D3.2: Data Product User Manual (DPUM)

Project number
313238

Project title
LOTUS— Preparing Land and Ocean Take Up from Sentinel-3

Call (part) identifier
FP7-SPACE-2012-1

Funding scheme
Collaborative project

Deliverable Number D3.1
Title: “Data Product User Manual”
Nature: Report
Dissemination level: Public

Status: v1.0
Date: 15th May 2015

DOCUMENT CHANGE LOG				
Rev.	Date	Sections modified	Comments	Changed by
1	15th May 2015	All	First running version	T.M., P.T. (CLS)

1 Introduction

1.1 Scope and purpose

This document is the Data Product User Manual document (DPUM) providing information and assistance to users so that they can make effective use of the innovative data products which are generated in the frame of the LOTUS project. This User's Manual provides a comprehensive description of the content and format of the data products. In addition, it provides users an overview of the various processing algorithms developed in this project and how they are relate to the different data products.

The algorithms and data descriptions presented in this User's Manual cover the Levels 2, 3 and 4.

1.2 Document structure

This document is structured into an introductory chapter (this section) followed by six chapters:

- Section 2 provides a general description of the LOTUS products and the data processing schemes that are used to produce these demonstration products,
- Section 3 provides the complete list of models and standards used to correct geophysical parameters,
- Section 4 presents the data storage format used for LOTUS data products,
- Section 5 provides a detailed description of the content and format of the LOTUS products,
- Section 6 deals with data access and product naming to help users for selecting data products of interest,
- Section 7 gives general information about any further datasets release.

2 LOTUS User Products

LOTUS is a Copernicus project funded by the European Commission aiming at developing innovative products and applications of Sentinel-3 (based on the use of SAR-mode data) to complete the space observation infrastructures that are designed for land and ocean monitoring for Copernicus. This in turn will support operational ocean and land services, but also provide useful value-added products for commercial activities.

The purpose of this document is to assist users of the LOTUS demonstration data products by providing a comprehensive description of the product format and content.

2.1 Overview of the LOTUS Products

Processing chains have been set up based on new methodologies and new data processing described, to produce new data products for end users. Several kinds of demonstration products are defined and designed in the LOTUS project according to the different applications it is aimed to (for operational marine or land services) and the different users types which may be encountered (scientists or end-users may have different product levels of interest).

All these products address the different surface targets of the Sentinel-3 topography measurements:

- Sea surface heights, wave heights and wind speeds in the open oceans, coastal seas as well as in sea ice covered regions,
- In-land water in rivers and lakes,
- Soil moisture, and
- Snow water equivalent,

and meet the needs of the different users:

- New Level-2 data products, complementing the ESA Sentinel-3 L2 data products, and
- Higher level data products (Level-3 and Level-4), so that end users get an easy access to the specific higher-level information they need.

In the frame of the LOTUS project, data are processed to Levels 1, 2, 3, and 4 over the different targeted test areas, but only Levels 2, 3, and 4 data products are distributed where applicable and if relevant for end-users. Level-1 products are used internally to the project as inputs to Level-2 processing and are not available to the user community.

The following sketch shows the different types of user products that are generated for each application (ocean or land topography measurements, and water surface target) and processing level as described within the DPDD [LOTUS D3.1 report, 2014].

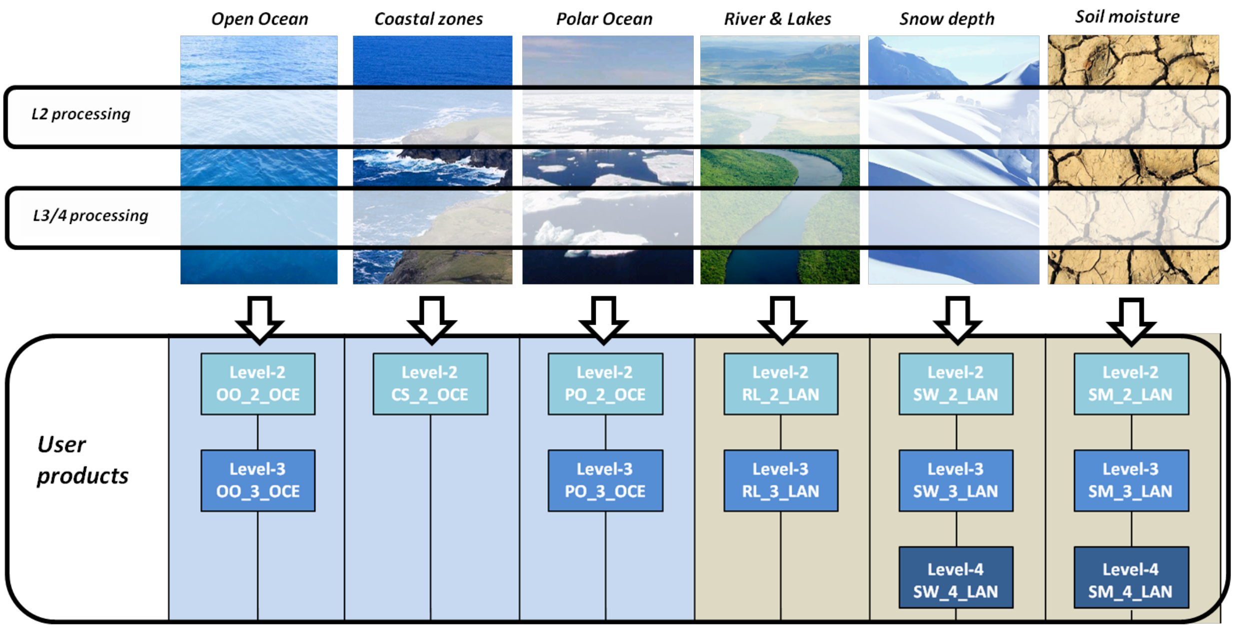


Figure 2.1: Overall view of the LOTUS prototype datasets.

2.2 Data Processing Overview

This section describes the overall design of the LOTUS processing schemes for the product generation. The data flow diagram in Figure 2.2 shows the principal segments of the LOTUS processing chain.

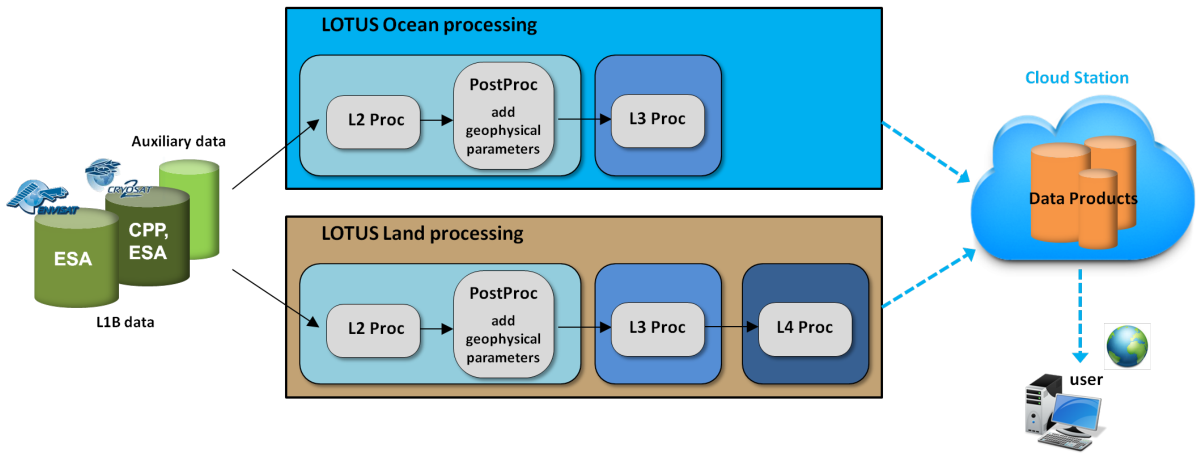


Figure 2.2: Flow chart describing the LOTUS data product generation.

The LOTUS processing chains consist of refined scientific algorithms that have been designed and developed for the different surface types of interest for the Sentinel-3 mission over both the ocean and land. Processing takes place at the premises of each partner responsible of an algorithm, and data products are coordinated by CLS for distribution to the science community through the CLS cloud server (linked to the LOTUS project website http://www.fp7-lotus.eu/). Those products only that are relevant to end-users and institutions are then disseminated by the web feature service.

CLS is thus responsible of the global processing prototype and data distribution. And each partner is responsible for algorithm studies and processing chain developments related to specific observed surfaces.

2.3 Processing Levels

As for all altimeter processing chains, the processing is split in different levels (from 1 to 4), and this, for any surface type identified as target of interest for the Sentinel-3 mission. Each processing level is associated to its data output tagged with the same level. The following diagram gives a classical processing architecture as developed for the LOTUS project, in which processing chains and products have been identified.

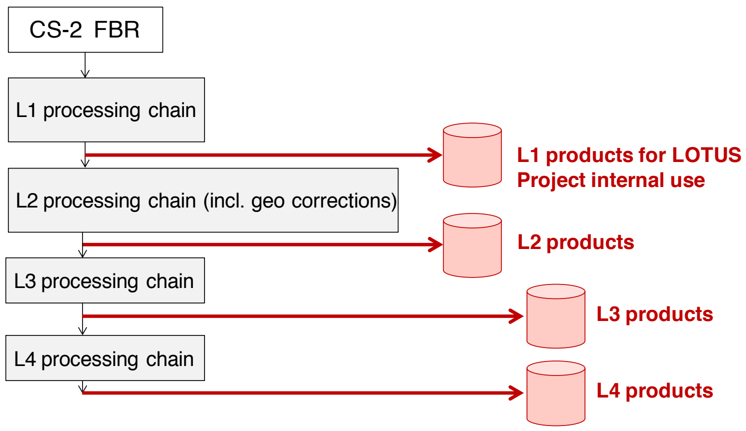


Figure 2.3: General flowchart of the LOTUS processing chain [D3.3 LOTUS report].

The data levels that are in use for the LOTUS project are described below according to their standard definitions.

2.3.1 Processing level conventions in altimetry

For delay-Doppler altimetry, **Level-1BS products** (not disseminated to the users) are geo-located complex (I and Q) echoes related to a given location on the ground, after slant range correction. No averaging of individual waveforms (i.e. multi-looking/stacking) is applied. These datasets that are of particular interest to the geophysical retrieval algorithm developers (over ocean, land and ice surfaces) of the LOTUS consortium are however not generated by the official processing chains of the Cryosat-2 mission. The only way to get at it and to successfully carry out these studies was to re-use the CryoSat-2 Processing Prototype chain (CPP) from CNES and to adapt it to the needs of each LOTUS partner and for each application (by developing several processing options depending on the observed target). This processing chain, inherited from the Sentinel-3 ground segments, allows to generate level-1BS Cryosat-2 SAR-mode data, and higher level products, starting from Full-Bit-Rate (FBR) data that are provided on request by ESA. Large amount of data has been already processed by the CPP over open ocean and shown high quality performances as such that it can be used with perfect confidence. The level-1BS product is used internally to the LOTUS project to produce data level 2 products with refined algorithms.

This product includes geo-located and calibrated radar echoes (i.e. stack of Doppler beams) with ancillary information (latitude and longitude coordinates, altitude, satellite velocity, roll and pitch mispointing angles of the antenna derived from the star tracker information, landmarks (land/water), tracker range, geophysical and environmental corrections) at high sampling rate (20-Hz). Also it should be noted that a precise orbit determination is used in the product (the reference surface for the altitude and the orbital rate of the satellite is the T/P reference ellipsoid and not WGS84).

Level-1BS products are identified to become operational on Sentinel-3 for promoting expertise activities in SAR altimetry among a greater number of users than any previous altimetric missions.

Other processing methods developed in the LOTUS project are based on the use of level-1B products (averaged power echoes, also known as multi-looked/stacked power echoes) for which ESA or CPP data providers are considered indifferently. The processing-center information given in the name of the dataset files will permit the users to identify the data sources that have been used to generate level-2 LOTUS data products.

Level-2 product: It includes high-resolution along-track geophysical quantities (altimeter range, orbital altitude and geophysical corrections, with significant wave height and wind-speed information for ocean) derived from the processing of the measurements data provided into the Level-1 product, then time-tagged, precisely located and corrected for geophysical and environmental effects (tropospheric correction, ionospheric correction, solid earth, ocean and pole tides, etc.). Different Level-2 processes are in use according to different targets and will thus generate more than one Level-2 product from the same Level-1 product. Data at this level are used by scientists or engineers that already work in the field.

Level-3 product: It is intended to ease the use of the CryoSat-2 measurements (as the ones that will be provided by Sentinel-3 SRAL measurements) for the end users. Along-track data are ready for immediate use in applications, validated (off-record data are edited), corrected and/or inter-calibrated. Two kinds of data products are designed regarding the target end applications.

- Ocean data products take benefit of multi-mission cross-calibration processing (based on combination with Jason-like altimeter mission) to ensure that all flows from all satellites provide consistent and accurate information. Furthermore, the resolution and sampling are optimised compared to Level 2 products.
- Land data products contain along-track filtered geophysical parameters. For river and lake studies, data points contaminated by land are excluded from the analysis by applying land-water mask.

Level 4 products are of two kinds:

- gridded sea level anomalies (SLA) over ocean. No ocean level-4 products are produced since this kind of products is only meaningful when merging different satellites which is out of scope of the LOTUS project. It is also important to note the limitations of these high-level satellite products when using them, since it is not taking benefit of the improvement in the along-track resolution anymore. The objective of LOTUS is precisely to demonstrate the potential improvement brought by the SAR processing in the along-track direction.
- Land data products have undergone rigorous selection and retracking processes that aim at extracting the most relevant observations on the hydrology of basins such as time series of water levels and discharge rates, which could be directly used by hydrologists for studies of regional climate variability as well as for socio-economic applications (e.g., water resources management, navigation, and flood hazards).

2.3.2 Level-1 processing

The objective of the LOTUS project is to develop new processing scheme for extracting high-resolution geophysical parameters from SAR-mode data in preparation to the upcoming Sentinel-3 mission. The SRAL instrument which will be embarked aboard Sentinel-3 is a new generation of radar altimeter, with new features and capabilities compared to conventional LRM altimeter. No data products based on this SAR-mode data are provided or used operationally yet. Fortunately Cryosat-2 mission that operates in-orbit for the first time a SAR-mode altimeter makes available (to any users and on request) raw-like data (full-bit-rate data) allowing investigations and developments within the LOTUS project.

However this mode does not operate in global and is activated mainly over sea-ice affected areas, but also in restricted open ocean zones and over in-land water basins (see http://cryosat.mssl.ucl.ac.uk/qa/mode.php for more details about the operational geographical mode mask of SIRAL/Cryosat-2 altimeter). Some processing methods (specially designed for soil moisture and snow depth) are applicable to other type of reflecting surfaces (typically in arid and semi-arid regions, and snow-covered terrains), which are unfortunately not covered in Cryosat-2 SAR-mode (or eventually in lack of ground-truth data). These particular methods will establish their geophysical retrieval feasibility and the development of downstream services by means of LRM altimeter data (Envisat, Jason-2...).

The development of the other processing methods is carried out from Cryosat-2 SAR-mode data (L-1BS or L1B products) that originate either from the adapted CPP chain or the ESA one.

Two kinds of Level-1BS product are employed regarding the target end applications to be considered:

- One is dedicated to open ocean investigations for which the nominal CPP processing chain remains unchanged (with no specific process). The original along track sampling resolution in SAR mode is kept unchanged (and not widened by a factor of approximately 1.3 in the along track direction with the azimuth Hamming weighting function) maintaining the benefits of SAR altimetry. In addition the pulse-to-pulse correlation properties as what is seen on RA-2 with the cross track PTR, is not modified.

In SAR mode, starting from the square-law detected Doppler beams (looks) of the stack (as shown in Figure 2.4), two types of signal are derived:

- The multi-looked SAR-mode power echo, which is simply obtained by an incoherent summation in azimuth direction of the beam waveforms of the stack. Adapted retracking algorithm fits the SAR echo with echo models (analytically computed or pre-computed by simulation) to retrieve the range, SWH and sigma-naught.
- The distribution of power across the beams after integrating the power in range. It is expected that this distribution of power reflects the along-track antenna gain pattern, i.e. Gaussian. Any distortion of the Gaussian shape may not be seen in the multilooked echo, but would give amount of relevant information to be analyzed, in particular the along-track mispointing angle and the ocean surface (ocean surface roughness, slope of the surface, high reflectivity surface). By fitting the distribution with a Gaussian-like model or a distorted one, the estimated parameters will inform about the beam behaviour. The characterization of the ocean surface (e.g. stack width, kurtosis, skewness) from stack data is one of the priorities raised by the expert team of the mid-term CP4O project [Cotton, 2014]. It has been implemented in the LOTUS processing chain.

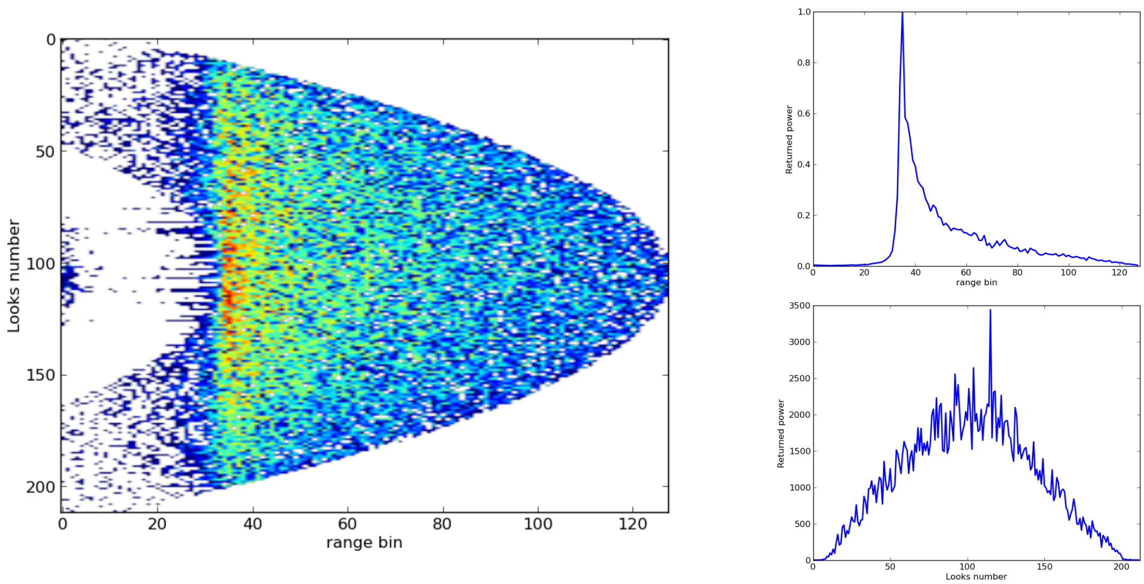


Figure 2.4: Left panel: Stack of co-located Doppler beams over open ocean. Upper-right panel: Multilooked SAR power echo. Lower-right panel: Distribution of power among looks in the stack.

- The other Level-1BS product is intended to provide relevant information to the Level-2 processing scheme for easing the extraction of high-resolution water body quantities in regions where the impact of land and ice elements having much higher brightness temperatures than water is significant (enhancing surface topography measurements in the coastal zone, sea ice regions and over in-land rivers, their tributaries and lakes).

A weighting function (e.g. Hamming function) is applied in azimuth direction in order to limit the effect of azimuth ambiguities originating from very high scattering targets that may aliased the received signal in azimuth. This function also allows cleaning the beams pointing off-nadir from spurious signals that may originate by the spread of high levels of energy contained in the main lobe into side lobes through the azimuthal impulse response. Nevertheless a side effect of the weighting is to degrade the azimuthal resolution (typically ~300m to ~450m in the CryoSat-2 configuration). A second process in the chain consists in zero padding the Doppler beam waveforms doubling their extension in order to, on the one hand, broad the leading edge of the waveform and, on the other hand, avoid aliasing of the signal that would occur by squaring the detected signal typically over specular surfaces in polar ocean or coastal

For such specular surfaces, the application of an along-track weighting function in the data processing is needed as it is underlined in the following figures. In left panel, a stack of co-located Doppler beams generated with no specific process shows spurious signals (nadir clutters) as parabola features present in the leading edge of the SAR echo waveforms that lead to an error of estimation of the water height (see in the bottom panel). The application of a Hamming window in azimuth direction eliminates clearly the nadir clutters and thus the side lobes effects as shown in the right panel and in the SAR-mode echoes displayed in the bottom panel.

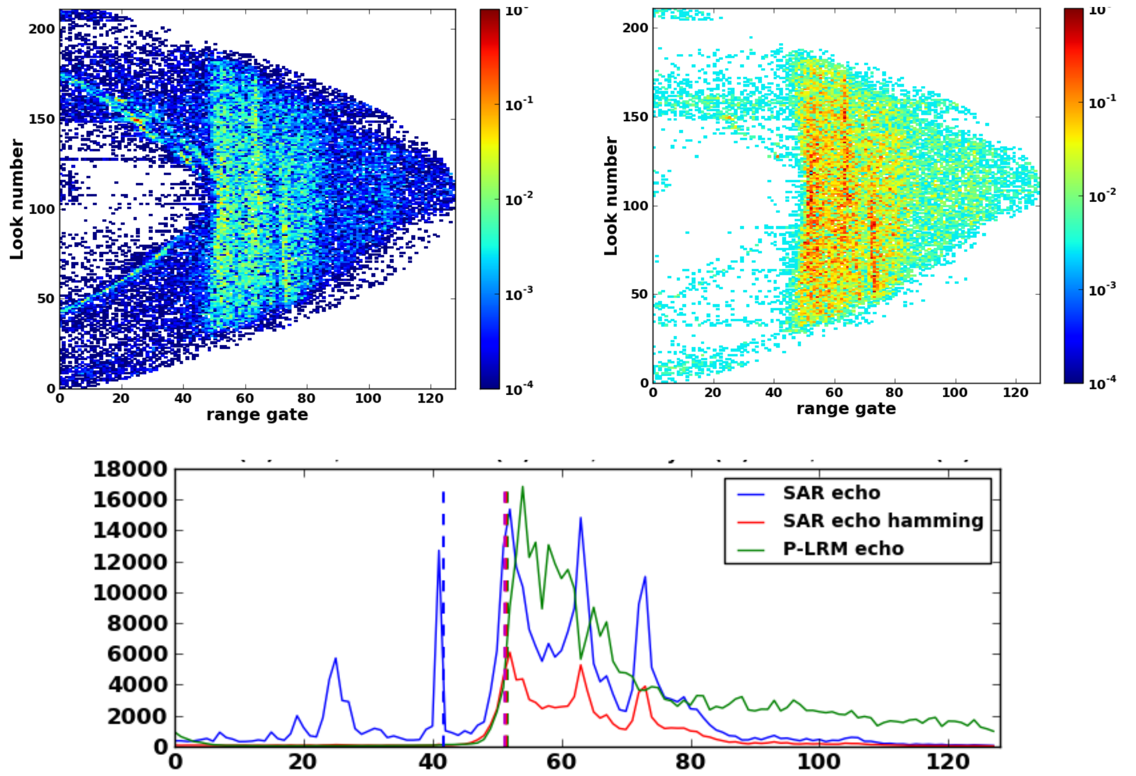


Figure 2.5: Upper panels: Stack of co-located Doppler beams over river basin with (right) and without (left) application of an along-track Hamming weighting function. Lower panel: Associated multilooked SAR power echoes and Pseudo-LRM waveform.

The Cryosat-2 multilooked SAR power echoes are made available by CPP or ESA with however some possible discrepancies due to the separate design and development of these chains. No comparisons have been done yet between them.

2.3.3 Higher-level processing

In the LOTUS activities of development, CLS is responsible for ocean processing chains, DTU for polar ocean and River and Lakes processing chains, Starlab for coastal ocean and Snow Depth processing chains and NU for soil moisture processing chain.

For each product the associated algorithm is described in the Algorithm Theoretical Basis Documents (ATBDs) written in WP1 and WP3. A more detailed overview of all the processing levels is given in the D3.3 deliverable.

All reports are available on the LOTUS's website at http://www.fp7-lotus.eu/Publications/Deliverables.

2.4 Products generation and data flow

As depicted in the Figure 2.2, the entire LOTUS data processing system is a highly complex undertaking which consists of separate entities that are each involved in the global data flow bringing the data up to the users.

In first, the CPP chain generates level-1BS products with specific processing options (weighting function and oversampling method applied or not according to the observed target) over the selected geographical zones. Other data sources also contribute to supply the project's needs with level-1B data and eventual ancillary data. After acquisition of the level-1BS data in CLS database, geophysical and environmental corrections are added to these products to enable the computation of corrected water level quantities at higher level.

The set of level-1BS (or level-1B) products responds to the requirements of the target application that is studied. The level-2 processing is then conducted individually by each partner. Higher-level products (level-3 and level-4 if needed) are generated afterwards. Each prototype dataset is finally stored into the CLS server to be disseminated to the users.

2.5 Water products classification

User community can access to the disseminated demonstration products from Level-2 to higher levels. In order to better serve the user community with products having geophysical information contents that are in line with their needs, but also enable faster and more effective data use, the products are divided into independent datasets based on the type of water surface they contain (e.g., open ocean, sea ice, river/lake). For each targeted test area, a specific algorithm that is relevant for the retrieval of most valuable geophysical parameters is applied.

Several water products are thus defined as follows:

- Level 2 and 3 Open oceans product,
- Level 2 Coastal seas product,
- Level 2 and 3 Polar ocean product,
- Level 2 and 3 Rivers and lakes product,
- Level 2, 3 and 4 Soil moisture product, and
- Level 2, 3 and 4 Snow depth (Snow water equivalent) product.

Each product dedicated to specific applications is demonstrated in the selected case study regions.

2.6 User products list

To ease the online dissemination and the users' data handling, the type of the data is clearly identifiable in the filename of the datasets. For the ocean data collection, the filename will contain the identifier "_OCE", and for the land data products the identifier "_LAN". The identifiers of the surface target and level of processing are also indicated. For example "PO_3_OCE" designates the Level-3 Polar ocean products.

The following table lists the products generated in the LOTUS project for each application and processing level (Level 2, 3, and 4 where they are applicable).

Table 2-1: User products list

Product ID	Level	Description
OO_2_OCE	2	1Hz and 20Hz SIRAL CryoSat-2 parameters (SAR/PLRM) over Open Ocean
OO_3_OCE	3	Cross-calibrated SIRAL CryoSat-2 parameters (SAR/PLRM) over Open Ocean
CS_2_OCE	2	20Hz SIRAL CryoSat-2 parameters (SAR/PLRM) over Coastal areas
PO_2_OCE	2	1Hz and 20Hz SIRAL CryoSat-2 parameters (SAR/PLRM) over Polar Ocean
PO_3_OCE	3	Cross-calibrated SIRAL CryoSat-2 parameters (SAR/PLRM) over the Polar Ocean (using global arc cross calibration)
RL_2_LAN	2	20Hz SIRAL CryoSat-2 parameters (SAR/PLRM) over in-land Rivers and Lakes
RL_3_LAN	3	Edited 20Hz SIRAL CryoSat-2 parameters (SAR/PLRM) over in-land Rivers and Lakes and time series at virtual stations
SW_2_LAN	2	Not for release
SW_3_LAN	3	20Hz filtered Snow Depth estimates (from SIRAL Cryosat-2 SAR data)
SW_4_LAN	4	Snow depth time series
SM_2_LAN	2	Not for release (data too noisy for 20Hz spatial resolution to be useful)
SM_3_LAN	3	Along-track soil moisture mean estimates (SMMEs) at spatial resolutions determined by data high frequency variation and locations bounded by DREAM extents
SM_4_LAN	4	Time series for each SMME

2.7 Datasets assessment

The quality of the Cryosat-2 ocean and land prototype data have been assessed in comparison with independent datasets (used as reference) that include multi-mission conventional low resolution altimetric data (i.e. Jason-2 and Altika), output from ocean models as well as in-situ data (tide gauges, wind observations from offshore station). The results of the assessment are given in deliverable D4.5, which is available on the LOTUS's website at http://www.fp7-lotus.eu/Publications/Deliverables.

3 Models and standards

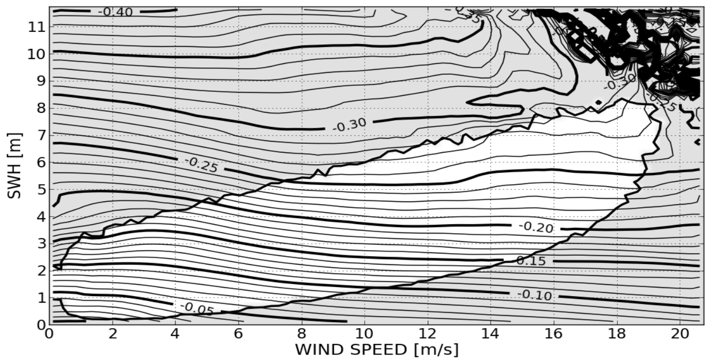
The LOTUS datasets adopt corrections from up-to-date models and standards that are relevant for the retrieval of geophysical parameters (notably for the correction of range measurements). Most of these corrections are used in Jason-2 GDRs products and have been extensively validated by Ocean Surface Topography Science Team (OSTST), demonstrating excellent quality and consistency with Jason-1 GDRs.

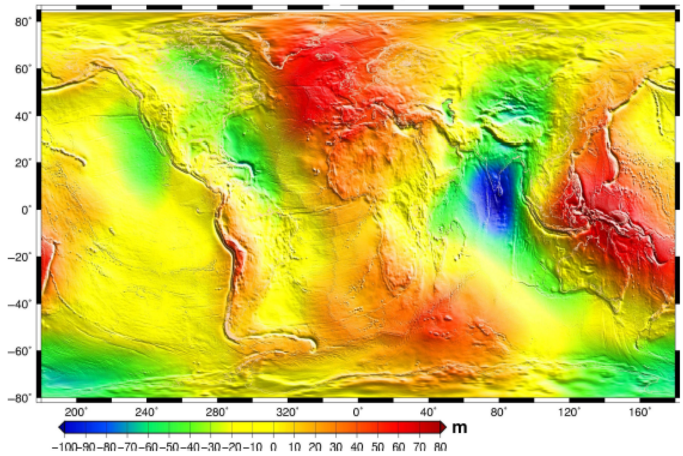
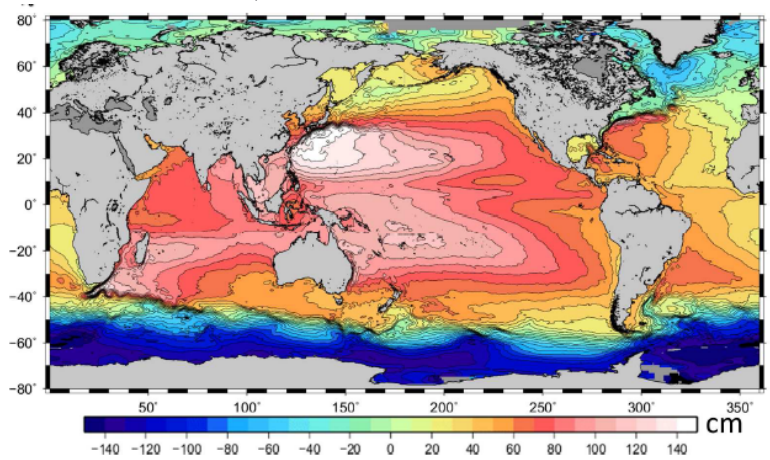
Specific dynamic geophysical correction data (model corrections for wet troposphere and ionosphere) are used to process sea height measurements for the particular Cryosat-2 mission, since the satellite does not carry a microwave radiometer to correct for water vapour path delay, and only operates at a single frequency (dual frequency allowing a direct estimate of the ionospheric delay) . It is also of high importance for any other missions which possess onboard MWR and C/Ku-bands radar antenna (e.g., Sentinel-3), to use model corrections since these corrections are less subject to instrumental instabilities and possible drifts that affect the corrections of some altimetric missions. For sea ice covered regions, model is also preferred to radiometer observations that are frequently distorted due to the presence of sea ice.

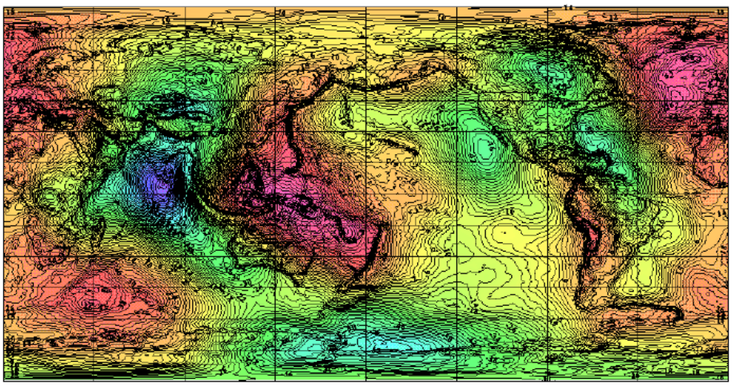
The table below summarizes the models and standards that are adopted in the LOTUS data products.

Table 3-1: Products corrections and models overview

Parameter	Description
Dry troposphere	Model dry tropospheric correction is computed at the altimeter time-tag from the interpolation of 2 meteorological fields that surround the altimeter time-tag. A dry tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for dry tropospheric range delays of the radar pulse. From European Center for Medium Range Weather Forecasting
Wet troposphere	Model wet tropospheric correction is computed at the altimeter time-tag from the interpolation of 2 meteorological fields that surround the altimeter time-tag. A wet tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for wet tropospheric range delays of the radar pulse. From European Center for Medium Range Weather Forecasting

Parameter	Description
Ionosphere	GIM ionospheric correction from NASA/JPL must be added (negative value) to the instrument range to correct this range measurement for ionospheric range delays of the radar pulse.
Ocean tide and loading tide	Geocentric ocean tide height (solution 1): GOT4.8 from GSFC Includes the loading tide and equilibrium long-period ocean tide height. The permanent tide (zero frequency) is not included in this parameter because it is included in the geoid and mean sea surface.
Solid Earth tide	Solid earth tide height is calculated using Cartwright and Taylor tables and consisting of the second and third degree constituents. The permanent tide (zero frequency) is not included. From Cartwright and Edden [1973] Corrected tables of tidal harmonics - J. Geophys. J. R. Astr. Soc., 33, 253-264.
Pole tide	Computed from Wahr [1985] Deformation of the Earth induced by polar motion - J. Geophys. Res. (Solid Earth), 90, 9363-9368.
Combined atmospheric correction	Also known as high frequency fluctuations of the sea surface topography which contains the combined atmospheric corrections (from MOG2D model + inverse barometer)
Sea State Bias	A sea state bias correction must be added (negative value) to the instrument range to correct this range measurement for sea state delays of the radar pulse. A SSB correction has been computed in 2D domain (SWH and wind speed) for the open ocean datasets in both SAR-mode and LRM. Since the SAR CryoSat-2 SSB cannot be directly calculated within the restricted areas, the LRM SSB produced from a non-parametric method applied at crossovers is used to derive the sea state bias in SAR mode. The SSB model is approximately -3.6% SWH and -3.3% SWH for SAR-mode and LRM/PLRM respectively. See D3.4 deliverable for more details about this model. 
Mean Sea Surface CLS	MSS_CNES_CLS-2011 from CLS/CNES 2011 is the mean sea surface height above T/P reference ellipsoid computed from 16 years (1993-2009) of satellite altimetry data from a variety of missions. The model is available in global between 80°S to 84°N.

Parameter	Description
	 Refer to http://www.aviso.oceanobs.com/en/data/products/auxiliary-products/mss/index.html for more details on this model.
Mean Sea Surface DTU	MSS_DTU10 from DTU 2010 is the mean sea surface height above T/P reference ellipsoid computed from 16 years (1993-2009) of satellite altimetry data from a variety of missions, based on the use of dedicated retracking to edit non-ocean waveforms. The model is available in global with no voids in the Artic ocean. Refer to http://www.space.dtu.dk/english/Research/Scientific_data_and_models/Global_Marine_Gravity_Field for more details or this model.
Mean Dynamic Topography	MDT_CNES-CLS09_v1.1 from CLS/CNES: the sea surface height (mean sea surface above the geoid, corresponding to mean geostrophic currents and its changes. This model is computed from satellite altimetry data from a variety of missions and relative to a 7 years (1993-1999) mean profile.  Refer to http://www.aviso.oceanobs.com/en/data/products/auxiliary-products/mdt/index.html for more details on this model.
Geoid	Height above the T/P reference ellipsoid based on the EGM1996 earth gravity model from NASA/GSFC. The EGM96 geopotential model has been used to calculate point values of geoid undulation on a 0.25 x 0.25 degree grid that

Parameter	Description
	spans the latitude range +85.0 deg. to -85.0 deg.  More information on EGM96 can be found at http://cddisa.gsfc.nasa.gov/926/egm96/egm96.html
Bathymetry	Ocean depth or land elevation based on the digital topographic model DTM2000 from NASA/GSFC. The bathymetric information in DTM2000.1 (originating from Smith and Sandwell's [1994] global sea floor topography) has significant differences with the ETOPO5 bathymetric model. The mean and standard deviation of these differences is 10 m and 270 m, respectively.

Those additional quantities associated with geophysical parameters that may have an effect on the measurements or their application, are incorporated within datasets. These ancillary variables are the wind speed model, distance to coast, land-water mask, and dry earth model, to name but a few, and are provided within the L2 data products for those that are relevant for the end-users. See the data description section to get more information about these ancillary fields.

4 Data format

This section presents the data storage format used for LOTUS data products.

4.1 NetCdf format and CF convention

All LOTUS data products are designed as GDR-like product including the retrievals from dedicated and innovative algorithms and all relevant corrections (wet and dry tropospheric corrections ...) that are used to derive physical measurements. These products are based on the use of the network common data format (NetCDF) to store data in an extremely flexible and self describing manner. This format was originally developed by the Unidata Program Center in Boulder, Colorado, and has been extended by user contributions, in an effective way to store large scale scientific and engineering datasets. A wide range of supporting tools is made also available (free and open-source) to read and write netCDF files for many languages. NetCDF is adopted as a standard data format and is used to a large extent in many operational oceanography systems. NetCDF data are largely preferred because of the advantages noted below:

- Self-Describing. A netCDF file includes information about the data it contains.
- Architecture-independent. A netCDF file is represented in a form that can be accessed by computers with different ways of storing integers, characters, and floating-point numbers.
- Direct-access. A small subset of a large dataset may be accessed efficiently, without first reading through all the preceding data.
- Appendable. Data can be appended to a netCDF dataset along one dimension without copying the dataset or redefining its structure. The structure of a netCDF dataset can be changed, though this sometimes causes the dataset to be copied.
- Sharable. One writer and multiple readers may simultaneously access the same netCDF file.

See Unidata NetCDF pages for more information, and to retreive NetCDF software package on:
http://www.unidata.ucar.edu/packages/netcdf/index.html

The data files also follow the Climate and Forecast (CF) NetCDF metadata conventions which are widely used for atmospheric, ocean, and climate data, providing a practical standard for storing and sharing of scientific data.

4.2 The Common Data Language

The Common Data Language (CDL) is used to describe the content of a dataset.

The CDL is an ASCII description of the binary data in a netCDF file that is human readable. CDL text can be generated from netCDF file via the command *ncdump*. The netCDF utility *ncgen* converts the CDL text file to the netCDF binary file.

A CDL description of a netCDF dataset takes the form:

```
netCDF name {
dimensions: …
variables: …
// global attributes:
data: …
}
```

which contains:

- the file name,
- the dimensions part that is used to define the shape of one or more of the multidimensional variables,
- the variables part contains the attributes about the data providing information for each variable (name, data type and shape),
- the data values for each variable,
- global attributes are also assigned to the netCDF dataset, providing general comments regarding the production date, the creation, contact and other relevant information. The global attributes keyword is preceded by double slash character to indicate commentary part in the CDL text file.

The dimensions, variables and global attributes parts may be displayed from the dataset file using "*ncdump –h*" command.

An example of a netCDF L2 Open Ocean file is given below (only a small part of the full CDL text is represented herein):

```
netcdf CS_CPP_OO_2_OCE_12E20E40N46N_20130305T001352_20130305T001514 {
dimensions:
        time_20hz = 1101 ;
        time_1hz = 57 ;
variables:
        double time_20hz(time_20hz) ;
                time_20hz:long_name = "20Hz time (sec. since 1950-01-01)" ;
                time_20hz:standard_name = "time" ;
                time_20hz:calendar = "gregorian" ;
                time_20hz:units = "seconds since 1950-01-01 00:00:00.0" ;
                time_20hz:tai_utc_difference = -35. ;
                time_20hz:leap_second = "0000-00-00 00:00:00" ;
        double time_1hz(time_1hz) ;
                time_1hz:long_name = "1Hz time (sec. since 1950-01-01)" ;
                time_1hz:standard_name = "time" ;
                time_1hz:calendar = "gregorian" ;
                time_1hz:units = "seconds since 1950-01-01 00:00:00.0" ;
                time_1hz:tai_utc_difference = -35. ;
                time_1hz:leap_second = "0000-00-00 00:00:00" ;
        int lat_20hz(time_20hz) ;
                lat_20hz:_FillValue = -2147483648 ;
                lat_20hz:long_name = "20Hz latitude" ;
                lat_20hz:units = "degrees_north" ;
                lat_20hz:scale_factor = 1.e-06 ;
                lat_20hz:comments = "Positive latitude is North latitude, negative latitude
              is South latitude" ;
        int lat_1hz(time_1hz) ;
                lat_1hz:_FillValue = -2147483648 ;
                lat_1hz:long_name = "1Hz latitude" ;
                lat_1hz:units = "degrees_north" ;
                lat_1hz:scale_factor = 1.e-06 ;
                lat_1hz:comments = "Positive latitude is North latitude, negative latitude is
              South latitude" ;
        int lon_20hz(time_20hz) ;
                lon_20hz:_FillValue = -2147483648 ;
                lon_20hz:long_name = "20Hz longitude" ;
                lon_20hz:units = "degrees_east" ;
                lon_20hz:scale_factor = 1.e-06 ;
                lon_20hz:comments = "East longitude relative to Greenwich meridian" ;
        int lon_1hz(time_1hz) ;
                lon_1hz:_FillValue = -2147483648 ;
                lon_1hz:long_name = "1Hz longitude" ;
                lon_1hz:units = "degrees_east" ;
                lon_1hz:scale_factor = 1.e-06 ;
                lon_1hz:comments = "East longitude relative to Greenwich meridian" ;
        byte flg_val_20hz(time_20hz) ;
                flg_val_20hz:_FillValue = 127b ;
                flg_val_20hz:long_name = "20Hz quality flag compute from an iterative
              process. A low pass filter is
                   applied on SAR SLA (fc=50km), then measurements higher than 3*std(SLA-
              SLAfiltered) are
                   edited. If = 0: valid measurement. If =1: edited measurement" ;
                flg_val_20hz:flag_values = 0, 1 ;
                flg_val_20hz:coordinates = "lon_20hz lat_20hz" ;
```

```
...
// global attributes:
        :CONVENTION = "CF-1.6" ;
        :title = "Ocean L2 GDR-like product" ;
        :source = "radar altimeter" ;
        :history = "Thu Apr  9 12:37:11 2015" ;
        :contact = "CLS: Thomas.moreau@cls.fr, Matthias.Raynal@cls.fr,
CNES:Francois.Boy@cnes.fr";
        :reference = "CPP V14" ;
        :processing_center = "CLS/CNES" ;
        :mission_name = "CryoSat-2" ;
        :altimeter_sensor_name = "SIRAL" ;
        :first_meas_time = "20130305T001352" ;
        :last_meas_time = "20130305T001514" ;
data:

 time_20hz = 1997396607.97912, 1997396608.0263, 1997396608.07347,
    1997396608.12065, 1997396608.16782, 1997396608.21499, 1997396608.26216,
    1997396608.30933, 1997396608.3565, 1997396608.40368, 1997396608.45085,
    1997396608.49802, 1997396608.54519, 1997396608.59236, 1997396608.63953,
    1997396608.6867, 1997396608.73388, 1997396608.78105, 1997396608.82822,
...
}
```

5 Structure of the LOTUS data files

In this section, a description of individual data products is provided.

As it is previously explained, the LOTUS products are organized into different targets and processing levels in order to implement a packaging scheme suited to archiving and dissemination to users. Each product package contains geophysical information that are retrieved by the new algorithms developed in the frame of the LOTUS project, around which associated auxiliary parameters (orbital information, roll/pitch mispointing, environmental/geophysical corrections, models, land/water mask) are embedded. The correction parameters and models that are necessary to compute the SLA and water level are listed in previous section.

5.1 OO_2_OCE Open Ocean datasets

5.1.1 Description of the products

The OO_2_OCE product contains level-2 ocean geophysical parameters derived from Cryosat-2 SAR-mode data. This product also includes pseudo pulse-limited (PLRM) estimates derived from an incoherent processing of the same returning echoes, that aim at assessing the in-orbit performances of the SAR-mode data compared to well-known LRM-like data. Both retrievals are made at the same points along the ground tracks with thus identical time and location allowing for their direct comparison without the need to apply any environmental and geophysical corrections that may contribute to some potential differences and lead to unclear conclusions regarding the comparison between the different processing approaches. SAR-mode and PLRM processing for open ocean are described in the deliverable document 1.3.

The altimeter estimates are obtained using a specific ocean retracker in PLRM and in SAR mode.

A conventional Brown ocean retracker based on unweigthed least square estimations (also known as MLE) which are traditionally used with LRM echoes, is applied to the PLRM power waveforms for estimating the different geophysical parameters. A 4-parameter estimator is considered for adequately fitting the measured waveforms with the return power model since the CryoSat-2 satellite exhibits unstable off-nadir mispointing angle in flight.

In SAR mode, a 3-parameter ocean retracker fits the numerical multilooked waveform model to a 20 Hz level-1B SAR waveform using the LSE method and retrieves the corresponding geophysical variables (range, significant wave height, backscattering coefficient) and fitting quality information.

Data are provided at two frequencies, 1 Hz and 20 Hz, and are sequentially split into NetCDF files, with one file per pass (one pass is a half-orbit, descending or ascending orbit). For each file, the datasets are organised in a standard structure that is close to the one used for Sentinel-3 products where 1 Hz and 20 Hz parameters are stored into separated variables. This way ensures ease of use for the user.

5.1.2 Dataset Areas

Three areas of interest have been selected for the prototype open ocean datasets: NE Atlantic, Bay of Singapore and Adriatic Sea.

Table 5-1: NE Atlantic open ocean dataset

Parameter	Value
Geographical Coverage	N.E. Atlantic 13W – 15E, 48N – 59N
Temporal Coverage	1st May 2012 – 30th April 2014

Table 5-2: Bay of Singapore open ocean dataset

Parameter	Value
Geographical Coverage	Bay of Singapore 98E – 121E, 4S – 25N
Temporal Coverage	1st May 2012 – 30th April 2014

Table 5-3: Adriatic Sea open ocean dataset

Parameter	Value
Geographical Coverage	Adriatic Sea 12E – 20E, 40N – 46N
Temporal Coverage	1st May 2012 – 30th April 2014

5.1.3 Product Parameters

This product contains several parameters:

- The main ocean geophysical variables (Sea surface height, wave height, sigma-nought and wind) for SAR-mode and PLRM. File names appended with "_sarm" are related to SAR-mode parameters, and those appended with "_plrm" are related to PLRM parameters
- Correction measurements (meteo fields, tides, …) that are used to compute the corrected sea level height
- A 1Hz valid flag to remove the main part of data outliers. This flag is built from the validation task of Cryosat-2 performed by the CLS Space Oceanography Division. On-board retracked data are used for generating the flag in SAR-mode areas. With noise statistics and the shape of SAR altimeter waveforms so markedly different from those of the on-board LRM, this flag might be not efficiently adapted to edit SAR mode data.
- A 20Hz valid flag to remove the main part of data outliers. This flag is built from an iterative process removing outliers. This flag has been computed from SAR SSHA estimations

Table 5-4: L2 Open Ocean parameters description

Short Name	Long Name	Unit	Description
Point characteristics			
time_20hz time_1hz	20 Hz and 1 Hz time	second	seconds since 1950-01-01 00:00:00.0
lat_20hz lat_1hz	20 Hz and 1 Hz latitude	degree north	Latitude of the point
lon_20hz lon_1hz	20 Hz and 1 Hz longitude	degree east	Longitude of the point
Ocean parameters			
range_20hz_sarm range_1hz_sarm	20Hz and 1Hz corrected altimeter range in SAR mode	m	Ocean SAR numerical retracking. All instrumental corrections included. Constant bias (accounting for internal path delay, ..) inherited from preliminary analyses done on LRM CY2 data (at crossovers with Jason-2) is applied to the SAR range values
range_20hz_plrm range_1hz_plrm	20Hz and 1Hz corrected altimeter range in PLRM	m	Ocean LRM retracking. All instrumental corrections included Constant bias (accounting for internal path delay, ..) inherited from preliminary analyses done on LRM CY2 data (at crossovers with Jason-2) is applied to the PLRM range values
sig0_20hz_sarm sig0_1hz_sarm	20Hz and 1Hz corrected ocean backscatter coefficient in SAR mode	dB	Constant bias -283.8dB (accounting for instrumental corrections, ...) is applied to SAR sigma0 values to make these data consistent with Jason-2

Short Name	Long Name	Unit	Description
sig0_20hz_plrm sig0_1hz_plrm	20Hz and 1Hz corrected ocean backscatter coefficient in PLRM	dB	Constant bias -288.4dB (accounting for instrumental corrections, ...) is applied to P-LRM sigma0 values to make these data consistent with Jason-2
swh_20hz_sarm swh_1hz_sarm	20Hz and 1Hz significant waveheight in SAR mode	m	All instrumental corrections included
swh_20hz_plrm swh_1hz_plrm	20Hz and 1Hz significant waveheight in PLRM	m	All instrumental corrections included
ssha_20hz_sarm ssha_1hz_sarm	20Hz and 1Hz sea surface height anomaly in SAR mode	m	The Sea Level Anomaly (aka SSHA) is computed as follows: SSHA = Altitude of satellite - corrected ocean altimeter SAR range - model ionospheric correction - model dry tropospheric correction - model wet tropospheric correction - solid earth tide height - geocentric ocean tide height solution1 (GOT solution) - geocentric pole tide height - combined corrections from MOG2D model - mean sea surface – SAR SSB
ssha_20hz_plrm ssha_1hz_plrm	20Hz and 1Hz sea surface height anomaly in PLRM	m	The Sea Level Anomaly (aka SSHA) is computed as follows: SSHA = Altitude of satellite - corrected ocean altimeter PLRM range - model ionospheric correction - model dry tropospheric correction - model wet tropospheric correction - solid earth tide height - geocentric ocean tide height solution1 (GOT solution) - geocentric pole tide height - combined corrections from MOG2D model - mean sea surface – LRM SSB
wind_speed_sol1_20hz_sarm wind_speed_sol1_1hz_sarm	20Hz and 1Hz wind speed in SAR mode	m/s	Non parametric wind speed estimation based on Labroue & Tran approach
wind_speed_sol1_20hz_plrm wind_speed_sol1_1hz_plrm	20Hz and 1Hz wind speed in PLRM	m/s	Non parametric wind speed estimation based on Labroue & Tran approach
Atmospheric/Ocean corrections			
tropo_dry_corr_model_20hz tropo_dry_corr_model_1hz	20Hz and 1Hz model dry tropospheric correction	m	Computed at the altimeter time-tag from the interpolation of 2 meteorological fields (ECMWF) that surround the altimeter time-tag. A dry tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for dry tropospheric range delays of the radar pulse.

Short Name	Long Name	Unit	Description
tropo_wet_corr_model_20hz tropo_wet_corr_model_1hz	20Hz and 1Hz model wet tropospheric correction	m	Computed at the altimeter time-tag from the interpolation of 2 meteorological fields (ECMWF) that surround the altimeter time-tag. A wet tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for wet tropospheric range delays of the radar pulse
iono_corr_gim_20hz iono_corr_gim_1hz	20Hz and 1Hz GIM ionospheric correction	m	An ionospheric correction must be added (negative value) to the instrument range to correct this range measurement for ionospheric range delays of the radar pulse
sea_state_bias_20hz_sarm sea_state_bias_1hz_sarm	20Hz and 1Hz sea state bias correction in SAR mode	m	A sea state bias correction must be added (negative value) to the instrument range to correct this range measurement for sea state delays of the radar pulse. The SAR sea state bias is derived from a non parametric solution fitted on SAR-mode Cryosat-2 CPP data.
sea_state_bias_20hz_plrm sea_state_bias_1hz_plrm	20Hz and 1Hz sea state bias correction in PLRM	m	A sea state bias correction must be added (negative value) to the instrument range to correct this range measurement for sea state delays of the radar pulse. The PLRM sea state bias is derived from a non parametric solution fitted on LRM Cryosat-2 CPP data.
mean_sea_surf_20hz mean_sea_surf_1hz	20Hz and 1Hz mean sea surface height above reference ellipsoid	m	MSS_CNES_CLS-2011
hf_fluct_corr_20hz hf_fluct_corr_1hz	20Hz and 1Hz high frequency fluctuations of the sea surface topography	m	Sea surface height correction due to air pressure and wind at high frequency
ocean_tide_sol1_20hz ocean_tide_sol1_1hz	20Hz and 1Hz geocentric ocean tide height (solution 1)	m	GOT4.8 from GSFC includes the loading tide and equilibrium long-period ocean tide height. The permanent tide (zero frequency) is not included in this parameter because it is included in the geoid and mean sea surface
solid_earth_tide_20hz solid_earth_tide_1hz	20Hz and 1Hz solid earth tide height	m	Calculated using Cartwright and Taylor tables and consisting of the second and third degree constituents. The permanent tide (zero frequency) is not included
pole_tide_20hz pole_tide_1hz	20Hz and 1Hz geocentric pole tide height	m	Deformation of the Earth induced by polar motion from Wahr
wind_speed_sol2_20hz wind_speed_sol2_1hz	Wind speed	m/s	Wind speed derived from ECMWF model
Additional information			
orb_poe_20hz	20Hz orbital altitude	m	Altitude of satellite above the reference ellipsoid
orb_poe_1hz	1Hz orbital altitude	m	Altitude of satellite above the reference ellipsoid

Short Name	Long Name	Unit	Description
flg_val_20hz	20Hz quality flag		20Hz quality flag compute from an iterative process. A low pass filter is applied on SAR SLA (fc=50km), then measurements higher than 3*std(SLA-SLAfiltered) are edited. If = 0: valid measurement. If =1: edited measurement
flg_val_1hz	1 Hz quality flag		1 Hz quality flag copied from the DUACS Cryosat-2 editing. If = 0: valid measurement. Else: edited measurement
dist_coast_20hz	20Hz distance to the closest coast	m	20Hz computed distance to the closest coast based on the use of a 1/16 deg map (GSHHG)
dist_coast_1hz	1Hz distance to the closest coast	m	1Hz computed distance to the closest coast based on the use of a 1/16 deg map (GSHHG)

Note that the sea state bias (SSB) correction that is applied to the SSHA was calculated to fit to SAR (or P-LRM) data generated from a previous processing version, and may be not optimized for the actual SAR (and P-LRM) data use in the LOTUS products. The SSB correction may be easily removed from the SSHA, as well as all those geographical corrections.

5.2 OO_3_OCE Open Ocean datasets

Not available at that time. It will be added in a new release of this document.

5.3 CS_2_OCE Coastal Sea datasets

5.3.1 Description of the product

This data product contains 20Hz Coastal Sea parameters over selected areas.

The altimeter estimates are obtained using a specific ocean retracker for SAR mode. The retracker is based in a modification of the SAMOSA model, a fully analytical waveform model, that performs the best fit to the SAR-waveform in order to simultaneously estimate the following geophysical variables: range, significant wave height, and backscattering coefficient. A Goodness of fit (GOF) parameter is also given in order to evaluate the quality of the model fit to the data.

Given that the distance to the coast is critical in this environment, two other parameters are provided within this product, i.e. the minimum distance to the coastline, and the minimum across-track distance. This is the key distance value, as it is the observation direction of the radar.

Given the requirement of high spatial resolution in coastal areas, all the estimated geophysical parameters are provided at 20 Hz.

The data is store in a NetCDF file using CF convention. Data has been processed according to the algorithm, which is described in the deliverable D1.3.

5.3.2 Dataset Areas

Two areas have been selected for the prototype coastal area datasets: NE Atlantic and Adriatic Sea.

Table 5-5: NE Atlantic coastal ocean dataset

Parameter	Value
Geographical Coverage	N.E. Atlantic + North Sea 15W – 17E, 46N – 61N
Temporal Coverage	1st May 2012 – 30th April 2013

Table 5-6: Adriatic Sea coastal ocean dataset

Parameter	Value
Geographical Coverage	Adriatic Sea 10E – 22E, 38N – 48N
Temporal Coverage	1st May 2012 – 30th April 2013

5.3.3 Product Parameters

The CS_2_OCE products contain several parameters:

- Ocean parameters: Sea surface height, significant wave height, sigma-naught
- Correction measurements (meteo fields, tides, …) that are used to compute the corrected sea level height
- The distance to the closest coast in across track direction and in all directions
- The goodness of fit of the SAMOSA retracker that evaluate the fit of the simulated waveform to the real input waveform

Table 5-7: Coastal Sea parameters description

Short Name	Long Name	Unit	Description
Point characteristics			
time	time	seconds since 1950-01-01 00:00:00.0	Timestamp
lat	latitude	degree north	Latitude of the point
lon	longitude	degree east	Longitude of the point
Ocean parameters			
ssh	Sea Surface Height	m	Sea Surface Height
swh	Significant Wave Height	m	Significant Wave Height

Short Name	Long Name	Unit	Description
sigma_nought	Backscatter coefficient	m	Backscatter coefficient
Atmospheric/Ocean corrections			
Iono_corr_gim	GIM ionospheric correction	m	An ionospheric correction must be added (negative value) to the instrument range to correct this range measurement for ionospheric range delays of the radar pulse
tropo_dry_corr_model	Model dry tropospheric correction	m	Computed at the altimeter time-tag from the interpolation of 2 meteorological fields that surround the altimeter time-tag. A dry tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for dry tropospheric range delays of the radar pulse
tropo_wet_corr_model	Model wet tropospheric correction	m	Computed at the altimeter time-tag from the interpolation of 2 meteorological fields that surround the altimeter time-tag. A wet tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for wet tropospheric range delays of the radar pulse
solid_earth_tide	Solid earth tide height	m	Calculated using Cartwright and Tayler tables and consisting of the second and third degree constituents. The permanent tide (zero frequency) is not included
ocean_tide_sol1	Geocentric ocean tide height (solution 1)	m	Includes the loading tide and equilibrium long-period ocean tide height. The permanent tide (zero frequency) is not included in this parameter because it is included in the geoid and mean sea surface
pole_tide	Geocentric pole tide height	m	Deformation of the Earth induced by polar motion
hf_fluct_corr	High frequency fluctuations of the sea surface topography	m	Sea surface height correction due to air pressure and wind at high frequency
mean_sea_surf	20Hz mean sea surface height above reference ellipsoid	m	Mean Sea Surface
Additional information			
xt_dist	Across track distance to coast	m	Distance to the closest coast in the across track direction
min_dist	Minimum distance to coast	m	Distance to the closest coast
gof	Goodness of Fit		Goodness of the fitting of the SAMOSA retracker

5.4 PO_2_OCE Polar Ocean datasets

5.4.1 Description of the product

This data product contains 20Hz Polar Ocean parameters over selected areas.

The Level-2 SIRAL CryoSat-2 prototype data sets contain the typical altimetry parameters, like the altimeter range, the sea surface height, the wind speed, significant wave height and all required geophysical corrections. As the Polar Ocean is primarily under the SAR-mode mask and as radiometer observations are frequently distorted due to the presence of sea ice, the following parameters are preferred to those classically used over the open ocean.

- Model corrections for wet and dry troposphere,
- The DTU MSS as it has no voids in the Arctic Ocean.

The data is store in ascii files with corresponding names to the original L1 data files delivered by ESA. Data has been processed according to the algorithm, which is described in the deliverable D1.3 for the Polar Ocean.

5.4.2 Dataset Areas

The availability of auxiliary data useful for validation of altimetry observations in the Polar Oceans is very limited. The test and demo areas are therefore divided into two groups.

Table 5-8: Track polar ocean dataset

Parameter	Value
Geographical Coverage	Over North pole
Temporal Coverage	2011-2013

The first group comprises of individual tracks. The tracks consist of a repeated profile crossing the Arctic Basin and a number of tracks where NASA has carried out coordinated CryoSat-2 under-flights as part of the IceBridge project. These tracks are detailed below.

Table 5-9: Svalbard polar ocean dataset

Parameter	Value
Geographical Coverage	Svalbard 0E – 40E, 75N – 85N
Temporal Coverage	2012

The second group covers the area around Svalbard where high temporal tide gauge data is available from the Ny-Ålesund station on Svalbard, see following table. This ocean area is partly covered with sea-ice in the winter and will typical be ice free in late summer.

The picture below shows the two regions:

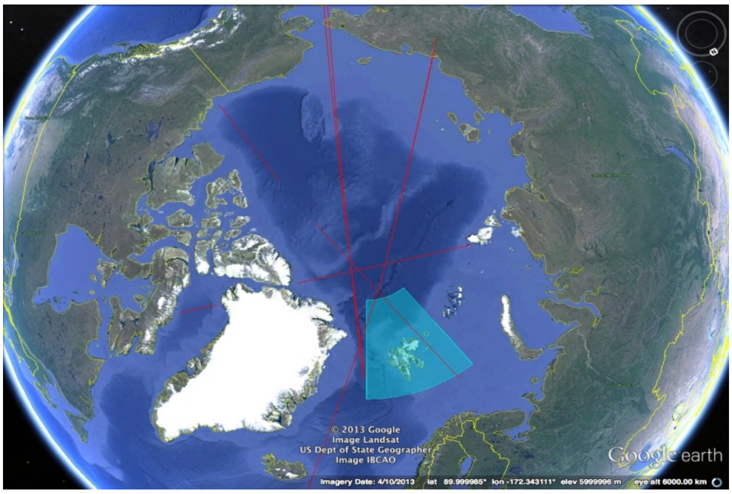


Figure 5.1: Polar Ocean datasets: Svalbard region, and along track region.

5.4.3 Product parameters

The following table lists the content of the PO_2_OCE products.

Table 5-10: Polar Ocean parameters description

Short Name	Long Name	Unit	Description
Point characteristics			
time	time		Timestamp
lat	latitude		Latitude of the point
lon	longitude		Longitude of the point
Ocean parameters			
	Combined Physical – Empirical Retracked SSHA	m	Retracked sea surface height prioritizing height from physical retracke. In case this is not available the primary peak empirical retracker is used
	Primary Peak Threshold Retracked SSHA	m	SSH from primary peak retracker
	ESA Retracked SSHA	m	SSH from ESA’s default retracker on Baseline B.
Atmospheric/Ocean corrections			
	Mean Sea Surface / Geoid	m	DTU10 MSS used
	Sum of all range corrections except	m	Sum of wet+dry troposphere (from ECMWF), tides, atmosphere

Short Name	Long Name	Unit	Description
	Ocean Tide Correction	m	Ocean tide from FES2004
Additional information			
	Retracker Type Flag		Retracker Type Flag for the combined Physical-Empirical Retracked SSHA: 1 for SAMOSA3-A Ocean Retracker 2 for SAMOSA3-A Lead Retracker 3 for Primary Peak Threshold Retracker NaN for missing points

5.5 PO_3_OCE Open Ocean datasets

Not available at that time. It will be added in a new release of this document.

5.6 RL_2_LAN River and Lake data sets

5.6.1 Description of the product

The Level-2 SIRAL CryoSat-2 prototype data sets for rivers and lakes contain 20 Hz water levels, a land-water mask value, an along-track mean river/lake levels and all required geophysical corrections.

The 20 Hz water levels are based on the empirical retracker, the Narrow Primary Peak Threshold retracker (NPPTR). For each waveform a subwaveform is extracted based on start and stop thresholds. These thresholds are found from the standard deviation of the power differences in consecutive bins (Jain et al., 2015). Once the subwaveform is extracted, the retracking bin is found by applying a threshold retracker (80 %) on the subwaveform.

In a post-processing procedure the MODIS mask is used to identify measurements over a given in-land water body. For each track that contains more than 5 measurements a robust mean water level is etimated in addition to the retracked water levels.

(References)

Maulik Jain, Ole Baltazar Andersen, Jørgen Dall, Lars Stenseng, Sea surface height determination in the Arctic using Cryosat-2 SAR data from primary peak empirical retrackers, Advances in Space Research, Volume 55, Issue 1, 1 January 2015, Pages 40-50, ISSN 0273-1177, http://dx.doi.org/10.1016/j.asr.2014.09.006.

The data is store in a NetCDF file using CF convention. Data has been processed according to the algorithm, which is described in the deliverable D2.3

5.6.2 Dataset Areas

The following areas have been selected for the prototype river and lake datasets.

Table 5-11: Denmark River and Lake dataset

Parameter	Value
Geographical Coverage	Denmark 8E – 13E, 54.5N – 58N
Temporal Coverage	16th July 2010 – 10th July 2014

The Danish area is an excellent test area to validate the algorithms that are used to process data over inland water bodies, due to a large amount of auxiliary data e.g. a high resolution DEM (10m).

Table 5-12: Thailand Chao Phraya River and Lake dataset

Parameter	Value
Geographical Coverage	Thailand Chao Phraya River 99E-101.5E, 13.25N-17N
Temporal Coverage	16th July 2010 – 8th July 2014

One of the main interests with the demo dataset is to derive time series of the water level for the river Chao Phraya. Another important application is to detect the flooding event that occurred in October 2011.

Table 5-13: Thailand Chao Phraya River and Lake dataset

Parameter	Value
Geographical Coverage	Brahmaputra River 89.5E-91.5E, 21.75N-24.25N
Temporal Coverage	13rd October 2012 – 3rd July 2014

The main interest with this demo dataset is to derive time series of water levels for Brahmaputra. To ensure coverage of the entire river all data modes (LRM, SAR, and SARIN) is processed.

Table 5-14: Amazon River and Lake dataset

Parameter	Value

Parameter	Value
Geographical Coverage	Amazon River -61E, -47E; -5S, 3N
Temporal Coverage	1st October 2012 – 9th July 2014

The Amazon River represents an excellent choice for the validation and development of the inland water algorithms. The area to be analysed is the downstream part of the Amazon River covered by the CryoSat-2 SAR mask.

5.6.3 Product Parameters

The parameters contained within this product are described below.

Table 5-15: L2 River and Lake parameters description

Short Name	Long Name	Unit	Description
Point characteristics			
time	time	second	seconds since 2000-01-01 00:00:00.0
lat	latitude	degrees_north	Latitude of the point
lon	longitude	degrees_east	Longitude of the point
Inland water/Land parameters			
wl_se	Water level/ Surface elevation	m	Water levels or surface elevations relative to WGS84, based on the Narrow Primary Peak Threshold retracker (NPPTR). The following corrections are included: model ionosphere, model dry troposphere, modelled wet troposphere, solid earth tide, ocean loading tide, geocentric pole tide.
mwl	Mean water level	m	Robust estimate of the along-track mean water level relative to EGM2008. This value is only returned if a water body contains 5 or more measurements otherwise the value 99999 is returned
Atmospheric/Land corrections			
Iono_corr_gim	GIM ionospheric correction	m	An ionospheric correction must be added (negative value) to the instrument range to correct this range measurement for ionospheric range delays of the radar pulse
tropo_dry_corr_mode l	Model dry tropospheric correction	m	Computed at the altimeter time-tag from the interpolation of 2 meteorological fields that surround the altimeter time-tag. A dry tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for dry tropospheric range delays of the radar pulse
tropo_wet_corr_model	Model wet tropospheric correction	m	Computed at the altimeter time-tag from the interpolation of 2 meteorological fields that surround the altimeter time-tag. A wet

Short Name	Long Name	Unit	Description
			tropospheric correction must be added (negative value) to the instrument range to correct this range measurement for wet tropospheric range delays of the radar pulse
solid_earth_tide	Solid earth tide height	m	Calculated using Cartwright and Tayler tables and consisting of the second and third degree constituents. The permanent tide (zero frequency) is not included
pole_tide	Geocentric pole tide height	m	Deformation of the Earth induced by polar motion
ocean_load_tide	Ocean loading tide	m	Deformation of the Earth due to the weight of the overlying ocean tide. The FES2004 loading tide model is used for this correction
geiod	Geoid	m	EGM2008
Additional information			
modis	MODIS mask value		Value that indicates the underlying surface type; water =1, land=0

5.7 RL_3_LAN River and Lake datasets

Not available at that time. It will be added in a new release of this document.

5.8 SW_2_LAN Snow Depth datasets

Conventionally, snow depth is measured at single points on the surface using snow rulers or ultrasonic ranging sensors with a centimeter accuracy. These measurements are extrapolated temporally and spatially in order to cover wider areas. Space-borne sensors, which can cover a wide swath and can provide rapid repeat global coverage, are ideally suited to augment the snow information on a regional and global scale. Snow depth estimation by means of altimeter measurements is based on the analysis of backscattering changes during a snow season.

For Sentinel-3 derived Snow Depth estimations two products are planned: a level 3 product, which contains the along-track filtered snow depth; a level 4 product which contains time series of snow depth evolution. The Level 2 product, i.e. the 20 Hz Backscatter coefficient is not delivered due to the high noise in the observable.

5.9 SW_2_LAN Snow Depth datasets

Conventionally, snow depth is measured at single points on the surface using snow rulers or ultrasonic ranging sensors with a centimeter accuracy. These measurements are extrapolated temporally and spatially in order to cover wider areas. Space-borne sensors, which can cover a wide swath and can provide rapid repeat global coverage, are ideally suited to augment the snow information on a regional and global scale. Snow depth estimation by means of altimeter measurements is based on the analysis of backscattering changes during a snow season.

For Sentinel-3 derived Snow Depth estimations two products are planned: a level 3 product, which contains the along-track filtered snow depth; a level 4 product which contains time series of snow depth evolution. The Level 2 product, i.e. the 20 Hz Backscatter coefficient is not delivered due to the high noise in the observable.

5.10 SW_3_LAN Snow Depth datasets

Not available at that time. It will be added in a new release of this document.

5.11 SW_4_LAN Snow Depth datasets

Not available at that time. It will be added in a new release of this document.

5.12 SM_2_LAN Soil Moisture datasets

Surface soil moisture estimates are derived over desert and semi-arid terrain by analysing each 20Hz waveform and computing the altimeter backscatter, applying all required corrections and scaling factors. These values are then compared with a DRy EArth Model (DREAM) which encodes the detailed variation in this parameter expected over the surface in dry earth conditions. The requirement for DREAM creation presently constrains this application to deserts and semi-arid terrain.

From a comparison of the measurement with the model it is then possible to calculate the surface soil moisture. Because the highest spatial frequency variations in this parameter (resulting from changes in small-scale roughness and surface composition) are not captured in the DREAMS, some filtering and averaging of the values is essential. Accordingly the SM_2_LAN product is not envisaged to be released. There are thus only two products envisaged for this parameter, at levels 3 and 4. The level 3 product contains along-track filtered and averaged mean soil moisture estimates (SMMEs) for each pass; the level 4 product contains time series for each SSME.

It is noted that Cryosat2 currently overflies all DREAM regions in LRM mode. However, this technique is applicable to both SAR and LRM mode data.

5.13 SM_3_LAN Snow Depth datasets

For SM-3-LAN a variable along-track averaging is applied to create Soil Moisture Mean Estimates (SMMEs) along the satellite track. The extent of this averaging may be different for each desert, but is fixed for all data derived using a single DREAM.

This product is not available at that time. It will be added in a new release of this document.

5.14 SM_4_LAN Snow Depth datasets

Because the soil moisture user community works with time series data at specific locations, the SM_4_LAN product contains time series of each Soil Moisture Mean Estimate (SMME) produced in SM_3_LAN. This product can be generated for Cryosat2 data, but the long repeat period means that this product is of benefit primarily for Sentinel3.

This product is not available at that time. It will be added in a new release of this document.

6 Product availability

All LOTUS data products are made available to the user community. The prototype datasets are hosted in a data center facility at CLS in Toulouse, France, and accessed over a Cloud server service.

To download LOTUS prototype datasets directly from the Cloud shared file service:
https://nas-ext.cls.fr:5001/fbsharing/tM1ATVZq

No authentication is needed to access to this server. However, you may need to authorize your browser entering the website by overriding the settings for this website.

The link is also accessible at the LOTUS's website:
http://www.fp7-lotus.eu/Publications/Prototype-data

Details of the content for each product are provided in the Data Product User Manual (DPUM) D3.2 deliverable.

6.1 Folders on the Cloud Server

Datasets are sorted by zones, date and level of processing, in a recursive directory list (represented like tree branches), making them easy for everyone to access (upon geographic and time criteria). No access restrictions are applied on folders giving you access to the entire LOTUS list of datasets.

An example of the user product tree is given below:

```
├── OpenOcean
│   ├── NE-Atlantic
│   ├── Adriatic-Sea
│   │   ├── 2012
│   │   │   ├── 01
│   │   │   ├── 02
│   │   │   │   ├── L2
│   │   │   │   └── L3
│   │   │   ├── 03
│   │   │   ├── 04
│   │   │   :
│   │   │   ├── 11
│   │   │   ├── 12
│   │   └── 2013
│   └── Singapore-Bay
├── PolarOcean
├── CoastalSea
```

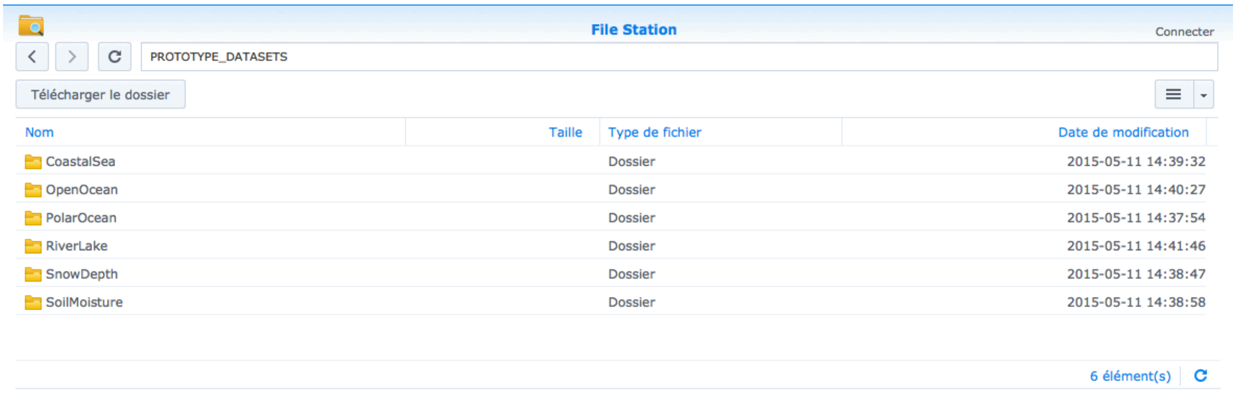
```
├── RiverLake
├── SnowDepth
└── SoilMoisture
```

From the web link, the default folder is /PROTOTYPE_DATASETS. Then the following naming scheme is used throughout the folders hierarchy:

- at the top level is the name of the product (OpenOcean, PolarOcean, CoastalSea, RiverLake, SnowDepth, or SoilMoisture),
- at the second level is the name of the zone (e.g. NE-Atlantic),
- the third and fourth levels are the year and month of the data, and
- last level directory is the data level (L2, L3 or L4).

Dataset files are put in their own fifth level directory, given a unique name (e.g. CS_CPP_OO_2_OCE_Adriatic-Sea_12E20E40N46N_20120506T152940_20120506T153053.nc is stored in the /PROTOTYPE_DATASETS/OpenOcean/Adriatic-Sea/2012/05/L2 directory).

The share file system pages look something like this:



To retrieve data from the Cloud server, right-click your selected file or folder and choose *Download*.

6.2 Product Naming

A naming convention has been applied to LOTUS data products to help users to identify the type of data, but also the time period and the area of interest (see D4.1 and D4.2 deliverables):

The following convention is employed to name all LOTUS data products:
[Product ID]_[Zone]_[Zone Coord]_[First Date/Time Stamp]_[Last Date/Time Stamp].[ext]

Product ID:
The Product ID is based on the short name for each LOTUS data product, for example: CS_CPP_OO_2_OCE. See D3.1 and D3.2 deliverables.
The short name contains the mission identifier ("CS" for Cryosat), the data processing identifier (for example "CPP" or "ESA"), the surface type ("OO" for Open Ocean), the data level and a designator for marine or land product (typically "OCE" for ocean and "LAN" for land).

Zone designator:
The filename contains the name of the selected test area and a geographical coordinate designator that references the southwest corner and the northeast corner of this bounding box.

First Date/Time Stamp:
The UTC date/time stamp of the first data measurement that appears in the product file. Format is: YYYYMMDDThhmmss

Last Date/Time Stamp:
The UTC date/time stamp of the last data measurement that appears in the product file. Format is: YYYYMMDDThhmmss

Extension
The products are in netcdf format; the extension will be .nc

Example file names are given below:
CS_CPP_OO_2_OCE_Adriatic-Sea_12E20E40N46N_20120506T152940_20120506T153053.nc
CS_CPP_CS_2_OCE_Adriatic-Sea_10E22E38N48N_20120506T152940_20120506T153053.nc

Note that the test area coordinates and date time designator provide the major means of identification.

7 News and Updates

To be kept informed on LOTUS activities (paper for a scientific journal or conference presentation) and on any dataset release or eventual dataset re-processing, please refer to the LOTUS's website:
http://www.fp7-lotus.eu/News-and-updates

All information about the prototype dataset release and notable changes made to the released datasets is reported in the latest version of the DPUM document (see the document change log on page 2 that lists the main changes). This document is available on the project's website:
http://www.fp7-lotus.eu/Publications/Deliverables

References

[LOTUS D1.3 report, 2014]: LOTUS report, "*Algorithm Theoretical Baseline Document for SAR mode processing over ocean*", D1.3, Version 1.0, 11st Jun. 2014.

[LOTUS D2.3 report, 2014]: LOTUS report, "*Theoretical Basis Document for river and lake levels algorithms*", D2.3, Version 1.0, 22nd May 2014.

[LOTUS D2.4 report, 2014]: LOTUS report, "*Cryosat2 Soil Surface Moisture Algorithm Theoretical Basis Document*", D2.4, Version 1.1, 18th Jun. 2014.

[LOTUS D2.5 report, 2014]: LOTUS report, "*Snow depth Theoretical Basis Document*", D2.5, Version 1.0, 20th Aug. 2014.

[LOTUS D2.6 report, 2014]: LOTUS report, "*Develop processing for snow depth*", D2.6, Version 1.0, 22nd Aug. 2014.

[LOTUS D3.1 report, 2014]: LOTUS report, "*Data Product Definition Document*", D3.1, Version 4, 3rd Nov. 2014.

[LOTUS D3.3 report, 2015]: LOTUS report, "*Algorithm Theoretical Baseline Document detailing the high level specification of the process*", D3.3, Version 1, 15th Apr. 2015.

[LOTUS D3.4 report, 2015]: LOTUS report, "*SAR Mode for Ocean Corrections Theoretical Basis Document*", D3.4, Version 1, 16th Jul. 2014.

[LOTUS D4.1 report, 2015]: LOTUS report, "*Processed ocean SAR data*", D4.1, Version 1, 4th May. 2015.

[LOTUS D4.2 report, 2015]: LOTUS report, "*Processed land SAR data*", D4.2, Version 1, 4th May. 2015.

[END OF DOCUMENT]