remote sensing

MDPI

Article

Challenges and Evolution of Water Level Monitoring towards a Comprehensive, World-Scale Coverage with Remote Sensing

Mélissande Machefer [1,*], Martí Perpinyà-Vallès [1], Maria José Escorihuela [2], David Gustafsson [3] and Laia Romero [1]

[1] Lobelia Earth S.L., Marie Curie, 8-14, A226, 08042 Barcelona, Spain; marti@lobelia.earth (M.P.-V.); laia@lobelia.earth (L.R.)
[2] isardSAT, Marie Curie, 8-14, A219, 08042 Barcelona, Spain; mj.escorihuela@isardsat.cat
[3] Sveriges Meteorologiska och Hydrologiska Institut (SMHI), Folkborgsvägen 17, SE-601 76 Norrköping, Sweden; david.gustafsson@smhi.se
* Correspondence: melissande@lobelia.earth

Abstract: Surface water availability is a fundamental environmental variable to implement effective climate adaptation and mitigation plans, as expressed by scientific, financial and political stakeholders. Recently published requirements urge the need for homogenised access to long historical records at a global scale, together with the standardised characterisation of the accuracy of observations. While satellite altimeters offer world coverage measurements, existing initiatives and online platforms provide derived water level data. However, these are sparse, particularly in complex topographies. This study introduces a new methodology in two steps (1) *teroVIR*, a virtual station extractor for a more comprehensive global and automatic monitoring of water bodies, and (2) *teroWAT*, a multi-mission, interoperable water level processor, for handling all terrain types. L2 and L1 altimetry products are used, with state-of-the-art retracker algorithms in the methodology. The work presents a benchmark between *teroVIR* and current platforms in West Africa, Kazakhstan and the Arctic: *teroVIR* shows an unprecedented increase from 55% to 99% in spatial coverage. A large-scale validation of *teroWAT* results in an average of unbiased root mean square error ubRMSE of 0.638 m on average for 36 locations in West Africa. Traditional metrics (ubRMSE, median, absolute deviation, Pearson coefficient) disclose significantly better values for *teroWAT* when compared with existing platforms, of the order of 8 cm and 5% improved respectively in error and correlation. *teroWAT* shows unprecedented excellent results in the Arctic, using an L1 products-based algorithm instead of L2, reducing the error by almost 4 m on average. To further compare *teroWAT* with existing methods, a new scoring option, *teroSCO*, is presented, measuring the quality of the validation of time series transversally and objectively across different strategies. Finally, *teroVIR* and *teroWAT* are implemented as platform-agnostic modules and used by flood forecasting and river discharge methods as relevant examples. A review of various applications for miscellaneous end-users is given, tackling the educational challenge raised by the community.

Keywords: remote sensing; satellite; altimetry; water level; water inland; essential climate variable; database

check for updates

Citation: Machefer, M.; Perpinyà-Vallès, M.; Escorihuela, M.J.; Gustafsson, D.; Romero, L. Challenges and Evolution of Water Level Monitoring towards a Comprehensive, World-Scale Coverage with Remote Sensing. *Remote Sens.* **2022**, *14*, 3513. https://doi.org/10.3390/rs14153513

Academic Editor: Konstantinos X. Soulis

Received: 7 June 2022
Accepted: 14 July 2022
Published: 22 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Due to their strong contribution to the water cycle, water surface dynamics are the predictive features for the eventuality of floods, hydric stress, water scarcity and severe drought [1–4]. Soaring needs of monitoring are being expressed by all types of end-users, such as hydrological and climate experts, administrations, intergovernmental and non-governmental agencies and financial investors. Through the requirements expressed by communities committed to water preservation, challenges of getting information in specific uncovered areas are being witnessed [5]. Significant indicators encompass water levels, water discharge and water volume changes of inland reservoirs and rivers [6–12].

Remote sensing, and especially the field of altimetry, offers a privileged access to the quantification of water availability and the computation of the relative change in terms of water resources [13].

Several web platforms, described in Table 1, allow access to water level observations at pre-determined locations, defined as virtual stations. Copernicus Global Operational Land Service [14] uses Jason-3 and Sentinel-3 data to obtain water level measurements over 12,901 virtual stations. Hydroweb [15] recently (November 2020) increased their number of virtual stations from 1733 to 11,336, in addition to 124 large lakes. Observations are refreshed at the latest 1.5 days after the availability of a new altimetric measurement in the Hydroweb platform. Dahiti [16] currently provides not only water level measurements for 4411 stations but also for some of them, surface areas, volume variations, bathymetry, water occurrence masks, land-water masks, hypsometry and river discharge with a latency of 1–2 days. G-REALM [17] monitors, in real time, lakes over 100 km^2 around the world, relying on Jason-3 and other past missions for historical data since 1992. However, the information provided suffer from spatio-temporal coverage sparsity at a global scale, as they are only available at predefined locations from existing web platforms. Not only large regions of the world are being underrepresented, but missing data are observed in complex topographies (mountainous landscapes, seasonal ice cover), which require more advanced processing algorithms. Cos et al. [18] give a truthful and suited listing of the inherent reasons behind these limitations.

Table 1. Open access water dynamics maintained datasets at predefined locations. Includes name of the dataset (Name), satellite missions used (Sat), latency (Lat), number of virtual stations (N_{vs}), method references (Ref). Satellite missions' acronyms are referred as TOPEX/Poseidon (T/P), Jason(JS) 1-2-3, Sentinel(S) 3, SARAL/Altika (SRL).

Name	Sat	Lat	N_{vs}	Ref.
Copernicus Global Land Operational Service (CGLOS) (https://land.copernicus.eu/global/themes/water, accessed on 1 December 2020)	T/P, JS-1, JS-2, JS-3, S-3	4 days	12,901	[15,19]
Hydroweb (https://www.theia-land.fr/en/product/water-levels-of-rivers-and-lakes-hydroweb, accessed on 1 December 2020)	T/P, JS-1, JS-2, GFO, ENVISAT	1.5 days	11,460	[15,20,21]
DAHITI (https://dahiti.dgfi.tum.de/en/products/water-level-altimetry/, accessed on 1 December 2020)	T/P, JS-1, JS-2, JS-3, GFO, ERS-1, ERS-2, Cryosat-2, SRL, ENVISAT	1–2 days	4411	[16]
Global Reservoirs and Lakes Monitor (G-REALM) (https://ipad.fas.usda.gov, accessed on 1 December 2020)	T/P, JS-1, JS-2, JS-3, GFO, ENVISAT e	7–10 days	379	[17]

When water level data are not available, their computation requires altimetry data access at different levels of processing and from the various satellite missions providing measurements. However, radar altimeter orbits and elevation retrieval technology were originally designed for observing oceans, implying a design of their spatial resolution limiting to 1000 m river width measurements for most of the missions [22]. Moreover, retrieving water heights from altimeter range data requires a great deal of processing for rivers, as surrounding land introduces noise in the system conceived for open water. As indicated in the Method column of Table 1, references of the algorithms used to extract water level heights from altimetry data show that these large datasets rely on L2 data-based products with ocean or ice retrackers [15,20,21,23] and applying post-processing outlier rejections, while DAHITI [16] uses L1 data-based algorithms. Although the BRAT toolbox [24] was conceived to process these raw data, downloading and processing the altimetry data for large-scale sites remain a challenge, which the AltEx software [25] intended to tackle in 2019. An open-source platform with a web-viewer was built to explore altimetry database availability and access water level time series information on-the-fly. However, considered missions were limited to Jason 2-3 and the web-viewer does not seem

to be maintained anymore, neither does the API access to data. Moreover, AltEx does not verify whether users have selected points over water.

Recently, [26] stated that the greatest challenge of extracting insights from altimetry is actually educational in order to set up synergies with hydrologists. There is a strong need for operational remote sensing-derived products with "continuity of data services, standardization of information, characterization of errors and accuracy" to engage miscellaneous types of end-users in benefiting from this information.

The call for world-scale tracking of water level heights firstly relies on mapping all possible water body areas that can be monitored on a regular basis, called virtual stations. In this study, a cutting-edge operational virtual station extractor algorithm (denominated *teroVIR*) has been developed for multi-altimetry missions going beyond predetermined location availability. Moreover, an operational, robust and consistent large geospatial–temporal data processor (denominated *teroWAT*) has been implemented, relying on a constellation of Level 2 (L2) altimetry mission data (Sentinel 3A-B, Jason 1-2-3 and SARAL/Altika) to obtain continuous water level information from 2001 over large areas barely covered until today, such as the Arctic or the Ishim and Nura river basins in Kazakhstan. The platform agnostic software *teroWAT* has then been adapted to systematically run in near real-time, covering all of West Africa and validated across the whole region (17 countries) with a standardized and state-of-the-art proposed metric. The community of scientific users of the *teroWAT* software has contributed to the development of products such as water volume variations and embedded visualisation, fostering the use of remote sensing information in hydrological applications. An L2 data-based algorithm [27] has been found to have limited accuracy in the Arctic region due to frozen water in winter time. Therefore, a Level 1 (L1) data-based algorithm [28] has been developed in an operational fashion and plugged into the existing software to deliver unrivalled water heights information.

In Section 2, the methodologies are introduced, firstly detailing the datasets. Secondly, a description of *teroVIR*, a world scale virtual stations extractor, is provided. Thirdly, the modulable and all terrain *teroWAT* software is presented by functional block highlighting its interoperability. Finally, we expose the conventional statistical metrics used to validate water level heights and propose a combined innovative one. In Section 3, the results, such as unprecedented water surface dynamics of unsurveyed areas, are revealed, as well as large-scale validation of water level time series in West Africa. Furthermore, we show improvements by the L1-based processor for complex terrain handling and illustration of the compliance to the mosaic of users' needs. In Section 4, the methods and results are discussed.

2. Material and Methods

2.1. Datasets

Twenty-five years of available satellite altimetry missions [26] provide the opportunity to build a long record of water level data. Figure 1 depicts the possibility of temporal continuity of observations thanks to overlapping missions' lifetime and high revisit time (roughly bi-monthly/monthly) at medium spatial resolution (300 m up to 2.2 km). However, each mission follows its specific nominal orbit, impeding the persistence of spatial coverage when the mission is over. The processors developed in this study only rely on currently flying missions, namely Saral/ALTIKA(SRL), Jason(J) missions 1-2-3 (which followed one after each other the nominal orbit) and Sentinel(S)3A-3B. While CryoSat-2 is currently active, this mission suffers from low revisit time (369 days) for most of the world since it is focusing on the polar regions, and is therefore not included as an input of the processor. As opposed to imaging satellite missions, the altimetry tracks are sparse and do not offer the possibility to fly over the entire world. Fortunately, water level heights disclose the advantage of being physical quantities spatially connected along the same water body and the sparsity of the measurements is then counterbalanced. As a result, observational requirements encompass permanent water bodies frequently flown over by

altimetry missions in favour of spatio-temporal consistency. Permanent water extents are extracted from the JRC GSWE dataset [29], which expresses the water occurrence value per pixel in percentage aggregated over thirty years (1979–2019) at 30 m resolution. Spatial resolution limitations of the altimetry missions and geometry of the water bodies entail constraints on their validity to extract water level heights. In this sense, the virtual stations extractor *teroVIR* aims at inducing all possible locations in the world which would allow for continuous water level observations, without the need for setting them up a priori.

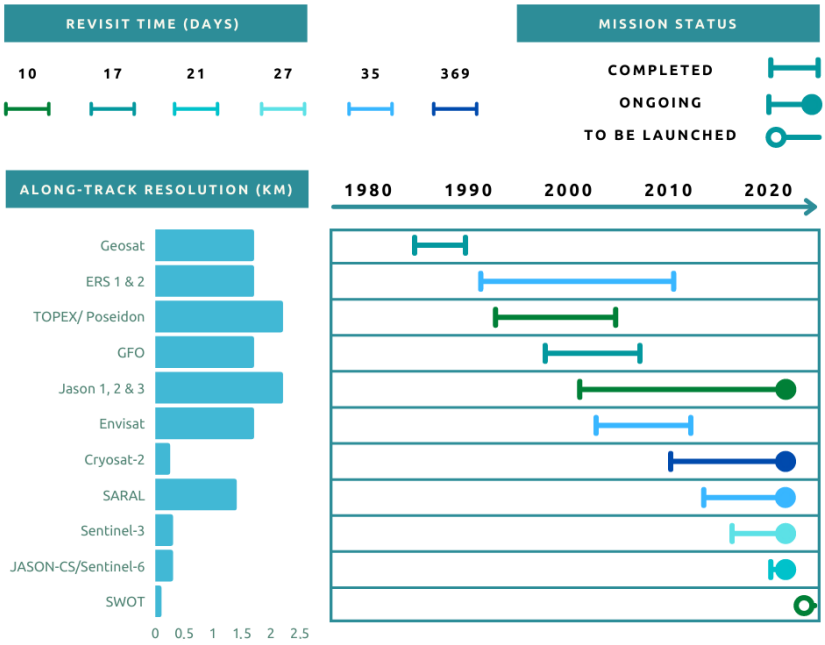


Figure 1. Satellite altimetry missions compatible with *teroVIR* and *teroWAT* processors, from first available operational data (GeoSat, starting 1986) until present and future launches.

2.2. teroVIR: World Scale Coverage Virtual Stations Extractor

The *teroVIR* processor derives virtual stations by crossing the altimeters' buffered nominal reference tracks with water masks extracted from JRC GSWE. In order to fulfill long data record requirements [26], no interleaved and drifting orbits are considered, since satellite missions follow their nominal reference tracks repeatedly during most of their operational life. In the following, we describe the built *teroVIR* framework, illustrated in Figure 2.

Satellite nominal tracks harmonisation

World coverage of the nominal satellite tracks is retrieved from the data providers of each satellite mission and read by the software as vector objects. Geometries of the tracks firstly have to be formatted to respect continuous coordinate sets of points encompassing latitudes between −90 and +90 degrees and longitudes between −180 and +180 degrees. Harmonisation of the multi-mission track geodata follows so that each satellite track object discloses the following metadata (essential for uniqueness denomination of the virtual stations): satellite name, absolute orbit, relative orbit.

User inputs

Virtual stations are extracted for the areas of interest provided as polygons by the user. In the case of regions with high water occurrence coverage (like in the Arctic), river beds can be provided as linestrings. In those regions, main rivers and reservoirs of interest drive the water dynamics impact within the hydrological catchment. By coarsely pre-setting the lines of geometry of the river beds of interest, only virtual stations along these water

streams are retrieved, significantly reducing the processing time. This process is called Geographic Selection and outputs filtering geometries of interest.

Surface water occurrence

Surface water mapping is conditioned by downloading the JRC GSWE occurrence maps. An adapter parsing the bounding box around the filtering geometries had to be implemented to fit the formatting system used to geo-query the JRC GSWE tiles by their name. Each tile is then downloaded and processed in order to extract, as explained in the following sections, firstly water bodies and secondly virtual stations. The codebase is modulable so that any water bodies dataset released with higher accuracy could be plugged in.

Water bodies module

For each tile downloaded, water occurence pixel values are thresholded (T) to derive a binary mask (water/land). Filtering geometries issued from the Geometric Selection step are intersected with the tile, forming the area of interest from which the water body geometries should be extracted. In case of string lines (representing the river beds), they are buffered by a distance in degrees (b_l) covering the full width potentiality of the river. The output collection of geometries is burned into a raster mask, to allow for the intersection with the water mask. The intersected mask is finally vectorised into a collection of water body geometries. As this collection could contain interrupted and disconnected components, which actually represent the same water body, traditional image processing operations of erosion followed by dilation are carried out. The water body dataset JRC GSWE has a resolution of 30 m, and parameters of erosion e and dilation d are chosen within the same order of magnitude. Finally, sanitised and non-empty geometries are kept as a collection of water bodies for the specified tile. This aforedescribed process represents the Water bodies module.

Virtual stations module

For each satellite mission, only harmonised tracks geometries lying in the tile are kept. As the track geometries are simplified compared to where the real altimetry signal could occur, their spatial expansion has to be modelled. To do so, each of these tracks is buffered by a certain distance in degrees (b_t), which covers the across-track resolution for JS and SRL nominal tracks and the derivation potentiality of the real track from the nominal track specific to the S3 mission. Then, water bodies and track geometries are intersected, outputting a collection of virtual stations. This collection is finally cleaned to only keep geometries within certain area extent boundaries (min_a and max_a). The inferior area extent bound ensures that the virtual station has a high probability to receive an altimetry pulse. JS and SRL work in Low Resolution Mode (LRM) operating mode whereas S3 works in Satellite Aperture Radar (SAR) mode, allowing for a higher resolution along track. Therefore, for JS and SRL missions, the minimum area extent is set as $min_a = \pi \times (\frac{al_{tr}}{2} \times \frac{ac_{tr}}{2})$, while for S3, $min_a = al_{tr} \times ac_{tr}$, where ac_{tr} and al_{tr}, respectively, are the across and along track resolution. The upper boundary max_a is used to discard virtual stations following into oceans or seas, which would then be way too big. All parameters and values are summed up in Table A1.

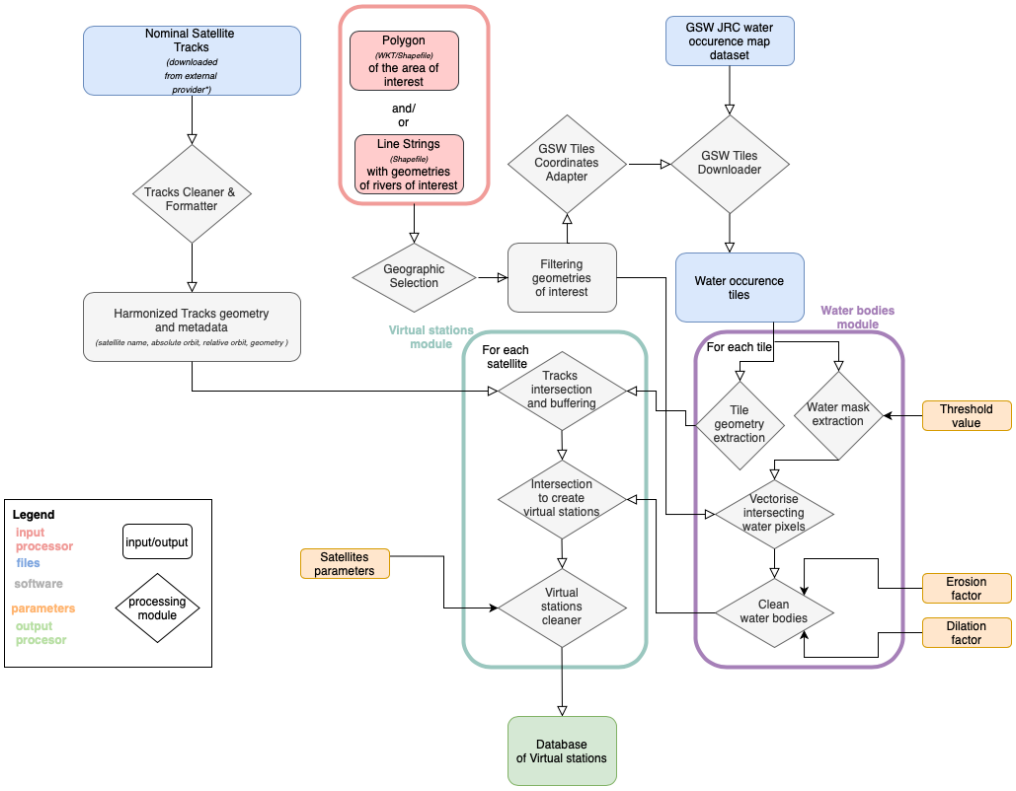


Figure 2. Virtual stations extractor (*teroVIR*) diagram.

2.3. teroWAT: Interoperable All Terrains Water Level Processor

The water level processor *teroWAT* is built under the following rationale. Altimeters send pulses at recurring points in time towards the earth's surface. Onboard trackers gather echoes that are scattered back, which are then accumulated under a power distribution function over time known as "waveform" [13]. Waveforms obtained over open water are accurately processed to obtain the satellite ranges (distance to the water) by estimating their "epoch" [30] with algorithms named "Retrackers" [31,32]. The altimetry L2 products [33] provide directly only the satellite ranges, whereas the L1 ones consist of geolocated waveforms generated by the signal received by the sensor on board of satellites. As shown in Figure 3, the water level processor *teroWAT* was built in blocks in order to ensure the modularity of the software. This way, additional features can be added to the processor so that it can be tailored to the users' needs. The necessary inputs mainly consist of surface water geometries (the so-called virtual stations or the entire water body) and the period of observation. It is then possible to further filter by missions or processing level (L1 or L2). Since the purpose of *teroWAT* is to obtain water level heights at a large number of relevant locations around the world, necessary automatic downloading of a large number of altimetry datasets has to be carried out. To optimize downloading times, only passes close to the virtual station geometries are queried to the data provider, with priority to the most consolidated processed products in terms of timeliness. Several virtual stations can be flown over by the same pass; therefore, duplicates of queries are removed to reduce downloading times. All passes are then downloaded in parallel for the case of multiple virtual stations ensuring an optimised processing. After this, the L2 processing module is also run in parallel for all virtual stations. In this module, for each virtual station, all geo-matching passes containing altimetry datasets are processed. Each altimetry dataset provides data for all pulses of the entire satellite pass crossing the globe but only data points intersecting the water mask geometry are kept for computation. The larger the water body is, the more data points are likely to be found. Satellite ranges and backscatter values for all retrackers available in L2 products and corresponding corrections (wet troposphere, dry troposphere, ionosphere, solid earth tide, geocentric pole tide and ocean loading tide) are

computed [27]. The best retracker is the one presenting the highest number of points falling into the mask and the smaller standard deviation of the basckatter values. An additional filtering step is performed to obtain valid points and reject outliers both within each water body, and within each water level time series. The Interquartile Range Rule [34] is used to first discard pulses with out of range backscatter values for one water body at a certain date, but also to discard water level values from the entire time series associated with a water body at the end of the processing. The output of the *teroWAT* consists of multiple time series of water level values for each of the surface water masks and each mission used as input.

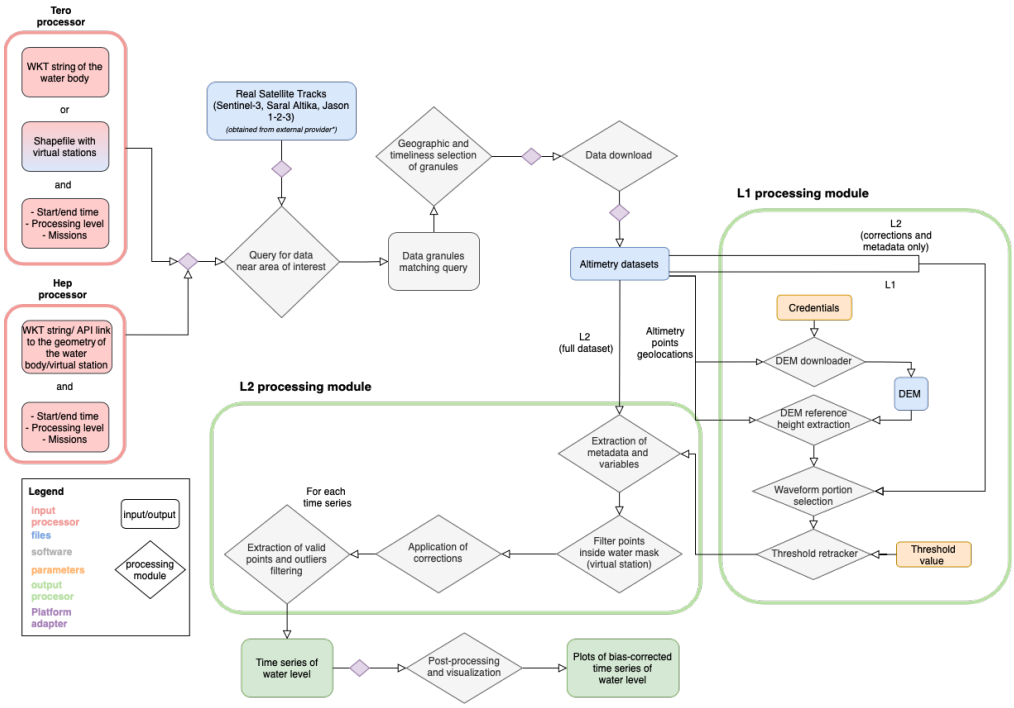


Figure 3. Water level processor *teroWAT* diagram block.

While the L2 product is sufficient for large open water bodies (oceans, lakes), the retracked satellite range can be a complete outlier when land reverberations or frozen water presence contaminate the waveform. In this case, the state-of-the-art algorithm by Gao et al. [28], based on L1 products, has been implemented to filter out too noisy waveforms (based on the number and amplitude of the peaks) and the ones with retracked epoch too far away from the corresponding local elevation value (obtained from a Digital Elevation Model). While in most of the regions of world, the Shuttle Radar Topography Mission (SRTM) DEM is available, for the northernmost and southernmest areas, an adapter has been developed to be able to use the Arctic DEM. The Copernicus DEM with global coverage will be considered in further work. Following the same strategy of modularity, other adapters can be included for regional DEM of higher resolution. As seen in Figure 3, the L1 processing module is stand alone and only requires communication with the altimetry datasets. The L1 processor computes the retracked satellite range value, which becomes an input of the L2 processor.

As all satellite missions operate on different orbits, biases can be introduced that require a cross-calibration of inter-track and inter-satellite biases. For each water surface geometry, mean differences between the overlapping time series of the different missions are set as bias and therefore removed to the water heights computed for each mission. By post-processing the time series with satellite mission bias removal [35], the final outputs of *teroWAT* are mere water level time series for each virtual station.

The *teroWAT* processor can be adapted to any online processing platforms facilitating tailored access as required by users [26]. As shown in Figure 3, platform adapters are only needed for reading the inputs, for communicating with the data providers available in the

platform and for visualisation of the results. This demonstrates the capability of the system to be implemented in on-line platforms by slightly changing the processor, while keeping the main steps unaltered. In other words, the system is platform agnostic.

2.4. Consistent and Inclusive Evaluation Metrics

The water heights are validated by calculating a collection of usual metrics. The unbiased root mean square error (ubRMSE) [28] allows an estimation of the errors between the in situ measurements and the water level datasets, removing the bias induced by the distance between the measurements (<50 km) on the same water body. $(X_1, \ldots X_N)$ are the in situ measurements corresponding to $(Y_1, \ldots Y_N)$, the water level matching N datapoints. ubRMSE is defined as:

$$ubRMSE = \sqrt{\sum_{i=1}^{N}[(X_i - \overline{X}) - (Y_i - \overline{Y})]^2} \quad (1)$$

The Median Absolute Deviation (MAD) provides an assessment of the errors by removing the impacts of potential outlier datapoints. MAD is defined as:

$$MAD = \text{median}(|(X - \overline{X}) - (Y - \overline{Y})|) \quad (2)$$

The Pearson coefficient (R) [36] gives an indication on the correlation of the time series compared. R is defined as:

$$R = \frac{N\sum_{i=1}^{N} X_i \times Y_i - (\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i)}{\sqrt{[N\sum_{i=1}^{N} X_i^2 - (\sum_{i=1}^{N} X_i)^2][N\sum_{i=1}^{N} Y_i^2 - (\sum_{i=1}^{N} Y_i)^2]}} \quad (3)$$

Considering the sparsity of the measurements in remote and difficult to access areas, not all in situ stations have the same sample size and reliability. The ubRMSE and R are more likely to be better for a time series of two datapoints than for more. Based on this observation, we propose to classify the time series extracted by virtual station-mission (VSM) in four reliability categories. If a virtual station is at the crossing of different mission tracks, each mission time series will be evaluated separately as a VSM. The normalised *TeroSCO* (T_s) grants a consistent comparison, combining different metrics, of all virtual stations and is evaluated as:

$$T_s = w_c \times \frac{1}{e^{ubRMSE}} \times R \quad (4)$$

where c is the reliability category and corresponding weight w_c. The definition of the categories is as such:

- Very low coverage category: $N = 1$ datapoint available and $w_c = 0$ since only one measurement does not define a reliable time series.
- Low coverage category: $1 < N \leq 3$ datapoints available, defining a minimum of points for a reliable time series and $w_c = 0.1$
- Medium coverage category: $3 < N \leq m$ datapoints available, defining around at least one year of data and $w_c = 0.35$. For Jason mission, $m = 35$, whereas for other missions $m = 12$, considering their revisit time.
- High coverage category: $N > 12$ datapoints available, defining more than one year of data and $w_c = 0.55$

The *TeroSCO* ensures an interpretable metric (normalised range), favouring datasets with the highest number of measurements (N), highest accuracies (ubRMSE) and highest correlations (R).

3. Results

3.1. Coverage

The *teroVIR* and *teroWAT* processors have been run for five areas of interest around the world. Virtual stations extracted with *teroVIR* are shown in Figure 4 in North of Kazakhstan, in the Arctic and in West Africa. Table 2 shows that Kazakhstan, Arctic in the West and East Russia were barely covered by CGLOS, Hydroweb or Dahiti, while North America and West Africa are at best respectively one sixth and half less covered than with *teroVIR*.

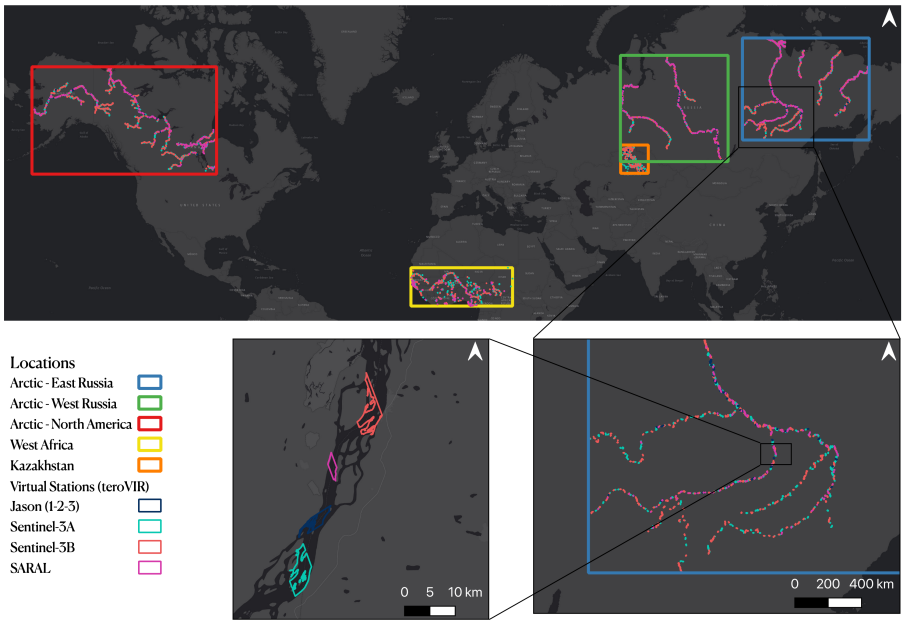


Figure 4. Virtual stations provided by *teroVIR* in the different Areas Of Interest (background from ESRI).

Table 2. Benchmark of virtual stations provided in different areas of interest by the different platforms.

AOI	*teroVIR*	CGLOS	Hydroweb	Dahiti
Arctic—North America	1677	269	132	62
Arctic—West Russia	828	6	90	9
Arctic—East Russia	1454	8	21	3
West Africa	1078	478	549	332
Kazakhstan	373	6	12	3

The total surface covered in Kazakhstan is 283,905 km^2 (roughly 10% of the total surface of this country), which has been suffering from severe flood events while its monitoring remains very sparse, as demonstrated. The Arctic area is covered with small water bodies, which may not influence much the water balance of a catchment, and therefore having prior information of major river lines helps to only run the *teroVIR* processor where needed. These river lines have been derived by hydrological experts and divided into three regions, namely North America, West Russia and East Russia. A description of the missions providing data for all areas of interest stations can be observed in Figure 5. While the JS 1-2-3 sequence ensures 10 years of continuous data, the spatial coverage is significantly lower than for S3A and S3B, which are only available since 2016 and 2018, respectively, and than for SRL, deprecated (on drifting orbit mode) since July 2016.

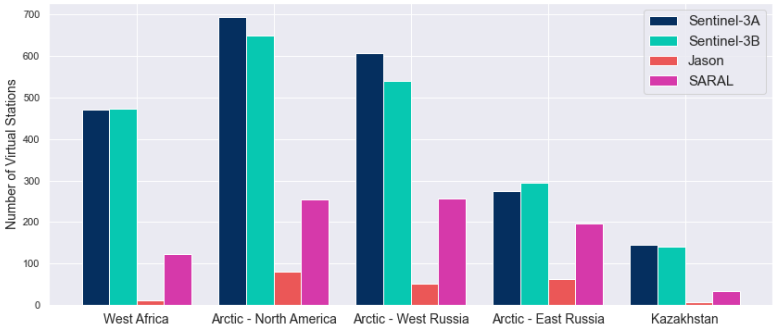


Figure 5. Virtual station counts by mission in each region.

3.2. Accuracy

In West Africa alone, 727 virtual stations provide data in near real time for the SRL, JS 1-2-3 and S3A-3B missions. In 17 countries, 19 in situ stations have overlapping datapoints with 36 virtual stations. All metrics have been computed by reliability category for the 46 VSMs available in West Africa and displayed in Table 3. On the one hand, the less reliable the category, the better the metrics (ubRMSE, MAD and R). The medium coverage category also has more than double the VSMs as other ones. On the other hand, *TeroSCO*'s are higher on average for the high coverage category, giving the evidence of its consistent design.

Table 3. Averaged metrics of all VSMs in West Africa by reliability category.

Category	ubRMSE (m)	MAD (m)	R	T_s	Number of VSMs
Very low	0	0	0	0	6
Low	0.488	0.262	0.999	0.061	8
Medium	0.748	0.461	0.816	0.161	21
High	0.886	0.678	0.890	0.216	11

For all VMSs, a mean score of 0.136 is found and a standard deviation of 0.104, revealing a large disparity of the VMSs accuracy, which is also verified for the metrics ubRMSE (mean of 0.638 m and standard deviation of 0.447 m) and MAD (mean of 0.418 m and standard deviation of 0.528 m). However, while the respectively worst ubRMSE and MAD reach 1.446 and 1.772 m, no more than 75% of the VMSs respectively go over 0.988 and 0.460 m. Therefore, most of the VMSs present state-of-the-art accuracy.

In order to assess the *TeroSCO* liability and validate the *teroWAT* algorithm againts existing datasets, we chose three in situ stations (Kirango and Koulikoro [Mali], Umaisha [Nigeria]) where all the data providers (CGLOS, Dahiti and Hydroweb) disclosed virtual stations. Table 4 reports the best ubRMSE, MAD and R and the worst T_s for *teroWAT*, since on average, the VMSs disclose the least datapoints. The *teroWAT* algorithm, through its outliers filtering processing steps, presents destructive behaviour, favouring higher accuracy and less datapoints. The detailed table by in situ station is found in Appendix A. The *TeroSCO* sheds light on the reliability of a VSM with Hydroweb disclosing the highest score.

Table 4. Averaged metrics of VSMs for 3 in situ stations (Kirango, Koulikoro, Umaisha) in West Africa for comparison between CGLOS, Dahiti, Hydroweb and teroWAT data providers. Bold values represent the best statistically significant results.

Data Provider	ubRMSE (m)	MAD (m)	N	R	TS
CGLOS	0.971	0.782	13.1	0.845	**0.161**
Dahiti	0.966	0.813	**13.2**	0.841	0.162
Hydroweb	0.961	0.778	13.0	0.846	0.163
teroWAT	**0.889**	**0.707**	10.42	**0.905**	0.153

While the *teroWAT* algorithm based on L2 products requires half the data to download and half the processing time, complex terrain can disclose severe outliers in the results. A comparison between L1 and L2 processing levels has been carried out for several virtual stations in the Lena river, one of the major rivers flowing into the Arctic Ocean. Figure 6 shows how L2 products can include outliers within the retracked values when the signals suffer from land pollution or frozen water backscattering. Unreliable values are being filtered out by the L1 *teroWAT* processor, especially in Kangalassy and Pokrovsk. S3A-derived water levels are only computed after 9 March 2019, due to the failing Open Loop mode in this region for on-board DEM prior to version 5.0. Table 5 reveals the improvement in all metrics for the studied stations, by using L1 instead of the L2 processor. Less datapoints are found for L1, since more outliers are discarded.

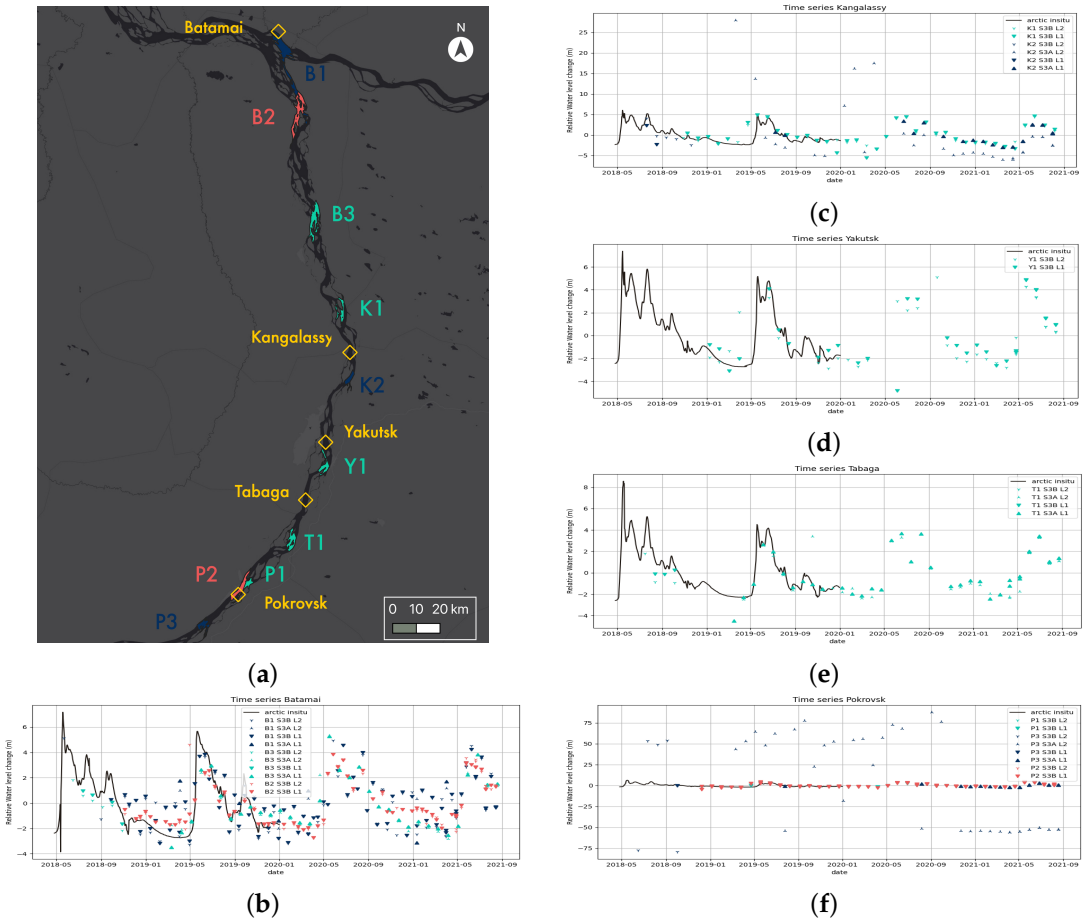


Figure 6. Water level relative change time series in meters for all virtual stations and corresponding in situ stations in the Lena River with (**a**) map with in situ stations (in yellow) and corresponding Virtual Station geometries (background from ESRI), (**b**) Batamai, (**c**) Kangalassy, (**d**) Yakutsk, (**e**) Tabaga, (**f**) Pokrovsk close-by station time series.

Table 5. Averaged metrics of VSMs for 5 in situ stations (Batamai, Kalangassy, Yakutsk, Tabaga, Pokrovsk) in the Lena River for comparison between L1 and L2 *teroWAT* processors. The in situ measurements are only available until 2020. Bold values represent the best statistically significant results.

teroWAT Processors	ubRMSE (m)	MAD (m)	N	R	TS
L1	**0.369**	**0.737**	9.25	**0.658**	**0.176**
L2	4.480	1.765	**10.6**	0.572	0.148

3.3. Fit for User Needs

The modularity of the processors allows for easily inputting their outputs into various applications required by the users. Common specification lies in retrieving long data records to retrace past scenario. Monthly water volume variations obtained from the combination of the output of the water level processor and water masks have, for instance, been computed for flood characterization in several unprecedented locations in Kazakhstan (https://labs.dhi-gras.com/eo-clinic-kazakhstan, accessed on 19 July 2021) and for water quality assessment in Lake Volta and Lake Victoria (http://sdg6-hydrology-tep.eu, accessed on 19 July 2021) Floods forecast is a topic of interest particularly in Africa where more extreme precipitation is observed at a higher frequency every year during rainy seasons. FANFAR [37] aims to produce an operational flood forecasting and alerts system in West Africa, gathering hydrological experts and civil protection units of 17 countries collaborating with the scientists developing the system. By allying direct user feedback and easy insight access (https://fanfar.eu, accessed on 19 July 2021), taking advantage of the H-TEP (https://hydrology-tep.eu/#!thematic, accessed on 19 July 2021), the all-in hydrology platform, FANFAR enhances the capacity of West African institutions to forecast, alert and manage floods. The flood forecast model uses as input the 727 water level time series (see Section 3.2) produced near real time in automatic production to overcome the sparsity of gauges in this region. Finally, fresh water monitoring for the arctic user community (Pan-arctic user-group: Arctic-HYCOS - national hydrological services in Arctic council member states/GRDC/WMO and Yakutian user-group: Federal, regional, and local stakeholders, and research institutes in Republic of Sakha (Yakutia), HYPE-ERAS (Belmont forum project)) has been carried out in the H-TEP with water discharge computed from water level time series from S3 obtained by the L1 processor.

4. Conclusions

Climate science often targets large-scale spatio-temporal phenomena depicted by long data records [38], many of which can be derived from Earth Observation satellites on a recurrent basis, such as water level related products. We developed an operational software conceptualised with a modulable view, not only allowing to fit the requirements of miscellaneous users but also delivering datasets at a global scale extracted from up and running altimetry missions (since 2001). Within this software, the interoperability of the *teroVIR* processor producing virtual stations and of the *teroWAT* processor delivering corresponding water level time series, guarantees the extension to past (Geosat, ERS 1-2, Topex/Poseidon, GFO, Envisat) and future (SWOT) missions. The recently launched Sentinel-6 mission is planned to be integrated shortly in the *teroVIR* and *teroWAT* processors.

The virtual stations extractor *teroVIR* was built on the scientific and physical knowledge of multi-mission/dataset types to map all possible water body areas that can be water level monitored on a regular basis. Unprecedented virtual stations in barely covered area to date have been found in high numbers in Kazakhstan and the Arctic. In West Africa, an increase of 55.6%, 49.0%, 69.2% of extracted virtual stations in comparison respectively with CGLOS, Hydroweb and Dahiti (dispensing predetermined stations), has been achieved. However, limitations are found for the virtual stations whose shape extension follows the same orientation as the satellite track in Figure 7. Since the bathymetry and width of the water body, as well as the geoid local values, can significantly vary along the river bed, the average water height value computed at each datapoint will not give an accurate representation of the distribution of all water level values of each pulse. Further refinement of the *teroVIR* based on the alignment of the virtual station with the track geometry should allow us to split it up. Moreover, the multitude of virtual stations found over large water bodies such as lakes should be merged into a unique one since they are all connected on a hydrological physical interpretation. Finally, since the *teroVIR* targets the production of time continuous data records, only permanent water bodies have been considered, though many regions present a high inter-seasonal variation in water availability such as

West Africa. An evolution of the processor considering seasonal water bodies would give unprecedented information on water heights over the years.

The *teroWAT* processor modularity ensured a smooth evolution guided by the needs of the users. Key stakeholders asking for data in remote locations can also be data providers. The access given by the hydrologists of West Africa to gauges allowed, to the best of our knowledge, the first large scale validation for water level products across such a large region. An average of ubRMSE of 0.638 m and MAD of 0.418 m have been found for 36 virtual stations for the L2 processor. The *TeroSCO* introduced, combining several metrics to assess with a normalised value, the reliability of a virtual station and corresponding water level heights accuracy, proved its effectiveness by favouring virtual stations with more datapoints (N), lower ubRMSE and higher Pearson coefficient (R). This score allows an unequivocal understanding of the quality of the validation of time series, by removing the bias introduced by the number of samples compared. Other datasets from CGLOS, Dahiti, Hydroweb providers have been compared against *teroWAT*, the latest showing better conventional metrics (ubRMSE, MAD, pearson) while lower *TeroSCO*. The accuracy of the L1 *teroWAT* processor has been improved, targeting complex terrain regions barely covered until today, like the Arctic, with the modular addition of the L1 module. Results significantly show the effectiveness of the L1 processor in discarding the polluted waveforms by terrain contamination and in improving the accuracy for the selected ones. The *teroWAT* undeniably discloses a destructive strategy with its outlier processing steps, removing unreliable datapoints benefiting for smaller errors but penalizing the *TeroSCO*. In parallel to removing the outliers postprocessing on water level heights, further adjustments on only retaining the high backscatter values (e.g., >15 dB in Ku band and >20 dB in S band [24]) for each pulse could help get rid of outliers that remain in the results. These values are obtained on very flat surfaces, such as deserts, large river basins or wetlands due to the specularity of the return radar echo.

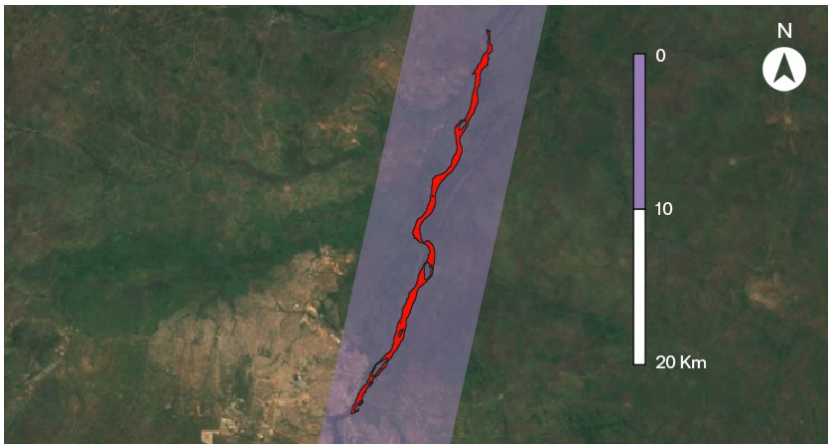


Figure 7. Virtual station (in red) leading to unreliable water level time series. The orientation of the river bed follows the S3 satellite track (in purple) over almost 30 km. For each date, all pulse observations falling into this geometry are averaged introducing artifacts due to great variations occuring along the river.

As demonstrated in Section 3.3, any region of the world can then be monitored in near real time, opening for scientific collaborations and actual usage of EO data as a driver for change implementation. Among the sundry purposes presented, water surface dynamics have disclosed their relation to flood forecast and river discharge estimation in complex terrains with the L1 processor. Since the flexiblity of the processors have been at the center of the development of the software, improvement of the computation core can easily be implemented while connection to any platform or viewer is facilitated. Finally, the developed processors, providing water level time series information on-the-fly, could be

adapted to an unrivaled platform that allows users to query the large global raw altimetry database dynamically.

Author Contributions: Conceptualisation, M.M.; data curation, M.M. and M.P.-V.; investigation, M.M. and D.G.; methodology, M.M.; supervision, L.R. and M.J.E.; validation, M.M., M.P.-V. and D.G.; visualisation, M.M. and M.P.-V.; writing—original draft, M.M. and M.P.-V.; writing—review and editing, M.M., L.R. and M.J.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: A dataset containing a sample of locations across Siberia and Africa is associated with this study. The data can be accessed (CC BY-NC-SA 4.0) at Mélissande Machefer, Martí Perpinyà-Vallès, Maria Jose Escorihuela, David Gustafsson and Laia Romero (2022). Virtual stations (TeroVIR) and water level time series (TeroWAT) in West Africa and Arctic regions [Dataset]. In Remote Sensing MDPI https://doi.org/10.5281/zenodo.6284704.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Parameters for the virtual station extractor.

Parameter	Name	Value		
T	Threshold	75%		
b_l	Line buffer	0.1 degrees		
e	Erosion	20 m		
d	Dilation	30 m		
		Sentinel 3	Saral/ALTIKA	Jason
bt	Track buffer	5000 m	1400 m	2200 m
ac_{tr}	Across track resolution	1640 m	1400 m	2200 m
al_{tr}	Along track resolution	300 m	1400 m	2200 m
min_a	Minimum area	0.5 km^2	1.6 km^2	3.8 km^2
max_a	Maximum area	10,000 km^2	10,000 km^2	10,000 km^2

Table A2. Detailed averaged metrics of VS-mission time-series for 3 in situ stations (Kirango, Koulikoro, Umaisha) in West Africa for CGLOS, Dahiti, Hydroweb and *teroWAT* data providers.. Bold values represent the best statistically significant results.

In Situ Station	Data Provider	ubRMSE (m)	MAD (m)	N	R	TS
Kirango	CGLOS	**1.260**	**1.393**	14.0	0.634	0.099
	Dahiti	1.296	1.548	**14.5**	0.647	0.098
	Hydroweb	**1.260**	**1.393**	14.0	0.634	0.099
	teroWAT	1.279	1.434	13.25	**0.787**	**0.109**
Koulikoro	CGLOS	0.741	0.351	14.0	0.963	**0.253**
	Dahiti	0.737	0.320	**14.5**	0.949	0.252
	Hydroweb	0.741	0.351	14.0	0.963	**0.253**
	teroWAT	**0.483**	**0.115**	8.0	**0.989**	0.215
Umaisha	CGLOS	0.834	0.602	11.3	0.938	0.132
	Dahiti	**0.746**	0.572	10.5	0.926	**0.137**
	Hydroweb	0.779	0.590	11.0	0.940	0.136
	teroWAT	0.812	**0.571**	10.0	**0.9401**	0.133

References

1. Vignudelli, S.; Kostianoy, A.G.; Cipollini, P.; Benveniste, J. *Coastal Altimetry*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 1–565. [CrossRef]
2. Haritashya, U.K.; Singh, P.; Kumar, N.; Singh, Y. Hydrological importance of an unusual hazard in a mountainous basin: Flood and landslide. *Hydrol. Process.* **2006**, *20*, 3147–3154. [CrossRef]
3. Huntington, T.G. Evidence for intensification of the global water cycle: Review and synthesis. *J. Hydrol.* **2006**, *319*, 83–95. [CrossRef]
4. Chen, X.; Hao, Z.; Devineni, N.; Lall, U. Climate information based streamflow and rainfall forecasts for Huai River basin using hierarchical Bayesian modeling. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 1539–1548. [CrossRef]
5. Alsdorf, D.; Birkett, C.; Dunne, T.; Melack, J.; Hess, L. Water level changes in a large Amazon lake measured with spaceborne radar interferometry and altimetry. *Geophys. Res. Lett.* **2001**, *28*, 2671–2674. [CrossRef]
6. Qiao, B.; Zhu, L.; Wang, J.; Ju, J.; Ma, Q.; Huang, L.; Chen, H.; Liu, C.; Xu, T. Estimation of lake water storage and changes based on bathymetric data and altimetry data and the association with climate change in the central Tibetan Plateau. *J. Hydrol.* **2019**, *578*, 124052. [CrossRef]
7. Duan, Z.; Bastiaanssen, W. Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sens. Environ.* **2013**, *134*, 403–416. [CrossRef]
8. Abileah, R.; Vignudelli, S.; Scozzari, A. A Completely Remote Sensing Approach To Monitoring Reservoirs Water Volume. *Int. Water Technol. J.* **2011**, *1*, 63–77.
9. Van Den Hoek, J.; Getirana, A.; Jung, H.C.; Okeowo, M.A.; Lee, H. Monitoring reservoir drought dynamics with landsat and radar/lidar altimetry time series in persistently cloudy eastern Brazil. *Remote Sens.* **2019**, *11*, 827. [CrossRef]
10. Schröder, S.; Springer, A.; Kusche, J.; Uebbing, B.; Fenoglio-Marc, L.; Diekkrüger, B.; Poméon, T. Niger discharge from radar altimetry: Bridging gaps between gauge and altimetry time series. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 4113–4128. [CrossRef]
11. Tarpanelli, A.; Santi, E.; Tourian, M.J.; Filippucci, P.; Amarnath, G.; Brocca, L. Daily River Discharge Estimates by Merging Satellite Optical Sensors and Radar Altimetry Through Artificial Neural Network. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 329–341. [CrossRef]
12. Bogning, S.; Frappart, F.; Blarel, F.; Niño, F.; Mahé, G.; Bricquet, J.P.; Seyler, F.; Onguéné, R.; Etamé, J.; Paiz, M.C.; et al. Monitoring water levels and discharges using radar altimetry in an ungauged river basin: The case of the Ogooué. *Remote Sens.* **2018**, *10*, 350. [CrossRef]
13. Vignudelli, S.; Scozzari, A.; Abileah, R.; Gómez-Enri, J.; Benveniste, J.; Cipollini, P. *Water Surface Elevation in Coastal and Inland Waters Using Satellite Radar Altimetry*; Elsevier Inc.: Amsterdam, The Netherlands, 2019; pp. 87–127. [CrossRef]
14. Buchhorn, M.; Lesiv, M.; Tsendbazar, N.E.; Herold, M.; Bertels, L.; Smets, B. Copernicus Global Land Cover Layers—Collection 2. *Remote Sens.* **2020**, *12*, 1044. [CrossRef]
15. Crétaux, J.F.; Arsen, A.; Calmant, S.; Kouraev, A.; Vuglinski, V.; Bergé-Nguyen, M.; Gennero, M.C.; Nino, F.; Abarca Del Rio, R.; Cazenave, A.; et al. SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Adv. Space Res.* **2011**, *47*, 1497–1507. [CrossRef]
16. Schwatke, C.; Dettmering, D.; Bosch, W.; Seitz, F. DAHITI—An innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 4345–4364. [CrossRef]
17. Birkett, C.; Reynolds, C.; Beckley, B.; Doorn, B. From research to operations: The USDA global reservoir and lake monitor. In *Coastal Altimetry*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 19–50. [CrossRef]
18. Coss, S.; Durand, M.; Yi, Y.; Jia, Y.; Guo, Q.; Tuozzolo, S.; Shum, C.K.; Allen, G.H.; Calmant, S.; Pavelsky, T. Global River Radar Altimetry Time Series (GRRATS): New river elevation earth science data records for the hydrologic community. *Earth Syst. Sci. Data* **2020**, *12*, 137–150. [CrossRef]
19. Calmant, S.; da Silva, J.S.; Moreira, D.M.; Seyler, F.; Shum, C.; Crétaux, J.F.; Gabalda, G. Detection of Envisat RA2/ICE-1 retracked radar altimetry bias over the Amazon basin rivers using GPS. *Adv. Space Res.* **2013**, *51*, 1551–1564. [CrossRef]
20. Santos da Silva, J.; Calmant, S.; Seyler, F.; Rotunno Filho, O.C.; Cochonneau, G.; Mansur, W.J. Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions. *Remote Sens. Environ.* **2010**, *114*, 2160–2181. [CrossRef]
21. Normandin, C.; Frappart, F.; Diepkilé, A.T.; Marieu, V.; Mougin, E.; Blarel, F.; Lubac, B.; Braquet, N.; Ba, A. Evolution of the Performances of Radar Altimetry Missions from ERS-2 to Sentinel-3A over the Inner Niger Delta. *Remote Sens.* **2018**, *10*, 833. [CrossRef]
22. Lettenmaier, D.P.; Alsdorf, D.; Dozier, J.; Huffman, G.J.; Pan, M.; Wood, E.F. Inroads of remote sensing into hydrologic science during the WRR era. *Water Resour. Res.* **2015**, *51*, 7309–7342. [CrossRef]
23. Calmant, S.; Crétaux, J.F.; Rémy, F. 4-Principles of Radar Satellite Altimetry for Application on Inland Waters. In *Microwave Remote Sensing of Land Surface*; Baghdadi, N., Zribi, M., Eds.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 175–218. [CrossRef]
24. Benveniste, J.; Rosmorduc, V.; Niemeijer, S.; Picot, N. Basic Radar Altimetry Toolbox. In Proceedings of the IGARSS 2008—2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 8–11 July 2008; Volume 3, pp. II-895–III-898. [CrossRef]
25. Markert, K.N.; Pulla, S.T.; Lee, H.; Markert, A.M.; Anderson, E.R.; Okeowo, M.A.; Limaye, A.S. AltEx: An open source web application and toolkit for accessing and exploring altimetry datasets. *Environ. Model. Softw.* **2019**, *117*, 164–175. [CrossRef]

26. Abdalla, S.; Abdeh Kolahchi, A.; Ablain, M.; Adusumilli, S.; Aich Bhowmick, S.; Alou-Font, E.; Amarouche, L.; Andersen, O.B.; Antich, H.; Aouf, L.; et al. Altimetry for the future: Building on 25 years of progress. *Adv. Space Res.* **2021**, *68*, 319–363. [CrossRef]
27. Garcia-Mondéjar, A.; Val, M.; Gao, Q.; Escorihuela, J.; Nilo García, P.; Yang, J.; Liao, J. Measuring the lake level evolution in the Qinghai-Tibet plateau with radar altimeters. In Proceedings of the DRAGON-3 Final Results Symposium, Wuhan, China, 4–8 July 2016; Volume 739.
28. Gao, Q.; Makhoul, E.; Escorihuela, M.J.; Zribi, M.; Seguí, P.Q.; García, P.; Roca, M. Analysis of retrackers' performances and water level retrieval over the Ebro River basin using sentinel-3. *Remote Sens.* **2019**, *11*, 718. [CrossRef]
29. Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [CrossRef]
30. Brown, G. The average impulse responce of a rough surface and its applications. *IEEE J. Ocean. Eng.* **1977**, *2*, 67–74. [CrossRef]
31. Makhoul, E.; Roca, M.; Ray, C.; Escolà, R.; Garcia-Mondéjar, A. Evaluation of the precision of different Delay-Doppler Processor (DDP) algorithms using CryoSat-2 data over open ocean. *Adv. Space Res.* **2018**, *62*, 1464–1478. [CrossRef]
32. Davis, C. A robust threshold retracking algorithm for measuring ice-sheet surface elevation change from satellite radar altimeters. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 974–979. [CrossRef]
33. Agency, E.S. Sentinel-3 ESAś Global Land and Ocean Mission for GMES Operational Services. 2012. Available online: https://sentinel.esa.int/documents/247904/351187/S3_SP-1322_3.pdf (accessed on 18 April 2020).
34. Dekking, F.M.; Kraaikamp, C.; Lopuhaä, H.P.; Meester, L.E. *A Modern Introduction to Probability and Statistics*; Springer: London, UK, 2005. [CrossRef]
35. Tong, X.; Pan, H.; Xie, H.; Xu, X.; Li, F.; Chen, L.; Luo, X.; Liu, S.; Chen, P.; Jin, Y. Estimating water volume variations in Lake Victoria over the past 22 years using multi-mission altimetry and remotely sensed images. *Remote. Sens. Environ.* **2016**, *187*, 400–413. [CrossRef]
36. Zaidi, A.Z.; Vignudelli, S.; Khero, Z.; Ghauri, B.M.; Muzaffer, R.; Naeem, B.; Panhwar, V.; Zafar, S.; ul Haque, S. Indus river water level monitoring using satellite radar altimetry. *Adv. Space Res.* **2021**, *6*, 641–651. [CrossRef]
37. Andersson, J.C.; Ali, A.; Arheimer, B.; Gustafsson, D.; Minoungou, B. Providing peak river flow statistics and forecasting in the Niger River basin. *Phys. Chem. Earth Parts A/B/C* **2017**, *100*, 3–12. [CrossRef]
38. Zhai, P.; Pirani, A.; Connors, S.; Péan, C.; Berger, S.; Caud, N.; Chen, Y.; Goldfarb, L.; Gomis, M.; Huang, M.; et al. *IPCC 2021: Climate Change 2021: The Physical Science Basis*; Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Technical Report; Cambridge University Press: Cambridge, UK, 2021.