Theia / Hydroweb

Hydroweb Product User Manual

Reference :	THEIA-MU-42-0282-CNES
Issue:	2. 1
Date:	2021,May.31

CLS headquarters
11 rue Hermès
Parc technologique du Canal
31520 Ramonville Saint-Agne
FRANCE

Tel. +33 (0)5 61 39 47 00
Fax +33 (0)5 61 75 10 14
Mail: info@cls.fr
www.cls.fr

CLS Brest Le Ponant
Zone du technopôle de Brest Iroise
Avenue La Pérouse
29280 Plouzané
FRANCE

Tel. 33 (0)2.98.05.76.80
Fax 33 (0)2.98.05.76.90

Chronology Issues:

Issue:	Date:	Reason for change:	Author
1.0	15/01/2016	initial version	V. Rosmorduc
2.0	09/04/2020	new format of river products	M. Vayre
2.1	31/05/2021	Correction of author name in lac reference article	N. Taburet

People involved in this issue:

Written by (*):	V. Rosmorduc, M. Vayre	
Checked by (*):	N. Taburet	[Checker]
Approved by (*):		[Approver]
Application authorized by (*):	Pacholczyk	2021/06/02 Pacholczyk Philippe Signature numérique de Pacholczyk Philippe DN : c=FR, o=CNES, ou=0002 775665912, ou=Agents CNES, 2.5.4.45=PacholczykP@cnes.fr, cn=Pacholczyk Philippe Date : 2021.06.02 12:38:57 +02'00'

In the opposite box: Last and First name of the person + company if different from CLS

Index Sheet:

Context:	
Keywords:	[Mots clés]
Hyperlink:	

Distribution:

Company	Means of distribution	Names
CLS	Notification	

FORM-NT-GB-10-1

List of tables, figures and equations

List of tables:

List of figures:

List of equations:

List of items to be confirmed or to be defined

Lists of TBC:

Aucune entrée de table des matières n'a été trouvée.

Lists of TBD:

Aucune entrée de table des matières n'a été trouvée.

Applicable documents

AD 1 Plan d’assurance produit de CLS
CLS-ED-NT-03-394

Reference documents

RD 1 Manuel du processus Documentation
CLS-DOC

List of Contents

1. Purpose

This document presents the information needed by users for the products provided in the frame of the Hydroweb project.

This document is organized as follows:

- Section 2: Introduction
- Section 3: Satellite radar altimetry general information
- Section 4: Processing
- Section 5: Product description
- Section 6: Web interface
- Section 7: bibliographical references

2. Introduction

Terrestrial waters represent less than 1% of the total amount of water on Earth. However, they have crucial impact on terrestrial life and human needs and play a major role in climate variability. Excluding the ice caps, fresh water on land is stored in various reservoirs: snow pack, glaciers, aquifers and other geological formations, root zone (upper few meters of the soil), and surface waters (rivers, lakes, man-made reservoirs, wetlands and inundated areas). Land waters are continuously exchanged with atmosphere and oceans through vertical and horizontal mass fluxes (evaporation, transpiration of the vegetation, surface and underground runoff). Although improved description of the terrestrial branch of the global water cycle is now recognized as being of major importance for climate research as well as for inventory and management of water resources, the global distribution and spatio-temporal variations of continental waters are still poorly known because routine in situ observations are not available globally. So far, global estimates of spatio-temporal change of land water storage essentially rely on hydrological models, either coupled with atmosphere/ocean global circulation models and/or forced by observations.

Concerning surface waters, in-situ gauging networks have been installed for several decades at least in some hydrographic basins. In situ measurements provide time series of water levels and discharge rates, which are used for studies of regional climate variability as well as for socio-economic applications (e.g., water resources allocation, navigation, land use, hydroelectric energy, flood hazards). Gauging stations, however, are scarce or even absent in parts of large river basins due to geographical, political or economic limitations. Moreover, since the beginning of the 1990s, numerous in-situ networks have declined or stopped working, because of political and economic factors.

Recently, remote sensing techniques have been used to monitor components of the water balance of large river basins on time scales ranging from months to decades. Among these, two are particularly promising: satellite altimetry for systematic monitoring of water levels of large rivers, lakes and floodplains and the space GRACE gravity mission for measurement of spatio-temporal variations of land water storage. Other remote sensing techniques, such as Synthetic Aperture Radar (SAR) Interferometry and passive and active microwave observations also offer important information on land surface waters, such as changing areal extent of large wetlands. By complementing in situ observations and hydrological modelling, space observations have the potential to improve significantly our understanding of hydrological processes at work in large river basins and their influence on climate variability, geodynamics and socio-economic life. Unprecedented information can be expected by combining models and surface observations with observations from space, which offer global geographical coverage, good spatio-temporal sampling, continuous monitoring with time, and capability of measuring water mass change occurring at or below the surface.

3. Satellite Radar Altimetry

3.1. Vocabulary

“Orbit” is one revolution around the Earth by the satellite.

A satellite **“Pass”** is half a revolution of the Earth by the satellite from one extreme latitude to the opposite extreme latitude.

“Repeat Cycle” is the time period that elapses until the satellite flies over the same location again.

“Range” is the satellite-to-surface pseudo distance given from the 2-way travel time of the radar pulse from the satellite to the reflecting water body

“Retracking” is the algorithm which computes the altimetric parameters (range, backscatter coefficient...) from the radar echoes

“Altitude” is considered as the height over the reference ellipsoid (orthometric height).

“Transect” high frequency measurement over a geographical area (river or lake) for a given satellite track

3.2. General principle

Radar altimetry from space consists of vertical range measurements between the satellite and water level (see Figure 1). Difference between the satellite altitude above a reference surface (usually a conventional ellipsoid), determined through precise orbit computation, and satellite-water surface distance, provides measurements of water level above the reference. Placed onto a repeat orbit, the altimeter satellite overflies a given region at regular time intervals (called the orbital cycle), during which a complete coverage of the Earth is performed.

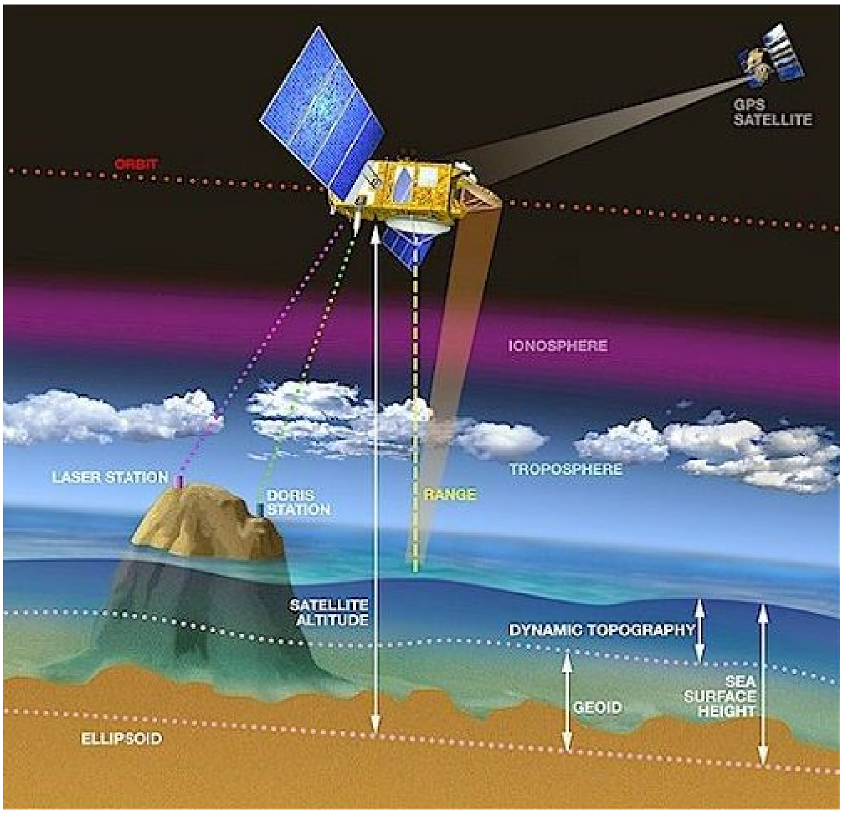

Figure 1: Altimetry principle (over ocean)

Water level measurement by satellite altimetry has been developed and optimized for open oceans. Over large water surface areas (same order than the footprints) so-called “Brown” echoes are obtained (one peak followed by a slow decrease). This is the classic waveform over ocean.

Over small lakes or rivers, the signal is reflected back from a mixture of surface such as open water, floodplains, sandbanks, forest, ect. It affects the shape of the echo adding multiple peaks. The smaller the ratio between water and emerged land surfaces in the footprint, the more complex the waveforms. In these areas the signal is thus more complex which makes the range estimation difficult. Two altimeter technologies are currently used:

- Low Resolution Mode (LRM) for the Jason-3
- Synthetic Aperture Radar (SAR) for Sentinel-3A/3B

The footprints are about 290 km² in LRM and 5 km² in SAR mode (the footprint is divided into “slices”). Figure 2 shows echoes in SAR or LRM modes according to the respective position of the satellite trajectory and the watercourse.

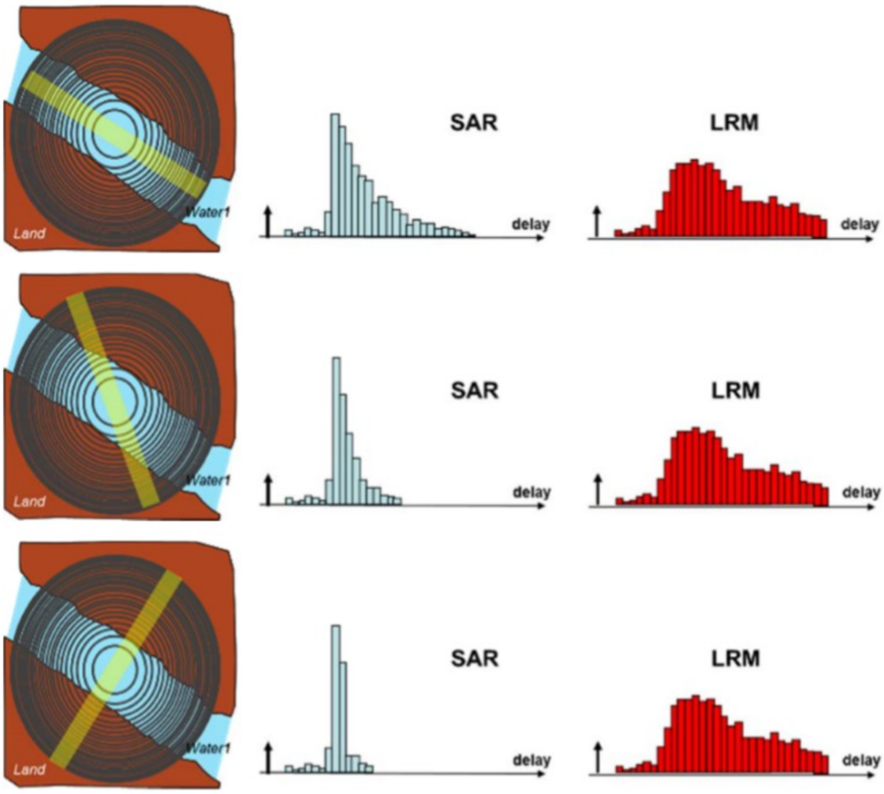

Figure 2: Echoes in SAR or LRM modes according to the respective position of the satellite trajectory (perpendicular to the rectangles of the SAR mode) and the watercourse. Footprints are represented as orange rectangles (SAR footprints) and black concentric circles (LRM footprints). (Credit: Aviso - CNES)

Several algorithms (ice1, ice3, ocean …) have been developed considering these different types of waveform to produce an accurate measurement of the surface height. This reprocessing step is called “retracking”.

Several satellite altimetry missions have been launched since the early 1990s, see Figure 3. The technique is now applied to obtain water levels of inland seas, lakes, rivers, floodplains and wetlands. Table 1 shows the altimetric missions used to perform such computation.

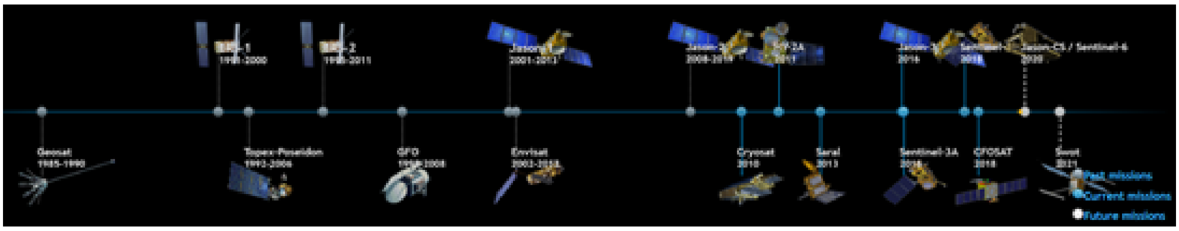

Figure 3: Historical timeline of the altimetric missions (Credit: Aviso)

	Date of launch	End date	Altitude	Inclination	Repetitivity	Agency	Objectives
Topex/Poseidon	10/08/1992	18/01/2006	1336 km	66°	9.9156 days	NASA - CNES	Measure sea surface height
GFO	10/02/1998	26/11/2008	800 km	108°	17 days	US Navy / NOAA	Measure ocean topography
Jason-1	07/12/2001	01/07/2013	1336 km	66 °	9.9156 days	NASA - CNES	Measure sea surface height
Envisat	01/03/2002	08/06/2012	782.4-799.8 km	98.55 °	30-35 days	ESA	Observe Earth's atmoshere and surface
Jason-2	20/06/2008	10/10/2019	1336 km	66 °	9.9156 days	NASA - CNES - EUMETSAT - NOAA	Measure sea surface height
Cryosat-2	08/04/2010		717 km	92 °	369 days with 30 day sub-cycle	ESA	Polar observation
SARAL	25/02/2013		800 km	98.55 °	35 days	ISRO - CNES	Oceans observation
Jason-3	17/01/2016		1336 km	66 °	9.9156 days	NASA -CNES - EUMETSAT - NOAA	Measure sea surface height
Sentinel-3A	16/02/2016		841.5 km	98.65 °	27 days	ESA	Deliver routine operation services to policy-makers and marine and land service users
Sentinel-3B	25/04/2018		841.5 km	98.65 °	27 days	ESA	Deliver routine operation services to policy-makers and marine and land service users

Table 1: Altimetric mission used in the water level computation over rivers and lakes. Current missions are specified with shaded “end date” cells

The combined global altimetry data set has more than 20 year-long history and is intended to be continuously updated in the coming decade. Combining altimetry data from several in-orbit altimetry missions increases the space-time resolution of the sensed hydrological variables.

3.3. Water surface height computation

For all satellites, the following operation is done:

$$corrected\ range = range + wet\ tropo + dry\ tropo + iono + polar\ tide + solid\ earth\ tide$$

range: *Range values*
wet tropo : correction to account for the radar signal delay
Due to cloud liquid water and water vapour in the atmosphere
dry tropo: correction to account for the radar signal delay
Due to the dry gas component of the atmosphere
iono: correction to account for the radar signal delay due to the electron content of the ionosphere
polar tide and ***solid earth tide***: *T*emporal deformation of the Earth's surface

Equation 1: Corrected range value

$$surface\ height = altitude - corrected\ range$$

Equation 2: Surface height value

All these corrections are provided within the input data (L2 products). Finally, the water surface height is expressed with respect to the geoid:

$$water\ surface\ height = surface\ height - geoid + product\ bias$$

Equation 3: Equation 2: Water surface height value

The **geoid** value is computed from a geoid model. For large lakes is calculated from "mean profiles". Those profiles are computed smoothing the geoid profile along the transect to reduce the WSH dispersion. After several cycles (at least 10 for a given ground track) geoid mean profiles are built. Geoid mean profiles are computed once from a given number of cycles (at least 10 cycles) and then used at each cycle for each satellite track. There is one mean profile by lake for each satellite ground track (whatever the cycle).

The "**product bias**" refers to constant values read for each lakes or VS from the master files:

- The lake product is a multi-satellite product. Data come from multiple satellite tracks and missions. Each satellite track samples a different area of the lake. An intercalibration step is therefore necessary. Data are processed independently by satellite ground track over each lake and unbiased with least square approach on the absolute reference TOPEX/Poseidon. Absolute biases are given by tracks (one bias value by lake and satellite track). These inter-tracks biases are written in the master file and used to compute a new lake WSH measurement.
- Dealing with VS, this product bias contains the necessary height correction to reference all VS dataset with a same ellipsoid. The reference ellipsoid is WGS84. A correction of about 70 cm is thus applied for the Jason-2 and Jason-3 measurements (referenced with the Topex/Poseidon ellipsoid).

4. Processing

4.1. Lake processing

A given lake can be overflown by several satellites, with potentially several passes. The lake product is thus a multi-satellite product.

The water level and volume time series is operationally updated less than 3 days after the availability of the input altimetry data for operational lakes. Other lakes are also monitored on a research mode basis.

4.1.1. Method

Past lake levels are based on merged Topex/Poseidon, Jason-1/2, Envisat, SARAL and GFO data provided by ESA, NASA and CNES data centers (see Table 1). Updates include Jason-3 and Sentinel-3A&B data. Processing chain input data are:

- Altimetry GDR and IGDR data (IGDR for near-real time processing), with the included corrections
- Along-track mean profiles over the lakes
- Location information, including missions and track number for all the processed lakes
- Maximum variation of the lake water level (for a given time step). These threshold values are included in a file (one value per lake)
- Hypsometric curves built using satellite imagery

The processing takes place according to an identical sequence of tasks. It consists of 3 main stages:

- Extract / select / read measurements from each of the IGDR files (Short Time Critical files) and calculate the water levels along the track.
- Water level calculation and filtering
- Validation and database update

These steps are further discussed for a given transect (GDR high frequency measurements over one lake at a given mission cycle).

4.1.2. First step: extraction of input data

Whatever the mission and whatever the lake, this step is unique. The altimetry database is explored for the latest measurements since the last update. Current lakes dataset and their processing specifications are indicated in the master file (retracking, use of 1Hz or high frequency measurements …).

All measurements and altimetric corrections are read from the L2 input files. The “ocean” and “ice1” retracking models are used for large and small lakes respectively.

Current along track data correspond to satellite passes over one of the “operational” lakes. Rejection criteria (see Table 2), based on altimetric corrections, are applied on along track measurements. The output corresponds to the measurements used in the water level computation.

Altimetric correction	Boundary values for editing (valid in between) [m]
Dry tropospheric correction	-2.5 / -1.2
Wet tropospheric correction	-0.8 / 0.01
Ionospheric correction	-0.4 / 0.004
Sigma0	7 dB / 40 dB for sigma0 ice1 (Jason-3)
	22 dB / 55 dB for sigma0 OCOG (S3A&B)
	7 dB / 40 dB for sigma0 Ocean (S3A&B)

Table 2: Boundary values of the different altimetry corrections

4.1.3. Second step: lake level computation and filtering

Once the measurements selection made, the water surface height computation can be done (see section 3.3). Then, the along track outliers are removed. This is performed by the following computation (recursive processing):

- Each height measurement is compared with the current lake median level. High frequency point is not selected if its deviation from the median measurement is greater than 1.5 times the value of the transect standard deviation. A new median value and standard deviation are then calculated. Other iterations are proceeded similarly but with a removal criterion equivalents to twice the standard deviation. Four iterations are carried out unless no measurement is rejected in one of them.
- If the standard deviation of the current selected measurements is upper than 2 meters the current median value is not used to update the lake timeseries

4.1.4. Third step: validation, volume computation and database update

At this stage one water level median value is available for the current transect. For some lakes, a maximum variation threshold is provided by the LEGOS hydrological experts. It is based on the history of the lake/reservoir level variations.

The variation between the current measurement and the last available in the lakes timeseries is computed. This variation is compared to the maximum variation thresholds (Equation 4).

$$\Delta_{WL} = \frac{WL_{t_{n+1}} - WL_{t_{last}}}{t_{n+1} - t_{last}}$$

$$\Delta_{WL} < \Delta_{\mathrm{WL\,max}} * 1.40$$

$\boldsymbol{WL}$: Lake water level
$\boldsymbol{\Delta_{WL}}$: lake water level variation
$\boldsymbol{\Delta_{WL\,max}}$: Threshold value of the lake water level variation (constant value, read from an input file)
$\boldsymbol{t_{n+1}}$: Current measurement date
$\mathbf{t_{last}}$: Last available measurement in the lake timeseries

Equation 4: Selection criterion from lake water level variation

The level variation between the most recent estimate and the last validated estimate is compared to the level variation threshold value. A weighting factor of 1.4 is applied to the maximum value ($\Delta_{\mathrm{WL\,max}}$) of the variation. This 40 % of the maximum value is thus applied to be conservative. Exceeding this value, the new estimate is not validated. Otherwise, the current water level will be used to update the lake water level timeseries.

The current water level value is now validated. For some lakes, a mathematical relation between the lake water level and the lake surface is available. Indeed, an hypsometric file contains all these relations. This file is provided by the LEGOS from satellite imagery data. The current lake surface is thus computed (Equation 5) if such data are available (for some lakes only the height will be computed). Volumes variations are then computed, see

$\boldsymbol{\Delta_{WL}}$: Water level variation between current and initial measurement
$\boldsymbol{S_{current}}$: Lake current area
$\boldsymbol{S_{init}}$: Lake initial area

Equation 6.

FORM-NT-GB-7-1

$$lake\ surface = poly(lake\ water\ lavel)$$

$\boldsymbol{poly}$: Mathematical relation between water lave and surface area

Equation 5: Lake surface from lake water level measurement

$$lake\ volume = \frac{\Delta_{WL} * \left(S_{current} + S_{init} + \sqrt{S_{current} * S_{init}}\right)}{3}$$

$\boldsymbol{\Delta_{WL}}$: Water level variation between current and initial measurement
$\boldsymbol{S_{current}}$: Lake current area
$\boldsymbol{S_{init}}$: Lake initial area

Equation 6: Lake volume from lake water level measurement and surface area

Finally, the lake water level is updated as well as the surface and volume timeseries for some lakes.

4.2. River processing

Radar echoes over land surfaces are hampered by interfering reflections due to water, vegetation and topography. Consequently, waveforms (e.g., the power distribution of the radar echo within the range window) may not have the simple broad peaked shape seen over ocean surfaces, but can be complex, multi-peaked, preventing from precise determination of the altimetric height. If the surface is flat, problems may arise from interference between the vegetation canopy and water from wetlands, floodplains, tributaries and main river.

The time series available in Hydroweb are constructed from “historical” altimetry data using Jason-2 and Envisat. Jason-3, Sentinel-3A and Sentinel-3B data are currently used to update the virtual stations dataset.

A virtual station (see Figure 4) is defined as the intersection between a satellite theoretical track and a river centerlines (representing the riverbed). From this central point, a geographical area is then defined considering the satellite “drift” (less than 1km across track around the reference ground track).

For a given satellite track, high frequency measurements within the virtual station geographical area are edited. One median value and its associated uncertainty (standard deviation) is thus obtained at each transect.

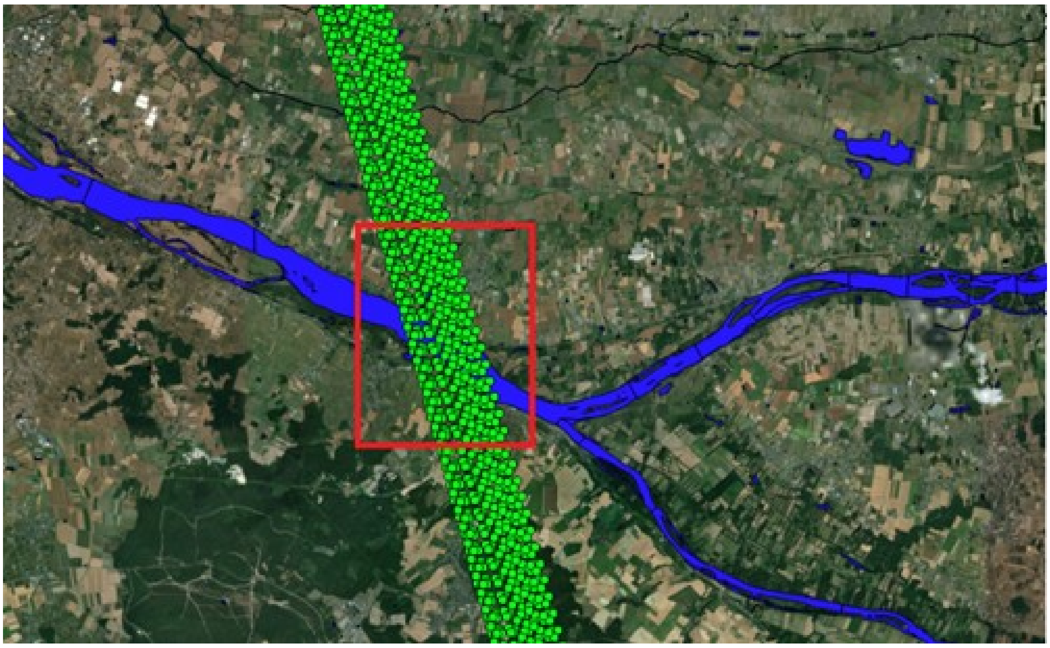

Figure 4: Virtual station concept. Example over the Loire river with Sentinel-3A. 20 Hz measurements are represented in green points (several cycles for a same pass number). An example of virtual station geographical area is shown as a red rectangle

Some stations are on a research mode basis. They are not updated from current altimetric data. Other virtual stations are operational i.e. operationally updated less than 3 days after the availability of the input altimetry data (Jason-3, Sentinel-3A and Sentinel-3B data).

Only repetitive orbits are used (no Jason-2 data when the satellite was put on a drifting orbit for instance).

4.2.1. Method

Past rivers levels are based on Jason-2 and Envisat data. Updates include Jason-3, Sentinel-3A and Sentinel-3B data. Processing chain input data are:

- Altimetry GDR and IGDR data (IGDR for near-real time processing), with the included corrections
- Virtual stations master file: virtual station name, mission, pass number, bias, threshold values of the processing and editing step
- Geographical area of the virtual station (GeoJSON file)
- the geoid grid (EGM2008), used to calculate the height of the geoid at the location of each Virtual Station. The grid resolution is 1x1 minute, and is downloaded from http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/egm08_wgs84.html and converted into NetCDF by the LEGOS
- Centerlines (CNES/LEGOS/CLS product), used to compute intersections with the processed GRD product. Once such coordinates available, the current distance from the river mouth is computed

The processing takes place according to an identical sequence of tasks. It consists of 3 main stages:

- Data reading and selection
- Water level calculation and filtering
- Validation and database update

These steps are further discussed for a given transect (GDR high frequency measurements over one virtual station at a given mission cycle).

4.2.2. First step: Data reading and selection

The current altimetric data (either IGDR or GDR) are read. The dataset of the virtual stations to update is obtained from the “master file”. For each virtual station (product of interest for the current GRD file) the current high frequency measurements are extracted over its geographical area.

Then the backscatter coefficient is used to select the most echogenic measurements. A threshold value is indicated for each virtual station in the master file. Then, a dynamic threshold is used to select the most consistent data over such water bodies (measurements selected as the higher backscatter coefficient values).

Water surface heights are then computed (see section 3.3).

4.2.3. Second step: Data filtering

At this step, high frequency measurements of the water surface heights are available for the current transect.

Overall WSH average and **Overall** standard deviations average are computed from each virtual station timeseries. They represent the WSH average and its associated standard deviation for a given virtual station.

For the current transect, the **current** WSH average value and its standard deviation are computed.

Moving average are computed from the virtual station timeseries and the current measurements. All WSH values for a given virtual station are merged on a “standard year”. It means that these values are only referenced by their annual day number (1st January equals to 1 and 1st February equals to 32 for instance). The annual day number corresponding to the current measurements is then computed. Moving average is computed over +/- 3 cycles from the current annual day number. New WSH average value and standard deviation are thus available.

The filtering is performed by the following computation (recursive processing):

- Measurements elimination (current GRD product) with respect to the **overall** average and **overall** standard variation, see Equation 7.

$$abs\left(WSH_{average} - WSH_{current}\right) < 3\ \sigma$$

$\boldsymbol{WSH_{average}}$: Overal l$water\ surface\ heights\ average$
$\boldsymbol{WSH_{current}}$: $Water\ surface\ heights\ measurements\ (high\ frequency)$
$\boldsymbol{abs}$: $Absolute\ value$
$\boldsymbol{\sigma}$: $Global\ water\ surface\ heights\ standard\ deviations\ average$

Equation 7: WSH filtering against global mean (virtual station)

- Elimination based on the **current** average and its standard deviation. A similar iteration (new water surface heights average and standard deviation are computed) is then done using a more restrictive value, see Equation 8.

$$abs\left(WSH_{average} - WSH_{current}\right) < 3\ \sigma\ \ \&\ \ abs\left(WSH_{average} - WSH_{current}\right) < 2\ \sigma$$

$\boldsymbol{WSH_{average}}$: $Current\ water\ surface\ heights\ average$
$\boldsymbol{WSH_{current}}$: $Water\ surface\ heights\ measurements\ (high\ frequency)$
$\boldsymbol{abs}$: $Absolute\ value$
$\boldsymbol{\sigma}$: $Current\ water\ surface\ heights\ standard\ deviations\ average$

Equation 8: WSH filtering against current mean (virtual station)

- Elimination from **moving average**: the corresponding average value and standard deviation can be used as a river template. A threshold coefficient (x_{th}) is indicated in the master file for each station to set the impact of this filtering, see Equation 9.

$$abs\left(WSH_{average} - WSH_{current}\right) < x_{th}\ \sigma$$

$\boldsymbol{WSH_{average}}$: *Global water surface heights average*
$\boldsymbol{WSH_{current}}$: *Water surface heights measurements (high frequency)*
$\boldsymbol{abs}$: *Absolute value*
$\boldsymbol{x_{th}}$: threshold coefficient, indicated in the master file
$\boldsymbol{\sigma}$: *Global water surface heights standard deviations average*

Equation 9: WSH filtering against moving average (virtual station)

- Standard deviation of the current selected measurements must be lower than 5 m. Otherwise this transect is not used to update the virtual station timeseries

Current water surface heights (from GDR product) are thus extracted within a geographical area and selected (from backscatter coefficients and water surface heights filtering). A median value is computed to update the virtual station time. Other variables are also read or calculated (non-exhaustive list):

- Current intersection between the satellite track and a river centrelines product. Mean location is computed if this intersection cannot be calculated
- Current geoid value (at the current intersection point location)
- Distance from the river mouth (at the current intersection point location)
- Several current transect constants are read: mission name, cycle, pass number, retracking used, GDR IPF …

4.2.4. Third step: Output data

The final stage consists in updating the virtual station database from the current measurement. Output files (CSV files) are also written. Some comments in the header section are thus updated. One can mention the first/last date in the dataset, the production date, minimal/maximal distance from the river mouth, the mean altitude and the number of measurements.

5. Products description

5.1. Description

Dataset Name	Production frequency	Temporal resolution	File format
Lake level, surface & volume variation time series (operational)	every day, depending on the tracks	depending on the altimetry mission coverage	".txt" with metadata and header as comment (one file per lake)
River level time series (operational)	every day, depending on the tracks	depending on the altimetry mission coverage	".csv" with metadata and header as comment (one file per virtual station)

Lake level, surface & volume variation time series (research)	once and for all	depending on the altimetry mission coverage	".txt" with metadata and header as comment (one file per virtual station)
River level time series (research)	once and for all	depending on the altimetry mission coverage	".csv" with metadata and header as comment (one file per virtual station)

Table 3: Products description

5.2. Lake products

5.2.1. Nomenclature

File are named:

- L_<*lakename*>.txt, <*lakename*>*: lake name (*in English or French)

5.2.2. Content

The lakes files include:

- a first line with descriptive metadata
- a header in a comment format (series of lines with #)
- the data themselves

5.2.2.1. Product first line

The product first line is:

lake=<*lakename*>;country=<*Countryname1*>&<*Countryname2*>;basin=<*basinname*>;lat=<*lat*>;lon=<*lon*>;date=<*yyyy/mm/dd*>;first_date=<*yyyy/mm/dd*>;last_date=<*yyyy/mm/dd*>;type=*(operational or research)*;diff=public

- <*lakename*> is the name of the lake (full name), in local language wherever possible (a choice might be done if the lake shore several different countries)
- <*CountrynameN*> are the names (full name) of the countries where the lake is located (several names possible, separated by &)
- <*basinname*> is the hydrological basin name (full name), in English
- <*lat*> latitude of the lake
- <*lon*> longitude of the lake
- "Date" is the processing date when the data have been computed
- "first_date" is the date of the first available measurement
- "last_date" is the date of the last available measurement
- type=*(operational or research)* operational data are produced daily. At a given location, data will be updated if measurements are available (satellite track over the area), not updated if no measurement over the area or all measurements filtered; research data are provided by Legos.

5.2.2.2. Header

Header lines begin with a “#”. The header includes several useful information:

- Geographical information on the lake (size, depth...)
- Missions and track number of each mission with measurements over this lake
- Corrections applied
- Reference surface used, with citation
- First date of measurement in the following format:
  <*YYYY*> <*MM*> <*dd*> yr month day <*HH*> hours <*MIN*> minutes
- Last date of measurement in the following format:
  <*YYYY*> <*MM*> <*dd*> yr month day <*HH*> hours <*MIN*> minutes
- The data formatting, with column number, content and unit or format if relevant: # (1): data1 name (unit), (2): data2 name = format, (3) data3 name
- Source and credits information, citation

5.2.2.3. Data

Each line corresponds to a different measurement.

For a given measurement (i.e. a line), associated data are separated by a “;”. Some fields can be empty. The value “9999.999” is used to indicate unavailable data.

	column number	**format**
decimal year	1	yyyy.ddddd
date	2	yyyy/mm/dd
time	3	hh.mm
height above reference surface	4	float
height standard deviation (mad)	5	float
surface	6	float
volume	7	float
flag	8	Currently empty

Table 4: Lake data

5.3. River products

5.3.1. Nomenclature

File are named:

R_<BASINNAME>_<RIVERNAME>_<KMXXXX>_.csv

Where:

- <BASINNAME>: Basin name
- <RIVERNAME>: River name
- <KMXXXX>: Mean distance from river mouth (curvilinear abscissa)

5.3.2. Content

The VS files include:

- a header in a comment format (series of lines with #)
- the data themselves

5.3.3. Header

Each header line begins with a "#" character. The river product header is composed as follows (divided hereafter in tree parts):

- Hydrological metadata:
  - #BASIN: Basin Name
  - #RIVER: River name
  - #ID: Virtual station ID (unique identifier)
  - #TRIBUTARY OF: Upstream river name (could be unavailable, equals to "NA" in this case)
  - #APPROX. WIDTH OF REACH (m): River width estimation (could be unavailable, equals to "NA" in this case)
  - #SURFACE OF UPSTREAM WATERSHED (km2): Estimation of the upstream watershed surface (could be unavailable, equals to "NA" in this case)
  - #RATING CURVE PARAMETERS A,b,Zo such that Q(m3/s) = A[H(m)-Zo]^b:: Rating curve parameters (could be unavailable, equals to "NA NA NA" in this case)

- Geographical metadata:
  - #REFERENCE ELLIPSOID: Ellipsoid of reference
  - #REFERENCE LONGITUDE: Virtual station longitude
  - #REFERENCE LATITUDE: Virtual station latitude
  - #REFERENCE DISTANCE (km): Distance from river mouth (curvilinear abscissa)
  - #GEOID MODEL: Reference geoid (geoid model version)
  - #GEOID ONDULATION AT REF POSITION(M.mm): Geoid value at the virtual station location
  - #MISSION(S)-TRACK(S): List of the used mission(s)-track(s) to build the timeseries
  - #STATUS: Operational (daily processed) or research
  - #VALIDATION CRITERIA: "Expert", "In-situ" or "Automatic" (virtual station validated from hydrological expert, in-situ data or automatically from statistical criteria)
  - #MEAN ALTITUDE (M.mm): Mean height
  - #MEAN SLOPE (mm/km): Mean slope over the river section (could be unavailable, equals to "NA" in this case)
- Product metadata:
  - #NUMBER OF MEASUREMENTS IN DATASET: Measurements number
  - #FIRST DATE IN DATASET: First date in the timeseries
  - #LAST DATE IN DATASET: last date in the timeseries
  - #DISTANCE MIN IN DATASET (km): Minimal distance from river mouth
  - #DISTANCE MAX IN DATASET (km): Maximal distance from river mouth
  - #PRODUCTION DATE: Production date
  - #PRODUCT VERSION: Product version
  - #PRODUCT CITATION: Product citation
  - #SOURCES: Product sources
  - #PRODUCT CONTENT: Description of the data content

5.3.4. Data

Each line corresponds to a different measurement. For a given measurement each column is separated by a space character. Some fields can be empty. The values “9999.999” or “NA” are used to indicate unavailable data.

Measurement data are provided as follows:

- “basic data”:
    - #COL 1: DATE(YYYY-MM-DD)
    - #COL 2: TIME(HH:MM)
    - #COL 3: ORTHOMETRIC HEIGHT (M) OF WATER SURFACE AT REFERENCE POSITION
    - #COL 4: ASSOCIATED UNCERTAINTY(M)
- “Expert data”:
    - #COL 5: LONGITUDE OF ALTIMETRY MEASUREMENT (in degree)
    - #COL 6: LATITUDE OF ALTIMETRY MEASUREMENT (in degree)
    - #COL 7: ELLIPSOIDAL HEIGHT OF ALTIMETRY MEASUREMENT (meters)
    - #COL 8: GEOIDAL ONDULATION (meters) at location [5,6]
    - #COL 9: DISTANCE OF ALTIMETRY MEASUREMENT TO REFERENCE POSITION (km)
    - #COL 10: SATELLITE
    - #COL 11: ORBIT / MISSION
    - #COL 12: GROUND-TRACK NUMBER
    - #COL 13: CYCLE NUMBER
    - #COL 14: RETRACKING ALGORITHM
    - #COL 15: GDR VERSION

On the Hydroweb portal, users can either download de “basic” files that contains only the metadata and first four columns of data, or the “expert” files containing all data columns.

5.4. Acknowledgments

When using those products, please cite:

- Research products: “Timeseries specified by LEGOS and computed by CLS on behalf of CNES”
- Operational products: “Timeseries specified by LEGOS and computed by CLS on behalf of CNES and Copernicus Global Land”

6. Web interface

6.1. Access service

The web site http://hydroweb.theia-land.fr/ is the point of access for all Hydroweb data.

6.2. Registration

To access the data, you must register. A link is provided top right of the website (sign up/s’inscrire).

Registration is common to all Theia products.

6.3. User guide and additional services

User guide section is available in the web site. The use of additional services is explained:

- product quicklook: Service available without authentication. It allows timeseries figures download.
- Authenticated data access: Automatic products download from a given URL
- CSW service: Provide products metadata without authentication
- SOS service: Provide products metadata without authentication

6.4. Visualisation tools

Some visualisation tools are available: http://hydroweb.theia-land.fr/hydroweb/tools?lang=en

KML and Shapefile, representing the locations of the available virtual stations and lakes, can be downloaded. These files can be used in all regular GIS softwares (Google Earth, QGIS …).

6.5. Contact

If you have any questions about Hydroweb, please contact: contact.hydroweb@cls.fr

7. References

7.1. Lakes

SOLS: A Lake database to monitor in Near Real Time water level and storage variations from remote sensing data, J. Adv. Space Res. Vol 47, 9, 1497-1507, doi:10.1016/j.asr.2011.01.004, 2011 Crétaux J-F , A. Arsen , S. Calmant , A. Kouraev , V. Vuglinski , M. Bergé Nguyen , M-C. Gennero, F. Nino, R. Abarca Del Rio , A. Cazenave , P. Maisongrande

Lake studies from satellite altimetry, C R Geoscience, Vol 338, 14-15, 1098-1112, doi: 10.1016/J.cre.2006.08.002, 2006 Crétaux J-F and C. Birkett

7.2. Rivers

EGM2008 Citation: The development and evaluation of the Earth Gravitational Model 2008 (EGM2008) - Nikolaos K. Pavlis, Simon A. Holmes, Steve C. Kenyon, John K. Factor; Journal of Geophysical Research: Solid Earth (1978-2012) Volume 117, Issue B4, April 2012

Frappart F., Calmant S., Cauhopé M., Seyler F., Cazenave A. (2006). Preliminary results of ENVISAT RA-2 derived water levels validation over the Amazon basin, Remote Sensing of Environment, 100(2), 252-264, doi:10.1016/j.rse.2005.10.027.

Santos da Silva J., S. Calmant, O. Rotuono Filho, F. Seyler, G. Cochonneau, E. Roux, J. W. Mansour, Water Levels in the Amazon basin derived from the ERS-2 and ENVISAT Radar Altimetry Missions, *Remote Sensing of the Environment*, doi:10.1016/j.rse.2010.04.020, 2010

Citation for data on specific watersheds:

Congo Basin:

Becker, M., J. Santos da Silva, S. Calmant, V. Robinet, L. Linguet et F. Seyler, Water level fluctuations in the Congo Basin Derived from ENVISAT satellite altimetry, *Remote Sensing*, 6, doi:10.3390/rs60x000x, 2014

Gange basin

FORM-NT-GB-7-1

Pandrey, R. K., J-F Crétaux, M Bergé-Nguyen, V. Mani Tiwari, V. Drolon, F. Papa and S. Calmant, Water level estimation by remote sensing for 2008 Flooding of the Kosi River, *IJRS*, 35:2, 424-440, doi:10.1080/01431161.2013.870678, 2014

Papa, F., F. Frappart, Y. Malbeteau, M. Shamsudduha, V. Venugopal, M. Sekhar, G. Ramillien, C. Prigent, P. Aires, R. K. Pandey, S. Bala, S. Calmant, Satellite-derived Surface and Sub-surface Storage in the Ganges-Brahmaputra River Basin, *Journal of Hydrology, regional studies*, doi:10,1016/j.ejrh,2015,03,004 2014.

Orinoco :

Frappart, F., F. Papa, Y. Malbeteau, J-G Leon, G. Ramillien, C. Prigent, L. Seoane, F. Seyler, S. Calmant, Surface freshwater storage variations in the Orinoco floodplains using multi-satellite observations, *Remote Sensing*, 7,89-110, doi:10,3390/rs70100089, 2015.

Other basins :

Santos da Silva, J., F. Seyler, S. Calmant, O. Correa Rotunno Filho, G. Cochonneau et W.J. Mansur, Water level dynamics of Amazon wetlands at the watershed scale by satellite altimetry, *Int. J. Of Remote Sensing*, 33:11, 3323-3353, doi 10.1080/01431161.2010.531914, 2012

Citation for empirical mean altimetric biases over rivers :

Calmant, S., J. Santos da Silva, D. Medeiros Moreira, F. Seyler, C.K. Shum, J-F. Crétaux, G. Gabalda, Detection of Envisat RA2 / ICE-1 retracked Radar Altimetry Bias Over the Amazon Basin Rivers using GPS, *Advances in Space Research*, doi: 10.1016/j.asr.2012.07.033, 2012.

Seyler, F., S. Calmant, J. Santos da Silva, D. Medeiros Moreira, F. Mercier, CK Shum, From TOPEX/Poseidon to Jason-2/OSTM in the Amazon basin, *Advances in Space Research*, doi: 10.1016/j.asr.2012.11.002, 2012

List of acronyms

AD	Applicable Document
RD	Reference Document
TBC	To be confirmed
TBD	To be defined
VS	Virtual station
WSH	Water Surface Height