the closest SWORD Node to each overflight.	m	do	time
other variables	*Cf. the SWORD/Nodes v15 documentation*			

4.2.4 SWORD/Reaches data group

This group contains the SWORD/Reaches metadata related to each of the “overflight groups” (groups of L2E meas. over uninterrupted sections of a waterbody). There is one record per overflight.

Var name	Description	units	Type	Dims
time	Time of the central L2E meas. of the overflight	seconds	do	time
lon	Longitude of the SWORD Reach (known as ‘x’ in the SWORD/Reaches database)	degrees	sl	time
lat	Latitude of the SWORD Reach (known as ‘y’ in the SWORD/Reaches database)	degrees	sl	time
distance_score	ECEF distance in between the SWORD Reach and the overflight (lon, lat, alt=0) point over WGS84. This score is used to select the closest SWORD Reach to each overflight.	m	do	time
other variables	*Cf. the SWORD/Reaches v15 documentation*			

4.3 AHL L3 Products description (Sentinel-3 only)

The Sentinel-3 A&B L3 products from AHL are organised as one file per ROI/basin (some ROI spans two basins and thus have two files). The basin information is extracted from HydroBASINS database making each file hydrologically consistent (i.e, the data in one file never crosses basin's boundaries).
Inside the L3 AHL files are a series of Virtual Stations data. For each virtual station, there are two time series of the river water level: one based on the DTU re-tracker and one based on the original ESA re-tracker.

The files have the following naming convention:
"project"_"producedBy"_"sat"_"river"_basin"XX"_"level"_"version".h5, for the Po river this would be: hydrocoastal_ahl_S3_po_basin21_L3_ v0_19.13.h5"
(XX is the Pfafstetter code for the river basin)

The products contains:

- Water level time series of selected virtual stations* on the rivers included the SWORD database and located within the geographic limits of the ROI polygon.
  (*) Time series not satisfying the two conditions below were discarded:
  - to span at least one year (>=365 days), 2)
  - to have a Sampling Loss Rate < 70% (cf. PVP document for details about the Sampling Loss Rate)
- Time series from two retrackers (DTU, ESA)
- Geoid height
- SWORD v15 metadata

The AHL L3 Products contents are organised in hierarchical groups as described below:

- global attributes
- group: "ts"
  - group: "ts_wl_<retracker1>"
    - group: "<vs_id1>"
    - …
    - group: "<vs_idN>"
  - group: "ts_wl_<retracker2>"
    - …
- group: "vs_info"

4.3.1 Global attributes

Table 4.3: AHL L3 product Global attributes

Attribute name	Description	Format
Product	Name of the product	string

Version	Version of the L3 processor	string
Author	Author of the products	string
Project	Name of the project, ESA contract number	string
Contact	Name and e-mail of contact person	string

4.3.2 Time series data groups

The data derived from each retracker, “ESA or “dtu”, is stored into a dedicated group, respectively, in groups “ts/ts_wl_ESA” and “ts/ts_wl_DTU”.
Inside these groups, the water level data from each Virtual Station is stored into a dedicated group, with the name of the group being the ID of the central SWORD Node of the VS, eg, group “ts/ts_wl_DTU/31362000100041”.

Below is the description of such a VS group:

Var name	Description	units	Type	Dims
time	Time in decimal years	decimal year	do	time
mlon	Mean longitude of obs w.r.t. WGS84 ellipsoid	signed degrees	sl	time
mlat	Mean latitude of obs w.r.t. WGS84 ellipsoid	signed degrees	sl	time
wl	Water level w.r.t. EIGEN-6c4	m	do	time
wlsd	Standard Deviation of the obs that produced wl value	m	do	time
geoid	EIGEN-6c4 geoid correction w.r.t WGS84	m	do	time
nobs	Number of obs used to produce the wl value	n.a.	ui	time
nrange	The number of different SWORD nodes related to the data set, always 1 for AHL products (SWORD Nodes are attached per pass over the river, not to obs)	n.a.	ui	time

4.3.3 Virtual Station information group

The group “vs_info” provides details about the availability of each retracker group for each VS ID along with VS metadata. There is one record in this group per VS.

Var name	Description	units	Type	Dims
vsid	Virtual Station ID = Node ID from the SWORD database	n.a.	ui	vs
tstart	The first time in the time series	decimal year	do	time
tend	The last time in the time series	decimal year	do	time
lon	Longitude of virtual station w.r.t. WGS84 ellipsoid	signed degrees	sl	vs
lat	Latitude of virtual station w.r.t. WGS84 ellipsoid	signed degrees	sl	vs
nobs	Number of obs used to produce the wl value	n.a.	ui	vs
nrange	The number of different SWORD nodes related to the data set, always 1 for AHL products (SWORD Nodes are attached per pass over the river, not to obs)	n.a.	ui	vs
nt	The number of times in the time series	n.a.	ui	vs
reach	Reach ID from the SWORD database	n.a.	ui	vs
ESA	Availability of a time series for the given retracker: 1= available, NA=not available	n.a.	ui	vs
dtu	Availability of a time series for the given retracker: 1= available, NA=not available	n.a.	ui	vs

4.4 DTU L3 Products description

This section describes the format of:

- river water level time series (L3) for Sentinel-3A (S3A) and -3B (S3B)

- extracted river water levels (L2) for S3A and S3B
- extracted and reduced river water levels from CryoSat-2
- a combined S3A, S3B, and Cryosat-2 lake level time series (L3) (so far just for Ireland)

4.4.1 Global Attributes

Table 4.4: DTU L3 product Global attributes

Attribute name	Description	Format
Project	Name of the project	string
Product	Name of the product	string
Author	Author of the products	string
Version	Version of the L3 processor	string
Contact	Name and e-mail of contact person	string

4.4.2 DTU River Water Level Time Series (L3) for S3A and S3B

The format of L3, S3 time series is detailed below.
The files have the following naming convention:
“project”_”producedBy”_”sat“_“river”_”level”_”version”.h5, for the Po river this would be: hydrocoastal_dtu_S3_po_L3_v1_4.h5”

The file format is “.h5” and contain 3 groups

'vs_info' : contains the meta data for each potential virtual station, and have following data sets

Variable	Description
basin	Hydrobasin level 6
lat	Latitude of virtual station
lon	Longitude of virtual station
model	Existence of time series: 0=yes, 1=No, modeled failes, 2=No, too few times available. If 1 or 2 please check raw observations (L2 file)
nobs	Number of observations used to produce the time series

nrange	The number of different SWORD nodes related to the data set
nt	The number of times in the time series
reach	Reach id from the SWORD database v14
tend	The last time in the time series
tstart	The first time in the time series
vsid	Node id from the SWORD database v14

The group 'ts': contains the time series. To access a specific time series, this can be done via the virtual station id'vsid'. 'ts' contain the following data sets:

Variable	Description
date	Date [ymd]
mgeoid	Mean geoid height of obs pr time (EIGEN 6c4 geoid correction w.r.t WGS84)
mlat	Mean latitude of obs pr time
mlon	Mean longitude of obs pr time
nobs	Number of obs pr time
nrange	The number of unique SWORD node values of the along track observations pr time
sat	Satellite name
sattrack	Track number of the satellite
sdobs	Standard deviation of the along track observations pr time
time	Time in decimal years
vsid	VS id (SWORD v14 node id)
wl	Model based water level
wlsd	Standard deviation of model based water level

The group “conv'” contain the model parameters for each time series and have the following data sets

Variable	Description

SigmaObs	Model based observation standard deviation
SigmaRW	Model based standard deviation of the process - random walk
vsid	VS id (SWORD v14 node id)

4.4.3 DTU Extracted River Water Levels (L2) for S3A and S3B

Format of extracted water levels (L2) with added SWORD information

The files have the following naming convention:
“project”_”producedBy”_”sat“_“river”_”level”_”version”.h5, for the Po river this would be: hydrocoastal_dtu_S3_po_L2_v1_4.h5”

The file format is “.h5” and contain 2 groups

The level 2 file contains the groups “raw_obs”, “vs_info”.

The “vs_info” group is the same as the L3 “vs_info”, but does not contain the data set “model”

The group “raw_obs” contains the raw observations extracted over the rivers. To access a specific data set, this can be done via the virtual station id 'vsid'.

Variable	Description
date	the data of the observation
flag	" Water level flag, Recommendation: Use data with flags 0 and 1. 0: MWaPP estimate within expected range 1: Not suitable for MWaPP, Narrow Primary Peak retracker used instead 2: MWaPP estimate more than 2 m from expected height over water body according to other waveforms. Use with care. 3: MWaPP estimate more than 5 m from expected height, do not use."
geoid	EIGEN 6c4 geoid correction w.r.t WGS84
height	" Water level w.r.t. EIGEN 6c4"
hsword	"The elevation related to the closest SWORD node id of the along-track water levels"
lat	"Latitude of observations"
lon	"Longitude of observations"

nodeid	"The closest SWORD node id of the along-track water levels"
OCval	"Global surface water occurrence value"
sat	"Satellite name"
sattrack	"Track number of the satellite"
time	"Time in decimal years [UTC]"
track	"Arbitrary track number"
vsis	"VS id"

4.4.4 DTU Extracted and Reduced River Water Levels from CryoSat-2 (L2)

The CryoSat-2 product contains 2 groups “raw” and “reduced”, The group “raw” contains raw water levels based on the DTU retracker extracted over the river.

Variable	Description
basin	Level 6 Basin id from the HydroBasin database
date	Date of along-track measurements
dtuh	Water level of the along-track measurements based on the DTU retracker
flag	Water level flag, Recommendation: Use data with flags 0 and 1. 0: MWaPP estimate within expected range 1: Not suitable for MWaPP, Narrow Primary Peak retracker used instead 2: MWaPP estimate more than 2 m from expected height over water body according to other waveforms. Use with care. 3: MWaPP estimate more than 5 m from expected height, do not use.
geoid	EIGEN 6c4 geoid correction w.r.t WGS84
lat	Latitude of along-track measurements
lon	Longitude of along-track measurements
name	Name of river
node	The closest SWORD node id of the along-track water levels

reach	Reach id from the SWORD database
sattrack	Track number of the satellite
time	Time in decimal years [UTC]
vsid	VS id (SWORD v14 node id)

The group “reduced” contains a water level summary for each crossing based on the DTU retracker extracted over the river.

Variable	Description
basin	Level 6 Basin id from the HydroBasin database
date	Date of along-track measurements
geoid	EIGEN 6c4 geoid correction w.r.t WGS84, Mean geoid value based on along-track measurements
lat	Latitude of virtual station
lon	Longitude of virtual station
model	Modeled Along-track mean, 0=TRUE, 1=FALSE"
name	Name of river at the VS
nobs	Number of observations used to produce the time series
nrange	The number of different SWORD nodes related to the data set
reach	Reach id from the SWORD database
sattrack	Cryosat-2 absolute orbit
time	Time of the reduced mean
vsid	Node id of the virtual station, from the SWORD database
wl	Mean water level based on along-track measurements
wlsd	Standard deviation of the Mean water level based on along-track measurements

4.4.5 DTU Combined Sentinel-3A, 3B and Cryosat-2 Lake Product

The format of the lake product is detailed below:

The files have the following naming convention: “project”_”producedBy”_”satellite”_“lake”_”area”_”level”_”version”.h5, for the Ireland lakes this would be: hydrocoastal_dtu_lake_Ireland__L3_v1_4.h5”

The file format is “.h5” and contain 4 groups: “conv”, “lake_info”, “raw”, and “ts”

'lake_info' : contains the meta data for each potential virtual station, and has the following data sets

Variable	Description
area	Lake area [km^2]
degrade	Discard data from one satellite if ntime < 10: 0=No, 1=yes, if one satellite only has few observations the bias cannot be estimated well.
lakeid	Lake id
lat	Latitude of lake centroid [dd]
lon	Longitude of lake centroid [dd]
model	Existence of time series: 0=yes, 1=No, modeled failes, 2=No, too few times available. If 1 or 2 please check raw observations (L2 file)
name	Lake name (if availale)
nobs	Number of observations available over the lake
nt	The number of times in the time series
tend	The last time in the time series
tstart	The first time in the time series

The group 'ts': contains the time series. To access a specific time series, this can be done via the lake id 'lakeid'. 'ts' contain the following data sets

Variable	Description
date	Date [ymd]
lakeid	Lake id

mgeoid	Mean geoid height of obs pr time (EIGEN 6c4 geoid correction w.r.t WGS84) [m]
mlat	Mean latitude of obs pr time [dd]
mlon	Mean longitude of obs pr time [dd]
nobs	Number of obs pr time
sat	Satellite name
sattrack	Track number of the satellite
sdobs	Standard deviation of the along track observations pr time
time	Time in decimal years
wl	Model based water level [m]
wlsd	Standard deviation of model-based water level [m]

The group “conv'” contain the model parameters for each time series and have the following data sets

Variable	Description
bias	Model based bias between CS and S3 [m]
SigmaObsCS	Model based observation standard deviation for CryoSat-2 [m]
SigmaObsS3	Model based observation standard deviation for Sentinel-3 [m]
SigmaRW	Model based standard deviation of the process - random walk [m]
lakeid	Lake id

The group “raw” contains the raw observations extracted over the lake. To access a specific data set, this can be done via the lake id 'lakeid'.

Variable	Description
date	the data of the observation
flag	" Water level flag, Recommendation: Use data with flags 0 and 1. 0: MWaPP estimate within expected range 1: Not suitable for MWaPP, Narrow Primary Peak retracker used instead 2: MWaPP estimate more than 2 m from expected height over water body according to other waveforms. Use with care. 3: MWaPP estimate

	more than 5 m from expected height, do not use."
geoid	EIGEN 6c4 geoid correction w.r.t WGS84
lakeid	lakeid
lat	Latitude of observations
lon	Longitude of observations
sat	Satellite name
sattrack	Track number of the satellite
time	Time in decimal years [UTC]
wlsd	Along track water level w.r.t. EIGEN 6c4

4.5 L3 Processing variants

L3 products are processed using two different L3 processors: one from AltiHydroLab.fr (AHL) and another one from DTU Space. Each processor can use different auxiliary datasets but in practice share most of them. This section describes these auxiliary datasets and the way they are used in the respective processors.

L3 Processing by AHL for the Sentinel-3 satellites
The L3 AHL products cover the Sentinel-3 A and Sentinel-3B missions. The products are confined by the actual satellite passes, the Global Surface Water Explorer (Pekel et al., 2016), HydroBASINS database (TODO:CITE) and the SWORD v15 Nodes & Reaches databases (TODO:CITE). More details are provided in the ATBD document of the project [RD-4], Section 6.

L3 Processing by DTU Space
L3 DTU products are confined by e.g. the Global Surface Water Explorer (Pekel et al., 2016), the Sentinel-3A/B ground track locations and a subset of a river mask/river centerline product e.g. The Global River Widths from Landsat (GRWL) Database (Allen et al., 2018). The second global product is a CryoSat-2 + Sentinel-3A/B Lake product also provided in 10° by 10° tiles. This product will be confined by a subset of a lake shapefile mask e.g. the hydroshed hydroLakes product https://hydrosheds.org/page/hydrolakes (Lehner et al., 2008).

5 L4 River discharge Format Specifications

5.1 Introduction

L4 water discharge products include the information from the L3 River product relevant to the discharge retrieving site.

Two L4 data sets have been produced, by NUIM and CNR-IRPI

NUIM provides two types of L4 products:

1) research product for three test sites, Salekhard (Ob R.), Memphis (Mississippi R.) and Pontelagoscuro (Po R.). For these sites the discharge was estimated by three methods: rating curve (RC), Bjerklie method (BJ) and Manning method (Man), using DTU L3 test dataset and river width retrieved by NUIM from Landsat-8 and Sentinel-1 images. This product does not contain the information about uncertainties.

2) Global L4 product for 11 large, medium and small rivers located in different climate conditions was retrieved using the rating curve method, which demonstrated the lowest errors among the methods evaluated on the test sites. The Q retrievals from several VS located within ~200 km river stretch were merged in one time series and sorted by **time**. The Q for an individual VS can be extracted based on **vs_id** variable (see Matlab code for example) The VS ID corresponds to the VS ID of Global L3 product. This product is developed for six large rivers, two medium size rivers (annual flow between 20 and 100 km^3/yr) and three small rivers with annual flow bellow 20 km^3/yr. The set of locations on large rivers includes : Khabarovs (Amur R.), Kolymskoe (Kolyma R.), Arctic Red River (Mackenzie R.), Salekhard (Ob R.), Chapeton (Parana R.), Igarka (Yenissei R.). The set of locations on medium size rivers includes : Mainz (Rhine R.), Nadym (Nadym R.). For small river the L4 product was produced at La Reole (Garonne R.,) Loc 1 Downstream (Murray R.) and at Pella Mission (Orange R.). Four large and medium size rivers belong to Arctic zone, one large river (Amur) is located in Boreal (seasonal ice/no permafrost) zone, three rivers are located in Temperate zone and two rivers drain the arid tropical areas.

CNR-IRPI provides two types of L4 products:

1) research product for 12 test sites: Salekhard (Ob R.), Chester, Thebes, Memphis (Mississippi R.), Piacenza, Cremona, Borgoforte, Sermide, Pontelagoscuro (Po R.), Worms, Kaub, Mainz (Rhein R.).

2) Global L4 product for 12 sites: Tortosa, Zaragoza, Gelsa, Ascò (Ebro R.), Bahadurabad, Baruria, Hardinge Bridge (Gange and Brahmaputra R.), Lokoja, Makurdi, Onitsha (Niger R.), Abu Tong, Malakal (White Nile R.).

For these sites the discharge was estimated by three methods: rating curve starting from the altimetry water level (RC), rating curve from the CM indices (CM), merge approach from altimetry and reflectance combination (Merge). The water level from altimetry is derived by DTU L3 dataset, whereas CM indices from MODIS Aqua and Terra and Sentinel-2 images provided by CNR-IRPI.

It is worth noting that the calculated discharge is referred to the location of the ground station used for calibration and not to the location of the satellite virtual stations. The ground station name and coordinates can be found in the Global L4 product only in the **Station information** global attribute.

5.2 L4 NetCDF format

There is one L4 file for each water body (or ROI) addressed by the project:

Table 5.1: Dimensions for Level 4 product

Dimension name	Description	Value
time	Number of water height values in the file (all time series merged), with a frequency related to the revisit period of the satellite over the water body.	# of points

5.3 L4 NUIM Products variables

Table 5.2: L4 NetCDF product variables

Var name	Description	units	Type	Dims	Filled by
time	Time of crossing	decimal year	do	time	DTU
lat	Mean Latitude of waterbody crossing [-90, +90]: Positive at North, Negative at South.	degrees	do	time	DTU
lon	Mean Longitude of waterbody crossing [-180, +180]: Positive at East, Negative at West.	degrees	do	time	DTU
geoid_height	Geoid Height from model	m	do	time	DTU
vs_id	Virtual station id. This will contain an A/B to indicate the mission. The same as in L3 “Virtual Station” product	NA	uc	time	DTU
water_level	Estimated Water Level w.r.t. Geoid height.	m	do	time	DTU
water_discharge_RC	Water Discharge estimated by Rating Curve approach	m3/s	do	time	NUIM

water_discharge_BJ	Water Discharge estimated by Bjerklie approach. The field contains NaN if the retrievals fail	m3/s	do	time	NUIM
water_discharge_Man	Water Discharge estimated by Manning approach. The field contains NaN if the retrievals fail	m3/s	do	time	NUIM
flag_floodplain_MAN	flag[0,1], ==1 if floodplain flow is included	NA	ui	time	NUIM
water_discharge_Merg	Water Discharge estimated by merging altimetry-based discharge retrievals at different Virtual Stations. The field contains NaN if the retrievals fail	m3/s	do	time	CNR
water_width	Water Width used in Bjerklie and Manning equations, retrieved from dynamic masks or from width- height relations. The field will contain NaN if both retrievals fail.	m	do	time	NUIM, ATK
water_slope	Water slope used in Bjerklie and Manning equations, retrieved from L3 river product.The field will contain NaN if both retrievals fail.	m/m	do	time	DTU
reflectance ratio_CM	Reflectance ratio C/M extracted by multispectral imageries	NA	do	time	CNR

5.4 L4 CNR Products variables

Table 5.3: L4 NetCDF product variables

Var name	Description	units	Type	Dims	Filled by
time	Time of crossing	decimal year	do	time	CNR
lat	Mean Latitude of waterbody crossing [-90, +90]: Positive at North, Negative at South.	degrees	do	time	CNR
lon	Mean Longitude of waterbody crossing [-180, +180]: Positive at East, Negative at West.	degrees	do	time	CNR

geoid_height	Geoid Height from model	m	do	time	CNR
vs_id	Virtual station id. This will contain an A/B to indicate the mission. The same as in L3 “Virtual Station” product	NA	uc	time	CNR
water_level	Estimated Water Level w.r.t. Geoid height.	m	do	time	CNR
water_discharge_RC	Water Discharge estimated by Rating Curve approach	m3/s	do	time	CNR
water_discharge_CM	Water Discharge estimated by CM approach. The field contains NaN if the retrievals fail	m3/s	do	time	CNR
water_discharge_Merg	Water Discharge estimated by merging altimetry and reflectance ratio C/M approach. The field contains NaN if the retrievals fail	m3/s	do	time	CNR
reflectance ratio_CM	Reflectance ratio C/M extracted by multispectral imageries	NA	do	time	CNR

5.5 L4 NUIM Global attributes

Table 5.4: L4 product Global attributes

Attribute name	Description	Format
Conventions	NetCDF convention	string
project_name	Name of the project ()	string
provider	Name of the product provider	string
Altimetric Missions	mission used for retrievals	string
Auxiliary Missions	missions and products used for retrievals	string
RC method	short description, reference to the document	string
Bjerklie method	short description, reference to the document	string

Manning method	short description, reference to the document	string
Q_Merging method	short description, reference to the document	string
Water level method	short description, reference to the document	string
Station information	Reference Ground station name and coordinates	String
Uncertainties	Information about validation period	String
Discharge time series validation scores	Validation scores and uncertainties (Rcorr, RMSE, NRMSE, BIAS, Nash- Nash-Sutcliffe efficiency and number of observations used)	String

5.6 L4 CNR Global Attributes

Table 5.5: CNR L4 product Global attributes

Attribute name	Description	Format
Conventions	NetCDF convention	string
project_name	Name of the project ()	string
provider	Name of the product provider	string
Altimetric Missions	mission used for retrievals	string
Auxiliary Missions	missions and products used for retrievals	string
RC method	short description, reference to the document	string
CM method	short description, reference to the document	string
Q merging method	short description, reference to the document	string
Station information	Reference Ground station name and coordinates	String

5.7 Discharge Merging, Validation and Uncertainties

The **Global L4_NUIM** file contains information about validation of the discharge retrievals. This information is provided in the Global attributes and indicates the validation period and the validation scores calculated for variable **water_discharge_Merg**. The variable **water_discharge_RC** contains sorted by time original discharge retrievals from all VS with

successful L4 processing. The variable **water_discharge_Merg** provides temporally smoothed discharge time series created from **water_discharge_RC**. The smoothing is based on Chebyshev window of specific size. The window size is provided in Uncertainties Global attribute. The window size for each TS was selected taking into account number of VS used and type of water regime (snow-fed, rain-dominated, arid). If the smoothing did not result in amelioration of validation scores, the window size was set to 1 (no smoothing applied).

5.8 Code for reading L4 Data

5.8.1 Matlab code to read L4_NUIM data

```
ncdisp(fileIn)

%====================================================
function [arrayName]=NetCDF_read_arrayout(fileIn);

ncid = netcdf.open(fileIn,'NOWRITE');

[~,nvars]=netcdf.inq(ncid);
scale_factor=ones(nvars,1);
add_offset=zeros(nvars,1);
FillValue=nan(nvars,1);

n=1;
for ind=0:nvars-1
    [varname vartype vardimIDs varatts] = netcdf.inqVar(ncid,ind);
    varName{n,1}=char(varname);
    for iat=0:varatts-1   % looking for scale_factor and add_offset
        attname= netcdf.inqAttName(ncid,ind,iat);
        if strcmp('_FillValue', attname)
           FillValue(ind+1,1) = double(netcdf.getAtt(ncid,ind,'_FillValue'));
        end
        if strcmp('scale_factor', attname)
        scale_factor(ind+1,1)                                                    =
double(netcdf.getAtt(ncid,ind,'scale_factor'));
        end
        if strcmp('add_offset', attname)
           add_offset(ind+1,1) = double(netcdf.getAtt(ncid,ind,'add_offset'));
        end
    end
    kill=double(netcdf.getVar(ncid,ind));,
    ifill=find(kill==FillValue(ind+1,1));,
    kill(ifill)=NaN;,  clear ifill
    kill1=double(kill)*scale_factor(ind+1)+add_offset(ind+1);
    eval(['OUT.',varname,'=kill1;'])
    n=n+1;
    clear kill* inan ifill
end
netcdf.close(ncid)
arrayName=OUT;
%================ end of function ==========================
```

```
disp('....Reading and Verifying the L4 file ...')

datacheck=NetCDF_read_arrayout(fileIn);

figure, % TS of Halti and Qalti
subplot    211,    plot(datacheck.time,    datacheck.water_discharge_RC,'*-'),
legend('Qrc')
subplot      212,      plot(datacheck.time,      datacheck.water_level,'*-'),
legend('Hdtu'),title('Halti from all VS')

% --- Discharge for individual VS ------------
figure % TS by VS
vs_ids = unique(datacheck.vs_id);
for ii=1:length(vs_ids)
    inVS=find(datacheck.vs_id== vs_ids(ii));
    cooplt=[datacheck.lon_VS(inVS(1)), datacheck.lat_VS(inVS(1)) ];
 subplot(length(vs_ids),1,ii),  plot(datacheck.time(inVS), Q(inVS,1),'*-')
    title(['VS ',num2str(vs_ids(ii)),'; coo ', num2str(cooplt) ]),
    xlim([2016 2023])
    clear inVS cooplt
end
```

5.8.2 Python code to read L4_NUIM data

```
from netCDF4 import Dataset, num2date
import  matplotlib.pyplot as plt
import  os


fileIn ='F:\yourpass\L4_Hydrocoastal\Hydrocoastal_Amur_Khabarovsk_L4_v1_1.nc'

file_obj=Dataset(fileIn)
print(file_obj)
print (file_obj.ncattrs())
print(file_obj.variables.keys())   #---- gives the parameters' names

var=[]
for item in file_obj.variables.keys():
      var.append(item);
      print('Variable:   \t', item)
      print('Dimensions: \t', file_obj[item].dimensions)
      print('Shape:      \t', file_obj[item].shape, '\n')

for item in var:
      exec(str(item)+ '= file_obj.variables[item][:]')

plt.figure(1)
plt.plot (time, water_discharge_RC,'.')
plt.show()

plt.figure(2)plt.plot (time, water_discharge_Merg,'r')
plt.show()
```

6 References

Allen, G. H., & Pavelsky, T. M. (2018). Global extent of rivers and streams. Science, 361(6402), 585–588. https://doi.org/10.1126/science.aat0636

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. Eos, 89(10), 93–94. https://doi.org/10.1029/2008EO100001

Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540(7633), 418–422. https://doi.org/10.1038/nature20584

7 List of Acronyms

ACE2 Altimeter Corrected Elevations (vers. 2)
AD Applicable Documents
AGC Automatic Gain Control
AH Alti-Hydro
AHP Alti-Hydro Product(s)
AI Action Item
AIM Action Item Management (tool)
AltiKa Altimeter in Ka band and bi-frequency radiometer instrument
AMSR-E Advanced Microwave Scanning Radiometer-Earth Observing System
ANA Agência Nacional de Águas (National Water Agency, Brazil)
AoA Angle of arrival
API Application Programming Interface
AR Acceptance Review
ASAP As Soon As Possible
ASCII American Standard Code for Information Interchange
ATBD Algorithm Technical Basis Document
ATK Along-Track S.A.S.
AVISO Archivage, Validation et Interprétation des données des Satellites Océanographiques
BIPR Background Intellectual Property Right
CASH Contribution de l'Altimetrie Spatiale à l'Hydrologie (Contribution of Space Altimetry to Hydrology)
CCN Contract Change Notice
CFI Customer Furnished Item
CLASS NOAA/Comprehensive Large Array-Data Stewardship System
CoG Centre of Gravity
CPP CryoSat-2 Processing Prototype (CNES)
CryoSat-2 Altimetry satellite for the measurement of the polar ice caps and the ice thickness
CRISTAL Copernicus polaR Ice and Snow Topography ALtimeter
CRUCIAL CRyosat-2 sUCcess over Inland wAter and Land
CSV Comma Separated Values
CTOH Centre de Topographie des Océans et de l'Hydrosphère (Centre of Topography of the Oceans and the Hydrosphere)
DAO Data Access Object
DARD Data Access Requirement Document
DDM Delay-Doppler Map
DDP Delay-Doppler Processor
DEM Digital Elevation Model
DGC Doppler Ground Cell
DPM Detailed Processing Model
DPP Data Procurement Plan
DTC Dry Tropospheric Correction
DTU Danmarks Tekniske Universitet (Technical University of Denmark)
DVT Data Validation Table
ECMWF European Centre for Medium-Range Weather Forecasts
ECSS European Cooperation for Space Standardisation
EGM Earth Gravitational Model
ENVISAT ENVIronment SATellite
EO Earth Observation
EOEP Earth Observation Enveloppe Programme
EOLi Earth Observation Link
EOLi-SA EOLi-Stand Alone
EPN EUREF Permanent Network
ERA ECMWF ReAnalysis
ESA European Space Agency
EUREF IAG Reference Frame Sub-Commission for Europe
FBR Full Bit Rate
FFT Fast Fourier Transform
FR Final Review
FTP File Transfer Protocol
FCUP (from portuguese) *"Faculdade de Ciências da Universidade"*, Science faculty of the University of Porto
GDAL Geospatial Data Abstraction Library
GDR, [I-,S-] Geophysical Data Record, [Interim-, Scientific-]
GFZ Deutsche GeoForschungsZentrum (German Research Centre for Geosciences)
GNSS Global Navigation Satellite System

GOCE Gravity field and steady-state Ocean Circulation Explorer
GPD GNSS-derived Path Delay
G-POD Grid Processing on Demand
GPT2 Global Pressure and Temperature model (vers. 2)
GPP Ground Processing Processor
GPS Global Positioning System
GRACE Gravity Recovery And Climate Experiment
GRDC Global Runoff Data Centre
GRGS Groupe de Recherche de Géodésie Spatiale (Space Geodesy Research Group)
GRLM Global Reservoir and Lake Monitor
GTN-L Global Terrestrial Network - Lakes
HDF-EOS Hierarchical Data Format - Earth Observing System
HGT A SRTM file format
HWS High Water Stage
HYCOS Hycos Hydraulics & Control Systems
HYPE Hydrological Predictions for the Environment model
IAG International Association of Geodesy
IDAN Intensity-Driven Adaptive-Neighbourhood
IE Individual Echoes
IGS International GNSS (Global Navigation Satellite Systems) Service
IM Internal Meeting (e.g. not with the client)
IODD Input Output Data Document
IPF Integrated Processing Facility
ISD isardSAT
ITRF International Terrestrial Reference Frame
IRF Impulse Response Function
Jason-1 Altimetry satellite, T/P follow-on
Jason-2 Altimetry satellite, also known as the « Ocean Surface Topography Mission » (OSTM), Jason-1 follow-on
Jason-3 Altimetry satellite, Jason-2 follow-on
Jason-CS Jason Continuity of Service
KML Keyhole Markup Language
KO Kick Off
L1A Level-1A
L1B Level-1B
L1B-S, L1BS Level-1B-S (aka, Stack data)
L2 Level-2
L3 Level-3
L4 Level-4
LAGEOS Laser Geodynamics Satellite
LEGOS (french acr.) Laboratoire d'Études en Géophysique et Océanographie Spatiale (Laboratory for Studies in Geophysics and Spatial Oceanography)
LOTUS Preparing Land and Ocean Take Up from Sentinel-3
LPS Living Planet Symposium
LRM Low Resolution Mode
LSE Least Square Estimator
LWL Lake Water Level
LWS Low Water Stage
MARS Meteorological Archival and Retrieval System
MDL Minimum Description Length
MMSE Minimum Mean Square Error
MNDWI Modification of Normalised Difference Water Index
MoM Minutes of Meeting
MPC Mission Performance Centre
MRC Mekong River Commission
MTR Mid Term Review
MSS Mean Square Slope
MSS Mean Sea Surface
MWR Microwave Radiometer
NAVATT Navigation and Attitude
NDVI Normalised Difference Vegetation Index
NDWI Normalised Difference Water Index
netCDF Network Common Data Form
NOAA National Oceanic and Atmospheric Administration
NR New Requirement (w.r.t. the SoW)
NRT Near Real-Time
NWM Numerical Weather Model
OCOG Offset Centre of Gravity
OPC One per Crossing
OSTM Ocean Surface Topography Mission (also known as Jason-2), is also the name of the satellites series T/P, Jason-1, Jason-2 and Jason-3
OVS Orbit State Vector
PDF Probability Density Function
PEACHI Prototype for Expertise on AltiKa for Coastal, Hydrology and Ice
PEPS Sentinel Product Exploitation Platform (CNES)
PISTACH (french acr.) Prototype Innovant de Système de Traitement pour les Applications Cotières et l'Hydrologie

PMP Project Management Plan
POCCD Processing Options Configuration Control Document
PR Progress Report
PRF Pulse Repetition Frequency
PSD Product Specification Document
PTR Point Target Response
PVP Product Validation Plan
PVR Product Validation Report
PVS Pseudo Virtual Station(s)
PWF Pseudo Waveform
RADS Radar Altimeter Database System
RB Requirements Baseline (document)
RCMC Range Cell Migration Curve
RCS Radar Cross Section
RD Reference Document
RDSAR Reduced SAR (also known as Pseudo-LRM)
RF Random Forest
RGB Red, Green, Blue
RID Review Item Discrepancy
RIP Range Integrated Power (of the MLD) sometimes referred as Angular Power Response (APR)
RMS Root Mean Square
ROI (geographical) Region(s) Of Interest
RP Report Period (a month that is being reported into a Progress Report)
RSS Remote Sensing Systems
RWD River Water Discharge
RWL River Water Level
SAMOSA SAR Altimetry MOde Studies and Applications
SARAL In Indian "simple", in english "SAtellite for ARgos and AltiKa.
SARIn SAR Interferometric (CryoSat-2/SIRAL mode)
SAR Synthetic Aperture Radar
SARvatore SAR Versatile Altimetric Toolkit for Ocean Research & Exploitation
SCOOP SAR Altimetry Coastal & Open Ocean Performance
SDP Software Development Plan
SEOM Scientific Exploitation of Operational Missions
SHAPE Sentinel-3 Hydrologic Altimetry PrototypE
SI-MWR Scanning Imaging MWR
SME Small and Medium-sized Enterprise
SMHI Swedish Meteorological and Hydrological Institute
SNAP SeNtinel Application Platform
SOA State Of the Art
SOW Statement Of Work
SPR Software Problem Reporting
SPS Sentinel-3 Surface Topography Mission System Performance Simulator
SRAL SAR Radar Altimeter
SRTM Shuttle Radar Topography Mission
SSB Sea State Bias
SSMI/IS Special Sensor Microwave Imager (SSM/I) Sounder
SSO Single Sign-On
Stack Matrix of stacked Doppler beams
STD Standard Deviation
STM Sentinel-3 Surface Topography Mission
STR Star Tracker
SUM Software User Manual
SWBD SRTM Water Body Data
SWH Significant Wave Height
TAI Temps Atomique International (International Atomic Time)
TBC To Be Confirmed
TBD To Be Done
TCWV Total Column Water Vapour
TDS Test Data Set
TMI Tropical Rainfall Measuring Mission (TRMM) Microwave Imager
TN Technical Note
T/P Topex/Poseidon (altimetry satellite)
TR Technical Risk
UNESCO United Nations Educational, Scientific and Cultural Organization
URL Uniform Resource Locator
USGS United States Geological Survey
USO Ultra Stable Oscillator
UTC Coordinated Universal Time
UWM Updated Water Mask
VS Virtual Station(s)
VH Vertical-Horizontal polarisation
VV Vertical-Vertical polarisation
WBS Work Breakdown Structure
WF Waveform
WFR Water Fraction Ratio
WMO World Meteorological Organization

WP	Work Package(s)
w.r.t.	with respect to
WTC	Wet Tropospheric Correction
XML	eXtensible Markup Language
ZP	Zero Padding