Surface Water and Ocean Topography (SWOT) Project

SWOT Product Description
Long Name: Level 2 KaRIn high rate river average vector product
Short Name: L2_HR_RiverAvg

Revision B

Document Custodians:

Cassie Stuurman JPL Algorithm Engineer	Date	Claire Pottier CNES Algorithm Engineer	Date

Approved by:

Curtis Chen JPL Algorithm System Engineer	Date	Roger Fjørtoft CNES Algorithm System Engineer	Date

Concurred by:

Stirling Algermissen JPL SDS Manager	Date	Lionel Zawadzki CNES Hydrology Distribution Center Manager	Date

Paper copies of this document may not be current and should not be relied on for official purposes. The current version is in the JPL Engineering Product Data Management system (EPDM: https://epdm.jpl.nasa.gov) and the CNES Product Data Management System

See Sect 1.4 for how to cite this document

October 26, 2023
JPL D-56414

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology

© 2023 California Institute of Technology. Government sponsorship acknowledged.

CHANGE LOG

VERSION	DATE	SECTIONS CHANGED	REASON FOR CHANGE
Preliminary	2017-03-01	ALL	Preliminary version Approved for export (LRR-028890)
Initial Release	2021-09-03	ALL	Initial Release Approved for public release (URS301837/CL#21-3650)
Revision A	2022-10-17	ALL	Revision A Approved for public release (URS312223/CL#22-5807)
Revision B	2023-10-26	ALL	Revision B Fix reference numbering. Approved for public release (URS320831/CL#23-6140)

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table of Contents

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table of Figures

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table of Tables

© 2023 California Institute of Technology. Government sponsorship acknowledged.

List of TBC Items

Page	Section

List of TBD Items

Page	Section

© 2023 California Institute of Technology. Government sponsorship acknowledged.

1 Introduction

1.1 Purpose

The purpose of this Product Description Document is to describe the Level 2 Ka-band Radar Interferometer (KaRIn) high rate (HR) river average (RiverAvg) vector data product from the Surface Water Ocean Topography (SWOT) mission. This data product is also referenced by the short name L2_HR_RiverAvg.

1.2 Document Organization

Section 2 provides a general description of the product, including its purpose (2.1) and temporal latency (2.2).

Section 3 provides the structure of the product, including granule definition (3.1), file organization (3.2), file naming convention (3.3), spatial sampling and resolution (3.4), temporal (3.5) and spatial organization (3.6) of the content, and file size and overall data volume (3.7).

Section 4 provides qualitative descriptions of the information provided in the product.

Section 0 provides a detailed identification of the individual data fields within the L2_HR_RiverAvg product.

Section 6 provides references.

Appendix A provides a listing of the acronyms used in this document.

Appendix B provides a description of the format of the product metadata.

Appendix C provides a description of bitwise quality flags in the product.

1.3 Document Conventions

Where specific names of data variables and groups of the data product are given in the body text of this document, they are represented in *italicized text*.

1.4 Citing This Document

Please cite this document as follows:

> JPL D-56414, Revision B, “SWOT Product Description Document: Level 2 KaRIn High Rate River Average Vector (L2_HR_RiverAvg) Data Product,” Jet Propulsion Laboratory Internal Document, 2023.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

2 Product Description

2.1 Purpose

The L2_HR_RiverAvg product provides river data from an aggregate of continent-passes of the high-rate (HR) data stream of the SWOT KaRIn instrument. The aggregate is compiled for reaches in the prior river database (PRD) [1] using all valid SWOT passes that occur during each observation cycle, and distributed by basin. These data are generally produced for inland and coastal hydrology surfaces, as controlled by the configurable KaRIn HR mask.

The L2_HR_RiverAvg product specifically provides data for river reaches identified in the PRD. Each reach is divided into a number of nodes in the PRD, however the averaged product is provided at the reach level only. For more details, see Sections 4 and 0.

Only rivers that are identified in the PRD are included in the L2_HR_RiverAvg product. Information on lakes, either connected to rivers or not, as well as unidentified water features are given in the L2_HR_LakeAvg science data product [2].

2.2 Latency

The L2_HR_RiverAvg product is generated with a latency of less than 45 days counted from the end of each 21 day observation cycle. The latency allows for consolidation of instrument calibration and the required auxiliary or ancillary data that are needed to generate this product (e.g., orbit ephemeris and external data on geophysical corrections). Different versions of the product may be generated at different latencies and/or through reprocessing with refined input data, such as an updated version of the PRD. Note that the L2_HR_RiverAvg product is only generated for the 21 day (actually 20.86 day) repeat cycles of the SWOT science orbit. It is not generated for the 1 day repeat cycles of the SWOT calibration orbit.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

3 Product Structure

3.1 Granule Definition

The granule size of the data product defines the spatial and temporal extent of the information given in each set of L2_HR_RiverAvg product files. The L2_HR_RiverAvg product aggregates all passes in a 21 day observation cycle, with each set of files covering a Pfafstetter level 2 basin as defined by the PRD and illustrated in Figure 1 for North America. A "pass" is a half-revolution of the Earth by the satellite from pole to pole (south to north latitudes for ascending passes, and north to south latitudes for descending passes). Multiple observations of a given reach are typically made during different passes within the observation cycle.

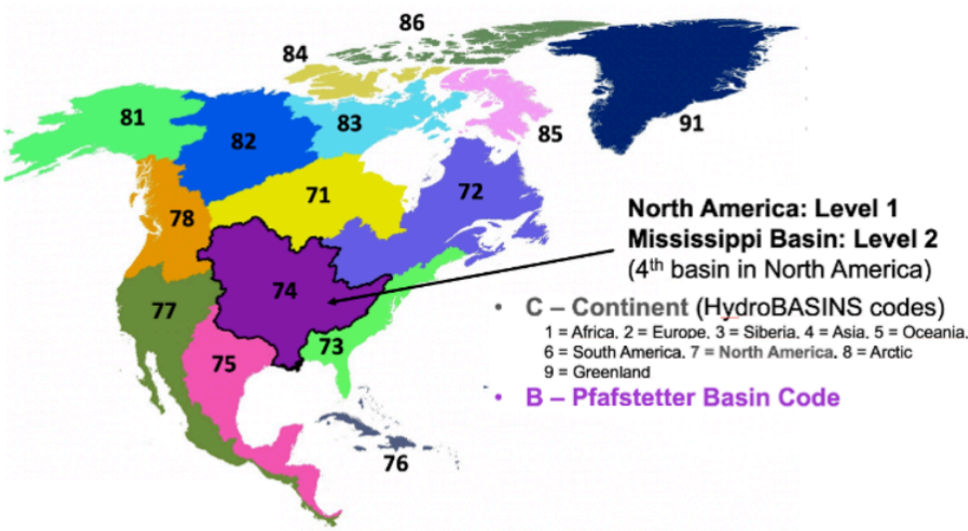

Figure 1. Level 2 Pfafstetter basins in North America, Arctic and Greenland used to define the spatial extent of each L2_HR_RiverAvg set of product files. A single granule represents a Level 2 basin such as the Mississippi basin (74). Basins on other continents are defined similarly.

3.2 File Organization

The L2_HR_RiverAvg product is distributed in the Esri geographical information system (GIS) vector shapefile format [3]. Each granule of the product consists of one shapefile. Each shapefile consists of the five associated files with file name extensions indicated in Table 1.

Table 1. Description of the files representing the L2_HR_RiverAvg reach shapefile.

File	Name	Description
1	Main shapefile with extension .shp	Provides coordinates (linestring shape) of the high-resolution centerline of each reach covered by the product granule.
2	Index file: .shx	Stores the index to each reach in the .shp file
3	Attributes file: .dbf	Provides attributes for each reach in the .shp file
4	Projection file: .prj	Provides map projection and coordinate reference description
5	Metadata file: .shp.xml	Provides metadata for the reach shapefile

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Each file in the shapefile set has the same prefix. The .shp file of each shapefile contains the basic geometry of the reaches as defined by the PRD, so it does not change from pass to pass (unless the PRD is updated). The .dbf file contains the SWOT observations of minimum, mean, median, and maximum river water surface elevation (WSE or height) for the observation cycle, the corresponding area, slope, width, and discharge values for each minimum, median, and maximum WSE as well as the mean river width, slope, and discharge over the observation cycle. It also includes other information from the PRD as described in Section 4. The .prj file contains a projection description, using a well-known text (WKT) representation of coordinate reference systems (CRS). The .shp.xml files, which are not defined by the Esri specification [3], carry metadata applicable across shapefiles (e.g., SWOT cycle number) and per-attribute metadata (e.g., units for each attribute). The format of the .shp.xml files is described in Appendix B.

Note that the use of the term "attributes" in this document follows the shapefile nomenclature in referring to the variables associated with each feature in the .shp file. The term should not be confused with attributes as typically used in the context of NetCDF files. This document uses the term "attributes" in reference to the contents of the .dbf file and uses the term "metadata" in reference to characteristics of each attribute of the entire shapefile. Therefore, as an example, in this document the SWOT-observed average WSE is an attribute of a given reach representing the average water surface elevation in the observation cycle, and the metadata for the *wse_avg* attribute would indicate that the value is given in units of meters. A comprehensive list of attributes included in the product is given in section 5.2.

Note that the names of attributes in shapefiles can be no more than 10 characters, which explains the abbreviated or truncated names of many reach attributes. Owing to this restriction, the naming conventions of attributes in the L2_HR_RiverAvg product sometimes differ from the naming conventions of similar variables in other NetCDF-based SWOT data products. Names of variables in the L2_HR_RiverAvg product are generally abbreviated more compactly.

3.3 File Naming Convention

The L2_HR_RiverAvg product files adopt the following naming convention:

SWOTL2_HR_RiverAvg_<CycleID>_<ContinentID>_<BasinID>_<RangeBeginningDateTime>_<RangeEndingDateTime>_<CRID>_<ProductCounter>.<extension>

The value of <CycleID> above is the number of the observation cycle contained in the shapefile, the <ContinentID> represents the continent identifier from Table 2, the <BasinID> identifies the basin contained in the shapefile through a two-digit basin identifier (see Figure 1). The <CRID> above contains the composite release identifier. It contains the version code of the data product, which changes if the processing software and/or the PRD is updated. The <ProductCounter> identifies the version of product that may have been generated multiple times with the same version of processing software. The <extension> above indicates which of the five parts of the shape file it is (.shp, .shx, .dbf, .prj, or .shp.xml file), per Section 3.2.

An example filename is:

`SWOT_L2_HR_RiverAvg_001_AF_13_20210612T072103_20210612T075103_PGA2_03.shp`

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table 2. Continent codes for the *reach_id* attribute and corresponding continent IDs for the filename.

Continent Code (*C*)	Continent	Continent ID
1	Africa	AF
2	Europe and Middle East	EU
3	Siberia	SI
4	Central and Southeast Asia	AS
5	Australia and Oceania	AU
6	South America	SA
7	North America and Caribbean	NA
8	North American Arctic	AR
9	Greenland	GR

3.4 Spatial Sampling and Resolution

The L2_HR_RiverAvg shapefile contains a collection of geometric elements of the type *polyline*. The coordinates of the polyline vertices correspond to points in the high-resolution (approximately 30 m) centerline represented in the PRD. The L2_HR_RiverAvg product contains a record for each reach in the PRD that is covered by the granule. The record for each reach provides the minimum, median, and maximum observations of water surface elevation and the corresponding values of surface area, cross-sectional area change, slope, width, and discharge over the length of the river covered by that reach. The reaches in the PRD are based on hydrological, morphological, and observational considerations [1]. Reach lengths of approximately 10 km are typical, but they can vary from 5 to 20 km. Most reaches include a water surface area of at least 1 km^2, to allow for enough spatial averaging of the SWOT observations to ensure that estimated WSEs and slopes meet an expected minimum precision [4, 5, 6]. Other considerations for setting reach boundaries include hydrological and morphological features such as tributaries, large islands or channel branching, and edges of the KaRIn measurement swath. Special short reaches are included for known significant WSE changes such as dams [5]. The number of reaches in each L2_HR_RiverAvg granule varies, from approximately 100 to over 20,000.

3.5 Temporal Organization

Each reach is observed a variable number of times from different orbit passes during each 21 day SWOT cycle. The majority of reaches are observed no more than 3 times per SWOT cycle, though this varies significantly with latitude [7, 8] and is illustrated by Figure 3. The minimum number of reach observations is 0 for sections of rivers within the nadir gap that are never sampled, and the maximum expected number of observations is 36 for several high-latitude reaches in Siberia. Note however that reaches with 0 observations throughout the cycle have fill values populating all SWOT-derived fields (see Section 0). See Section 3.5 of the L2_HR_RiverSP Product Description Document (PDD) [9] for information on how reach and node information for individual passes is handled.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

The L2_HR_RiverAvg product provides the mean over these observations of quantities such as WSE, width, slope, and discharge. Because WSE is the fundamental measurement of SWOT, the minimum, maximum, and median WSE values are also provided to better capture the distribution over the observation cycle. Other quantities such as width, slope, and discharge are also provided for the passes corresponding to minimum, maximum, and median WSE, though of course, for example, the slope for the pass with maximum WSE is not necessarily the maximum slope over the cycle. However, average attributes such as *wse_avg*, *wid_avg*, and *slp_avg* representing the water surface elevation, the reach width, and the water surface slope, respectively, are simple averages computed across the single-pass measurements occurring in the observation cycle, considering passes that contained valid WSE measurements. Validity of the single-pass data is defined by the quality flags associated with the data in the L2_HR_RiverSP product.

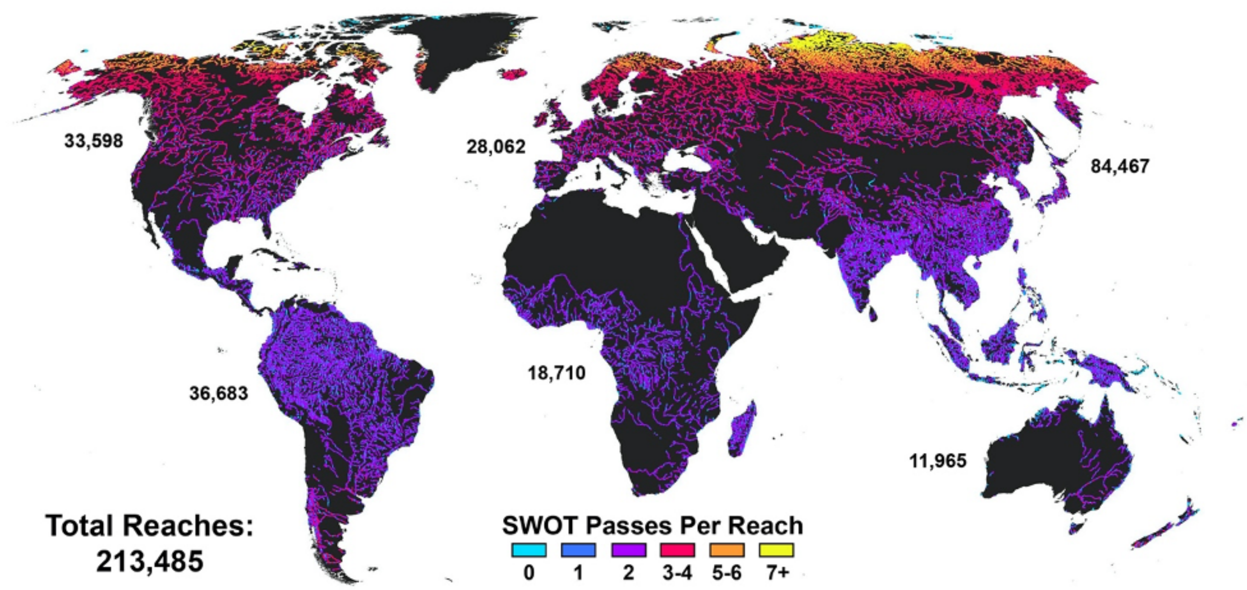

Figure 2. Number of observations per 21 day SWOT cycle for each reach contained in the PRD. Reproduced from [1]. Individual observations for a reach collected during the 21 day SWOT cycle are aggregated to form the river average product.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

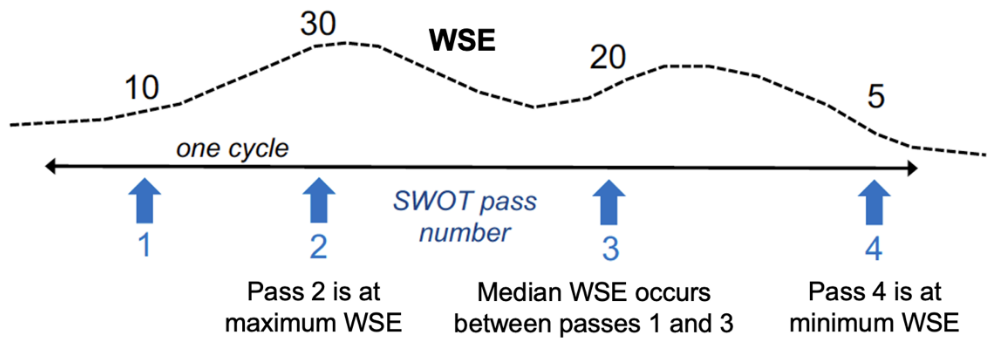

Figure 3. A notional example series of WSE measurements over one SWOT observation cycle. Over the course of one observation cycle, a given reach may be observed from multiple SWOT passes. The L2_HR_RiverAvg product will report the minimum, maximum, and median WSE and the mean WSE over all passes with valid WSE. The mean area, slope, width, and discharge values are computed using the corresponding passes with valid WSE. The area, slope, width, and discharge of the passes corresponding to maximum, minimum, and median WSE will also be reported.

3.6 Spatial Organization

As noted above, reaches are statically defined in the PRD based on hydrological and morphological considerations along each river course. For ease of indexing, each reach has as its first data element a unique *reach_id* attribute (described in Section 4.1). The reaches are sequentially written to the reach shapefile according to the order of the *reach_id* attribute. Information on reach connectivity, including the number and identifiers of upstream and downstream reaches, is available in the PRD.

3.7 Volume

The list of data elements in Section 5 and general assumptions about the number of reaches and the SWOT swath geometry allow an estimate of the expected data volume of the L2_HR_RiverAvg product. The estimated data product volume in Table 3 is based on the following assumptions:

1. There are 60 L2_HR_RiverAvg granules per observation cycle.
2. The number of reaches per granule varies from 83 to 23626 with an average number of 4433 reaches per granule.
3. Reaches are approximately 10 km long. Centerline points for reach .shp files are spaced at 30 meters (assume 350 points/reach). This assumption enables an estimate of the storage required for the centerline itself.
4. To store all data listed in Section 5, a single reach requires 6.3 kilobytes (KB)

With these assumptions, the average data volume of each shapefile is 27.3 MB per cycle.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table 3: Description of the data volume of the L2_HR_RiverAvg product.

Name	Expected Mean Volume / Granule (MB)	Maximum Volume / Granule (MB)
RiverAvg shapefile (all 5 files)	27.3	145.2

© 2023 California Institute of Technology. Government sponsorship acknowledged.

4 Qualitative Description

The L2_HR_RiverAvg vector product is derived from the associated L2_HR_RiverSP product [9], which contain arrays of the measured WSE, area, width, slope, and discharge data.

The files that make up the shapefile format are described in Section 3.2. The format of the .shp file is specified in [3]. The .shp file contains the basic geometry (latitudes and longitudes, defined relative to the parameters of the .prj file) defining the shape type (from the PRD, not the SWOT measurement). The content of the .shp file does not vary with observations for a given instance of the PRD used during processing. Even if a reach is partially observed or is not detected in the SWOT cycle, the .shp file still contains that reach, with fill values in the .dbf file indicating null SWOT measurements (Section 5.1).

Descriptions of the .dbf attributes are given below, with attribute names according to the following conventions:

- Prefix "*p_*" indicates that information is taken from the PRD,
- Suffix "*_f*" indicates a flag,
- Suffix "*_u*" indicates an estimated uncertainty on a corresponding quantity.

Measured or observed (the terms are used interchangeably in this document) values for reach attributes are calculated from the corresponding single-pass reach-level attributes (for a description of the SP reach-level attribute, see the L2_HR_RiverSP product description document [9] with the methods used in its computation described in [10]). Unlike the L2_HR_RiverSP product, the L2_HR_RiverAvg product does not distinguish between "Basic" or "Expert" attributes. The L2_HR_RiverAvg product was designed with simplicity in mind and is intended for global-scale studies. Each record in the .dbf file contains attributes that are conceptually grouped in the subsections below. There is one record for each attribute for each reach in the granule. In particular, note the L2_HR_RiverAvg product is not intended to provide information on how the WSE, width, or slope vary over a given reach.

Note that the distinction of what information comes from each file in a shapefile should be transparent to users who read the L2_HR_RiverAvg shapefiles with GIS software. It is simply provided in this document for completeness.

In addition to the cycle-average attributes, this product also includes three sets of measurements for each attribute and associated uncertainty, which are tied to the WSE statistics. Each of the three sets includes the values of time, WSE, river width, water surface slope, enhanced water surface slope, change in cross-sectional area from a reference value, water surface area, discharge, and gauge-constrained discharge observed. Each set is reported for the pass when the WSE was at its (1) minimum, (2) median, and (3) maximum values during the observation cycle. The WSE statistic represented by each set of measurements is identified in the attribute name by *_hmin*, *_hmed*, or *_hmax*, denoting a measurement taken at the time of the minimum WSE (*_hmin*), median WSE (*_hmed*), or maximum WSE (*_hmax*), respectively.

Typically, only the single-pass reach observations that are indicated as either "good" and/or "suspect" by their corresponding quality flags (see the *reach_q* variable in the L2_HR_RiverSP product) are considered when computing the mean, minimum, median, and maximum values reported in the L2_HR_RiverAvg product. If "suspect" single-pass observations are used, the quality flag in the L2_HR_RiverAvg product is set accordingly (see *reach_q* and *reach_q_b*). If

© 2023 California Institute of Technology. Government sponsorship acknowledged.

there are no "good" or "suspect" single-pass reach observations, then "degraded" single-pass reach observations are used to compute the reported values in the L2_HR_RiverAvg product, and the output quality flags indicate "degraded" results. "Bad" single-pass reach observations are never considered in producing the L2_HR_RiverAvg quantities, but a reach may be flagged as "bad" in the output if no "good," "suspect," or "degraded" single-pass observations are available.

For example, the product reports the reach slope at the minimum, median, and maximum WSE; it does **not** necessarily report the minimum, median, and maximum slope, as these may not coincide with the minimum, median, and maximum WSE given the non-monotonic relationship between WSE and slope.

As a more detailed example, the measurement set and associated uncertainties at the time of the minimum observed WSE (*t_hmin*) in the observation cycle contains: time stamp (*t_hmin, t_tai_hmin, t_str_hmin*), WSE (*wse_hmin, wse_hmin_u*), river width (*wid_hmin, wid_hmin_u*), water surface slope (*slp_hmin, slp_hmin_u*), enhanced water surface slope (*sl2_hmin, sl2_hmin_u*), change in cross-sectional area from a reference value (*dxa_hmin, dxa_hmin_u*), water surface area (*are_hmin, are_hmin_u*), discharge (*dsg_hmin, dsg_hmin_u*), and gauge-constrained discharge (*dsc_hmin, dsc_hmin_u*).

4.1 Reach Identifier (ID)

Each reach record is associated with a unique identifier (*reach_id*) copied from the PRD. The format of the identifier is a 11-character string of the form *CBBBBBRRRRT*, where *C* = continent, *B* = basin, *R* = reach, and *T* = type. The *reach_id* provides the link between the reach and its corresponding entry in the PRD.

- *reach_id*: Unique reach identifier from the prior river database.

The basin definition embedded in the *reach_id* is based on the Pfafstetter coding system [11] that assigns identifications based on the topology of the river network. The code allows digits 0-9 at each hierarchy level. SWOT *reach_id* values always include six Pfafstetter levels of basins including the Continent (C) and five basin levels (BBBBB) in decreasing hierarchical order. Continent (C) and water body type (T) codes are provided in Table 2 and Table 4, respectively. Note that in Table 4, lake water bodies that are connected to the river topology of the PRD ($T = 3$) are processed as river reaches and will appear in the L2_HR_RiverAvg product. These connected lakes (reservoirs, for example) will typically also appear in the L2_HR_LakeAvg product. Disconnected lakes ($T = 4$) are those that are not on the river network as defined by the PRD, which excludes very narrow rivers. Therefore, disconnected lakes may or may not be endorheic.

Table 4. Water body type codes for the *reach_id* attribute.

Type Code (*T*)	Water Body Type
1	River
2	Disconnected Lake (not processed)

© 2023 California Institute of Technology. Government sponsorship acknowledged.

3	Connected Lake
4	Dam
5	No Topology

The continent code (C) is level 1 in the Pfafstetter code. The continent codes and 2-character continent IDs in Table 2 are consistent with the continent coding used in the HydroBASINS product [12]. As indicated in the template above and illustrated in Figure 4, five additional basin levels (*BBBBB*) are used within each continent. Within each basin level, the reach (RRRR) is numbered in the upstream direction beginning with 0001 to a maximum of 9999 (i.e., a zero-padded four-digit number, represented as *RRRR* in the *reach_id*).

Within continents, up to 10 second-level basins are allowed. Figure 4 shows an example of the Pfafstetter decomposition, where the North American continent (Level 1, $C = 7$) is broken into 8 Level 2 basins and, as an example, the Mississippi River Basin (Level 2, $B = 4$) is further broken into 10 Level 3 basins. Each of the Level 3 basins will be further broken down into up to 10 Level 4 subbasins and so on, until Level 6 is reached.

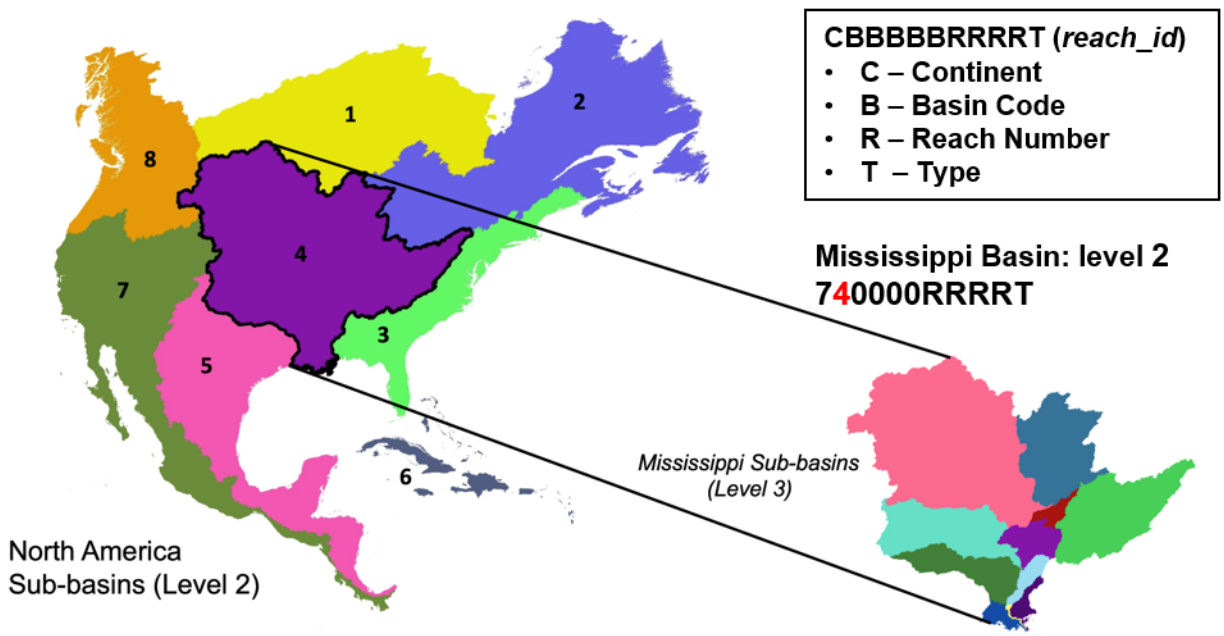

Figure 4. Example of the Pfafstetter basin coding system [11] at Level 2 over North America.

The common name of the river is given in the following variable:

- *river_name* (Basic): English language name of the river, adapted from Open Street Map (OSM) *[13]* river names. The OSM field “name:en” is used to extract the name of the local river in the English language. Prior to attaching to the PRD, OSM data are edited to correct errors. In cases where multiple names are given for a river, these are separated with a forward slash.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

4.2 Time

Numeric time tags for each measurement data record are provided in the UTC and TAI time scales using the attributes beginning with *t_* and *t_tai_* respectively, with a third time tag beginning with *t_str_*, which stores the UTC time as an ISO 8601 date string. A set of three time tags precedes each of the set of measurement statistics. The first set of time tags is associated with the cycle average observations. These give the simple average of the times of the individual pass observations that contribute to *wse_avg* with pass times represented as numbers of seconds since an absolute epoch. The attributes are defined as follows:

- *t_avg*: Average UTC time for all passes with valid water surface elevation in the observation cycle. The value is the average measurement time in seconds in the UTC time scale since 1 Jan 2000 00:00:00 UTC. The metadata field *tai_utc_difference* in the .shp.xml file gives the difference between the TAI and UTC reference time (in seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata *leap_second* is set to the UTC time at which the leap second occurs.
- *t_tai_avg*: Average TAI time for all passes with valid water surface elevation in the observation cycle. The value is the average measurement time in seconds in the TAI time scale since 1 Jan 2000 00:00:00 TAI. This time scale contains no leap seconds. The time difference (in seconds) between TAI and UTC at the time of the first measurement of the dataset is given by the metadata property *tai_utc_difference* of the *t_avg* attribute.
- *t_str_avg:* Average UTC time for all passes with valid water surface elevation in the observation cycle. This value is identical to *t_avg* but is represented as a string in the format YYYY-MM-DDThh:mm:ssZ, where the Z suffix indicates UTC time.

The second set of time tags represents the time instant when the water surface elevation was at its minimum level in the observation cycle. The three variables are named and defined as follows:

- *t_hmin*: UTC time corresponding to the minimum water surface elevation (*hmin*) in the observation cycle. The UTC time and metadata definitions are identical to that of *t_avg*.
- *t_tai_hmin:* TAI time corresponding to the minimum water surface elevation (*hmin*) in the observation cycle. The TAI and metadata definitions are identical to that of *t_tai_avg*.
- *t_str_hmin:* UTC time corresponding to the minimum water surface elevation (*hmin*) in the observation cycle. The UTC time definitions and string formatting are identical to that of *t_str_avg*.

The third set of time tags represents the time instant when the water surface elevation was at its median level in the observation cycle. The three variables are named and defined as follows:

- *t_hmed*: UTC time corresponding to the median water surface elevation (*hmed*) in the observation cycle. The UTC time and metadata definitions are identical to that of *t_avg*.
- *t_tai_hmed:* TAI time corresponding to the median water surface elevation (*hmed*) in the observation cycle. The TAI and metadata definitions are identical to that of *t_tai_avg*.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

- *t_str_hmed:* UTC time corresponding to the median water surface elevation (*hmed*) in the observation cycle. The UTC time definitions and string formatting are identical to that of *t_str_avg*.

The fourth set of time tags represents the time instant when the water surface elevation was at its maximum level in the observation cycle. The three variables are named and defined as follows:

- *t_hmax*: UTC time corresponding to the maximum water surface elevation (*hmax*) in the observation cycle. The UTC time and metadata definitions are identical to that of *t_avg*.
- *t_tai_ hmax:* TAI time corresponding to the maximum water surface elevation (*hmax*) in the observation cycle. The TAI and metadata definitions are identical to that of *t_tai_avg*.
- *t_str_hmax:* UTC time corresponding to the maximum water surface elevation (*hmax*) in the observation cycle. The UTC time definitions and string formatting are identical to that of *t_str_avg*.

The UTC time attributes have a metadata field named *tai_utc_difference*, which represents the difference between TAI and UTC (i.e., total number of leap seconds) at the time of the first measurement record in the product granule. For example, considering the *t_avg* time stamp:

- *t_tai_avg[0] = t_avg[0] + tai_utc_difference*

The above relationship holds true for all measurement records unless an additional leap second occurs within the time span of the product granule. To account for this, the UTC time attributes also have a metadata field named *leap_second* which provides the date when a leap second might have occurred within the time span of the product granule. The UTC time attributes exhibit a jump when a leap second occurs. If no additional leap second occurs within the time span of the product granule, the metadata field *leap_second* of the *time* attribute is set to "0000-00-00T00:00:00Z".

The Table 5 provides some examples of hypothetical values of *t_avg*, *t_tai_avg*, and *tai_utc_difference*. With this approach, the value of *t* has a 1 second regression during a leap second transition, while *t_tai* is continuous. That is, when a positive leap second is inserted, two different instances have the same value for the UTC time attribute, making UTC time attributes non-unique by themselves; the difference between *t_avg* and *t_tai_avg*, or the *tai_utc_difference* and *leap_second* metadata fields, can be used to resolve this.

Table 5. Examples of how UTC and TAI dates relate to *t_avg*, *t_tai_avg*, and the metadata field *tai_utc_difference*.

UTC Date	TAI Date	t_avg	t_tai_avg	tai_utc_difference
January 1, 2000 00:00:00	January 1, 2000 00:00:32	0.0	32.0	32
December 31, 2016 23:59:59	January 1, 2017 00:00:35	536543999.0	536544035.0	36
December 31, 2016 23:59:59.5	January 1, 2017 00:00:35.5	536543999.5	536544035.5	36
December 31, 2016 23:59:60	January 1, 2017 00:00:36	536543999.0	536544036.0	37
January 1, 2017 00:00:00	January 1, 2017 00:00:37	536544000.0	536544037.0	37
January 1, 2017 12:00:00	January 1, 2017 12:00:37	536587200.0	536587237.0	37

© 2023 California Institute of Technology. Government sponsorship acknowledged.

The UTC time corresponding to the numeric *t_avg* attributes are also given as a string attribute (*t_str_avg*) with the following format: YYYY-MM-DDThh:mm:ssZ (with 'Z' suffix to indicate UTC time). The *t_avg* and *t_tai_avg* attributes maintain sub-second precision, but *t_str_avg* is truncated to one-second precision. The same conventions apply to the time stamps associated with *hmin*, *hmed*, *hmax*.

4.3 Location

A single reference location for each reach is provided from the static PRD attributes *p_lat* and *p_lon* (i.e., the reference location is predefined, not a SWOT-measured quantity for a reach):

- *p_lat*: Geodetic latitude of the reach center in degrees, from the PRD. Positive latitude values increase northward from the equator. The latitude is defined with respect to the reference ellipsoid given by the .prj file.
- *p_lon*: Geodetic longitude of the reach center in degrees, from the PRD. The longitude values become more positive to the east and more negative to the west of the Prime Meridian.

4.4 SWOT Overpasses

Three attributes describe the number of SWOT overpasses during the observation cycle. The attribute *n_passes* indicates the number of passes in the cycle for which data were collected and an observation was possible, regardless of whether the WSE measurement was successful or not. This number depends on the location of the reach. Thus, *n_passes* changes in space, but, nominally, it does not change from cycle to cycle (for a given orbit phase and HR mask). The attribute *n_valid_p* describes the number of passes with valid water surface elevation; it is less than or equal to *n_passes*. The attribute *pass_list* provides a list of orbit passes. The attributes related to the number of SWOT overpasses for a given reach in an observation cycle are stored as follows:

- *n_passes:* Number of SWOT overpasses during the observation cycle. The value is nominally the same for each cycle unless the HR mask changes. It does not depend on the validity of the measurements. The value may be lower when data are missing.
- *n_valid_p:* Number of SWOT overpasses used for the computation of the cycle average quantities, which include only valid measurements.
- *pass_list:* Comma-separated list of the valid SWOT overpasses in *n_valid_p*.

4.5 Measured Hydrology Parameters

The L2_HR_RiverAvg product provides four sets of statistics for the hydrological attributes and estimates of their associated measurement uncertainties computed over an observation cycle. The first set represents the mean conditions and uncertainties reported during the observation cycle: WSE (*wse_avg, wse_avg_u*), river width (*wid_avg, wid_avg_u*), water surface slope (*slp_avg, slp_avg_u*), enhanced water surface slope (*sl2_avg, sl2_avg_u*), change in cross-sectional area from a reference value (*dxa_avg_mean, dxa_avg_u*), water surface area (*are_avg, are_avg_u*), discharge (*dsg_avg, dsg_avg_u*), and gauge-constrained discharge (*dsc_avg,*

© 2023 California Institute of Technology. Government sponsorship acknowledged.

dsc_avg _u; see Section 4.6). The second set reports the values of the hydrological attributes measured at the time of minimum WSE recorded within an observation cycle. These values are identified by the suffixes *_hmin* and *_hmin_u* for each attribute and their associated uncertainty. The third and fourth sets provide hydrological attributes measured at the times of the median (*_hmed*) and maximum (*_hmax*) WSE recorded within the cycle. Unless otherwise specified, all uncertainties represent one-sigma or 68th-percentile uncertainty estimates.

The WSE and slope provided in this product are reported with respect to the Earth Gravitational Model 2008 (EGM2008) geoid model [14]. For details on the transformations between ellipsoid-relative height and geoid-relative height, the application of geophysical range corrections, and instrument corrections, see Section 4.1 of the L2_HR_RiverSP PDD. The attributes included in this product use the corresponding reach attributes in the L2_HR_RiverSP product to which such corrections were already applied. The SP reach attributes are computed using SP node quantities as detailed in [10].

A list of hydrology parameters, except for discharge-related attributes, which are described in section 4.6, followed by a short description of each attribute is presented below:

Average set:

- *wse_avg*: Average water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
- *wse_avg_u*: Estimated uncertainty in the average water surface elevation observed in the cycle.
- *wid_avg*: Average width in the observation cycle.
- *wid_avg_u*: Estimated uncertainty in the average width in the observation cycle.
- *slp_avg*: Average slope in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as the WSE. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower. Slopes are dimensionless and reported in meters/meter.
- *slp_avg_u*: Estimated uncertainty in the average slope in the observation cycle.
- *sl2_avg*: Average enhanced slope in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, with the same corrections and geophysical fields applied as the WSE. It is produced using smoothed node WSE in the L2_HR_RiverSP product. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower. Enhanced slopes are also dimensionless and reported in meters/meter
- *sl2_avg_u*: Estimated uncertainty in the average enhanced slope in the observation cycle.
- *dxa_avg*: Average cross-sectional area change in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
- *dxa_avg_u*: Estimated uncertainty in the average cross-sectional area change in the observation cycle.
- *are_avg*: Average water surface area in the observation cycle. This variable is corrected for the potential presence of dark water.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

- *are_avg_u*: Estimated uncertainty in the average water surface area in the observation cycle.

Reach attributes at minimum WSE:

- *wse_hmin*: Minimum water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
- *wse_hmin_u*: Estimated uncertainty in the minimum water surface elevation observed in the cycle.
- *wid_hmin*: Measured width during the pass containing the minimum water surface elevation in the observation cycle.
- *wid_hmin_u*: Estimated width uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
- *slp_hmin*: Slope measured during the pass containing the minimum water surface elevation in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as the WSE. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
- *slp_hmin_u*: Estimated slope uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
- *sl2_hmin*: Enhanced slope measured during the pass containing the minimum water surface elevation in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, produced using smoothed node WSE in the L2_HR_RiverSP product. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
- *sl2_hmin_u*: Estimated enhanced slope uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
- *dxa_hmin*: Cross-sectional area change during the pass containing the minimum water surface elevation in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
- *dxa_hmin_u*: Estimated cross-sectional area change uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
- *are_hmin*: Water surface area measured during the pass containing the minimum water surface elevation in the observation cycle. This variable is corrected for the potential presence of dark water.
- *are_hmin_u*: Estimated water surface area uncertainty during the pass containing the minimum water surface elevation in the observation cycle.

Reach attributes at median WSE:

- *wse_hmed*: Median water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

- *wse_hmed_u*: Estimated uncertainty in the median water surface elevation observed in the cycle.
- *wid_hmed*: Measured width during the pass containing the median water surface elevation in the observation cycle.
- *wid_hmed_u*: Estimated width uncertainty during the pass containing the median water surface elevation in the observation cycle.
- *slp_hmed*: Slope measured during the pass containing the median water surface elevation in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as the WSE. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
- *slp_hmed_u*: Estimated slope uncertainty during the pass containing the median water surface elevation in the observation cycle.
- *sl2_hmed*: Enhanced slope measured during the pass containing the median water surface elevation in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, with the same corrections and geophysical fields applied as the WSE. It is produced using smoothed node WSE in the L2_HR_RiversSP product. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
- *sl2_hmed_u*: Estimated enhanced slope uncertainty during the pass containing the median water surface elevation in the observation cycle.
- *dxa_hmed*: Cross-sectional area change during the pass containing the median water surface elevation in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
- *dxa_hmed_u*: Estimated cross-sectional area change uncertainty during the pass containing the median water surface elevation in the observation cycle.
- *are_hmed*: Water surface area measured during the pass containing the median water surface elevation in the observation cycle. This variable is corrected for the potential presence of dark water.
- *are_hmed_u*: Estimated water surface area uncertainty during the pass containing the median water surface elevation in the observation cycle.

Reach attributes at maximum WSE:

- *wse_hmax*: Maximum water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
- *wse_hmax_u*: Estimated uncertainty in the maximum water surface elevation observed in the cycle.
- *wid_hmax*: Measured width during the pass containing the maximum water surface elevation in the observation cycle.
- *wid_hmax_u*: Estimated width uncertainty during the pass containing the maximum water surface elevation in the observation cycle.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

- *slp_hmax*: Slope measured during the pass containing the maximum water surface elevation in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as the WSE. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
- *slp_hmax_u*: Estimated slope uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
- *sl2_hmax*: Enhanced slope measured during the pass containing the maximum water surface elevation in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, with the same corrections and geophysical fields applied as the WSE. It is produced using smoothed node WSE in the L2_HR_RiverSP product. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
- *sl2_hmax_u*: Estimated enhanced slope uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
- *dxa_hmax*: Cross-sectional area change during the pass containing the maximum water surface elevation in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
- *dxa_hmax_u*: Estimated cross-sectional area change uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
- *are_hmax*: Water surface area measured during the pass containing the maximum water surface elevation in the observation cycle. This variable is corrected for the potential presence of dark water.
- *are_hmax_u*: Estimated water surface area uncertainty during the pass containing the maximum water surface elevation in the observation cycle.

4.6 Discharge

The L2_HR_RiverAvg product includes two types of discharge estimates. The first does not make use of in situ stream gauge measurements and is found in the attributes whose names begin with *dsg_*. The second type takes advantage of historical in situ discharge measurements from gauges, where available, to constrain the estimated discharge. Gauge-constrained discharge estimates can be identified by attribute names starting with *dsc_*. Both the unconstrained discharge estimates and the gauge-constrained discharge estimates are computed from algorithms that estimate flow-law parameters such as river bathymetry and friction coefficients. The flow-law-parameter estimates are based on mass-conserved flow-law inversion [15, 16], data assimilation [17, 18, 19], or big data analysis [20]. The gauge-unconstrained discharge estimates do not use any other ancillary information such as bathymetry or streamflow measured at streamflow gauges. The gauge-constrained discharge estimates are produced using the same methods, but additional, external information is also taken into account, including measurements from streamflow gauges globally, historical streamflow, and in-situ records. The gauge-constrained discharge estimates spread this information across river networks; therefore constrained discharge differs from the unconstrained discharge estimates both in reaches where gauges exist and in other reaches in the same river network. The gauge-constrained estimates will never fail to exist due to a lack of gauges (in a complete absence of gauges, the gauge-

© 2023 California Institute of Technology. Government sponsorship acknowledged.

constrained and unconstrained estimates would be equivalent). The gauge-constrained discharge estimates are recommended as the most accurate estimates available in the L2_HR_RiverAvg product. Users should be advised that these estimates are trained on many streamflow gauges, and therefore their accuracy evaluated against those gauges may not reflect the accuracy in basins where no gauges are available. As with other hydrology parameters, discharge is also reported as an average in the observation cycle and at the timing of the minimum, median, and maximum observed WSE. The fields containing discharge values and their associated uncertainties are:

- *dsg_avg*: Average unconstrained discharge in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
- *dsg_avg_u*: Estimated uncertainty in the average unconstrained discharge in the observation cycle.
- *dsc_avg*: Average gauge-constrained discharge in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
- *dsc_avg_u*: Estimated uncertainty in the average gauge-constrained discharge in the observation cycle.
- *dsg_hmin*: Unconstrained discharge estimated during the pass containing the minimum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
- *dsg_hmin_u*: Estimated unconstrained discharge uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
- *dsc_hmin*: Gauge-constrained discharge estimated during the pass containing the minimum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
- *dsc_hmin_u*: Estimated gauge-constrained discharge uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
- *dsg_hmed*: Unconstrained discharge estimated during the pass containing the median water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
- *dsg_hmed_u*: Estimated unconstrained discharge uncertainty during the pass containing the median water surface elevation in the observation cycle.
- *dsc_hmed*: Gauge-constrained discharge estimated during the pass containing the median water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
- *dsc_hmed_u*: Estimated gauge-constrained discharge uncertainty during the pass containing the median water surface elevation in the observation cycle.
- *dsg_hmax*: Unconstrained discharge estimated during the pass containing the maximum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

- *dsg_hmax_u*: Estimated unconstrained discharge uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
- *dsc_hmax*: Gauge-constrained discharge estimated during the pass containing the maximum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
- *dsc_hmax_u*: Estimated gauge-constrained discharge uncertainty during the pass containing the maximum water surface elevation in the observation cycle.

4.7 Quality Indicators

The product contains one flag indicating conditions that affect data quality. In general, flag values of zero indicate good data.

- *reach_q*: Quality flag with four possible states: 0 denotes a good measurement, 1 denotes a suspect measurement (a significant error is possible), 2 denotes a degraded measurement (a significant error is likely), and 3 denotes a bad measurement (the measurement should be ignored).
- *reach_q_b:* Bitflag that provides details on why the reach quality flag (*reach_q*) is set as it is. See Appendix C for details.

4.8 Geophysical References

The geoid height, from a model, in meters above the reference ellipsoid (defined by the .prj file) is provided for reference. This information enables the user to convert the observed WSE to a different representation. The geoid slope is provided as well.

- *geoid_hght*: Model for geoid height above the reference ellipsoid whose parameters are given in the .prj file. The geoid model is EGM2008 *[14]*. The geoid model includes a correction to refer the value to the mean tide system (i.e., it includes the zero-frequency permanent tide).
- *geoid_slop*: Geoid model slope in the along-stream direction as defined in the PRD, based upon a least-squares linear fit along the reach. The units are m/m. A positive slope means that the downstream geoid model height is lower.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

5 Detailed Content

The L2_HR_RiverAvg product adopts the Esri shapefile format and conventions [3]. The shapefile format stores geospatial data as primitive geometric shapes like points, polylines, and polygons representing locations, rivers, and lakes, respectively. These shapes, together with data attributes that are linked to each shape, create the representation of the geospatial data. In this section a description of the information in the .dbf file is given. This information is also stored in the .shp.xml files of the reach and node shapefiles. The .shp.xml files provide shapefile metadata information similar to what would be provided as global and per-variable attributes in a NetCDF format file. The format of the .shp.xml files are described in Appendix B.

5.1 Shapefile Information

5.1.1 Dimensions

The headers of the .shp and .shx reach and node files give the number records in the shapefiles. However, the .dbf file does not have an entry for the number of records. All attributes in the .dbf file are scalars (each attribute corresponds to only a single integer, floating-point value, or string). However, some string attributes are used to represent multiple values. For example, in the reach.dbf file, the *pass_list* attribute is given as a character string in a comma-separated list of passes when a particular reach was observed.

5.1.2 Attributes

The attributes of the .dbf file are assigned a name and a particular data type. Note that .dbf attributes are all stored as space-separated, formatted ASCII (ANSI) character strings rather than binary data types. Table 6 summarizes the type, field width and fill value for each data type.

Table 6. Attribute data types in shapefile products.

Data Type	Description	fill value
int4	integer (4-character storage)	-999
int9	integer (9-character storage)	-99999999
float	floating point (13-character storage)	-999999999999
text	maximum 254-character storage	“no_data”

The unique, descriptive metadata for each attribute (e.g., expected minimum and maximum values) and the global metadata (e.g., basin number) generally follow the conventions defined for other SWOT products and are given in Table 7 and

Table 8, respectively. Since metadata cannot be stored internal to the .dbf file, the shapefile .shp.xml file provides the metadata fields that apply to each shapefile attribute in the .dbf file. Not all metadata fields will be used for each shapefile attribute (e.g., the metadata field *leap_second* is unique to the time attributes). A description of the .shp.xml file format is given in Appendix B.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table 7. Metadata fields used to describe shapefile attributes.

Item	Description
fill_value	The value used to represent missing or undefined data.
calendar	Reference time calendar.
comment	Miscellaneous information about the attribute, or the methods to generate it.
coordinates	Coordinate variables associated with the attribute.
flag_meanings	The description of the meaning of each of the elements of flag_values.
flag_values	Values of the flag attribute. Used in conjunction with flag_meanings.
flag_masks	Used in conjunction with flag_meanings. Describes a number of independent Boolean conditions using bit field notation by setting unique bits in each flag_masks value. A flagged condition is identified by performing a bitwise AND of the variable value and each flag_masks value; a non-zero result indicates a true condition. Thus, any or all of the flagged conditions may be true, depending on the variable bit settings.
institution	Institution which generates the source data for the attribute, if applicable.
leap_second	UTC time at which a leap second occurs within the time span of the data represented in the attribute.
long_name	A descriptive name that indicates the content of the attribute.
quality_flag	Names of variable quality flag(s) that are associated with this attribute to indicate its quality.
source	Data source (model, author, or instrument).
standard_name	A standard name that indicates the attribute content.
tai_utc_difference	Difference between TAI and UTC reference time.
units	Units of attribute.
valid_max	Maximum theoretical value of the attribute (not necessarily the same as maximum value of actual data)
valid_min	Minimum theoretical value of the attribute (not necessarily the same as minimum value of actual data)
type	Attribute type (int4, int9, float or text)

Table 8. Global metadata fields of the L2_HR_RiverAvg product.

Item	Description
Conventions	Esri conventions as given in 'ESRI Shapefile Technical Description, an ESRI White Paper, July 1998' http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf
title	Level 2 KaRIn High Rate River Average Product
institution	Name of producing agency.
source	The method of production of the original data. If it was model-generated, source should name the model and its version, as specifically as could be useful. If it is observational, source should characterize it (e.g., 'Ka-band radar interferometer').
history	UTC time when file generated. Format is: 'YYYY-MM-DDThh:mm:ssZ : Creation'
platform	SWOT
references	Published or web-based references that describe the data or methods used to product it. Provides version number of software generating product.
reference_document	Name and version of Product Description Document to use as reference for product.
contact	Contact information for producer of product. (e.g., 'ops@jpl.nasa.gov').
cycle_number	Cycle number of the product granule.
continent_id	Two-letter continent identifier of the product granule.
continent_code	One-digit (C) continent code of the product granule.
basin_code	Two-digit (CB) basin code of the product granule (from HydroBASINS).
short_name	L2_HR_RiverAvg
crid	Composite release identifier (CRID) of the data system used to generate this file

© 2023 California Institute of Technology. Government sponsorship acknowledged.

product_version	Version identifier of this data file
pge_name	Name of the product generation executable (PGE) that created this file
pge_version	Version identifier of the product generation executable (PGE) that created this file
time_granule_start	Nominal starting UTC time of product granule. Format is: YYYY-MM-DDThh:mm:ss.ssssssZ
time_granule_end	Nominal ending UTC time of product granule. Format is: YYYY-MM-DDThh:mm:ss.ssssssZ
time_coverage_start	UTC time of first measurement. Format is: YYYY-MM-DDThh:mm:ss.ssssssZ
time_coverage_end	UTC time of last measurement. Format is: YYYY-MM-DDThh:mm:ss.ssssssZ
geospatial_lon_min	Westernmost longitude (deg) of granule bounding box
geospatial_lon_max	Easternmost longitude (deg) of granule bounding box
geospatial_lat_min	Southernmost latitude (deg) of granule bounding box
geospatial_lat_max	Northernmost latitude (deg) of granule bounding box
xref_l2_hr_riversp_files	Names of input Level 2 river single-pass files.

5.2 Attribute Description

Table 9 lists the reach .dbf shapefile attributes (bold left-most column), and their associated metadata fields from Table 7. The attributes are separated into the seven categories listed in Sections 4.1 through 4.7. Appendix B contains a description of the shp.xml format that captures metadata in the product files.

Table 9. Attributes of the reach shapefile of the L2_HR_RiverAvg product.

Attributes		
reach_id		
	type	text
	long_name	reach ID from prior river database
	short_name	reach_id
	coordinates	p_lon p_lat
	comment	Unique reach identifier from the prior river database. The format of the identifier is CBBBBBBRRRRT, where C=continent, B=basin, R=reach, T=type.
p_lat		
	type	float
	fill_value	-999999999999.0
	long_name	latitude of the center of the reach
	standard_name	latitude
	short_name	prior_latitude
	units	degrees_north
	valid_min	-80
	valid_max	80
	comment	Geodetic latitude of the reach center from the prior river database. Positive values increase northward of the equator.
p_lon		
	type	float
	fill_value	-999999999999.0
	long_name	longitude of the center of the reach
	standard_name	longitude
	short_name	prior_longitude
	units	degrees_east
	valid_min	-180

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	valid_max	180
	comment	Geodetic longitude of the reach center from the prior river database. The longitude values become more positive to the east and more negative to the west of the Prime Meridian.
river_name		
	type	text
	fill_value	no_data
	long_name	river name(s)
	short_name	river_name
	comment	English language name(s) of the river from the prior river database, which adapts the name(s) from Open Street Map. If there are multiple names, they are separated by a forward slash.
n_passes		
	type	int4
	fill_value	-999
	long_name	number of passes per cycle
	short_name	num_passes
	units	1
	valid_min	0
	valid_max	99
	comment	Number of expected SWOT overpasses during the observation cycle. The value is the same for each cycle. It does not depend on the validity of the measurements.
n_valid_p		
	type	int4
	fill_value	-999
	long_name	number of valid passes
	short_name	num_valid_passes
	units	1
	valid_min	0
	valid_max	99
	comment	Number of swot overpasses used for the computation of the cycle average quantities.
pass_list		
	type	int4
	fill_value	-999
	long_name	list of valid passes
	short_name	pass_list
	comment	Comma-separated list of passes used in the computation of the cycle average.
t_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average UTC time
	standard_name	time
	short_name	time_avg
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	units	seconds since 2000-01-01 00:00:00.000
	comment	Average UTC time for all passes with valid water surface elevation in the observation cycle. Time of measurement in seconds in the UTC time scale since 1 Jan 2000 00:00:00 UTC. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

t_tai_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average TAI time
	standard_name	time
	short_name	time_tai_avg
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	units	seconds since 2000-01-01 00:00:00.000
	comment	Average TAI time for all passes with valid water surface elevation in the observation cycle. Time of measurement in seconds in the TAI time scale since 1 Jan 2000 00:00:00 TAI. This time scale contains no leap seconds. The time difference (in seconds) between TAI and UTC at the time of the first measurement of the dataset is given by the metadata [tu_avgwse:tai_utc_difference].
t_str_avg		
	type	text
	fill_value	no_data
	long_name	average UTC time
	standard_name	time
	short_name	time_str_avg
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	comment	Average UTC time for all passes with valid water surface elevation in the observation cycle. The format is YYYY-MM-DDThh:mm:ssZ, where the Z suffix indicates UTC time. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
wse_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average water surface elevation
	short_name	wse_avg
	units	m
	quality_flag	reach_q
	valid_min	-1500
	valid_max	150000
	coordinates	p_lon p_lat
	comment	Average water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
wse_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the average water surface elevation
	short_name	wse_avg_uncert
	units	m
	valid_min	0
	valid_max	999999
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average water surface elevation observed in the cycle.
wid_avg		

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	type	float
	fill_value	-999999999999.0
	long_name	average reach width
	short_name	width_avg
	units	m
	quality_flag	reach_q
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Average width in the observation cycle.
wid_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the average reach width
	short_name	width_avg_uncert
	units	m
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average width in the observation cycle.
slp_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average water surface slope with respect to the geoid
	short_name	slope_avg
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Average slope in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
slp_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the average water surface slope
	short_name	slope_avg_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average slope in the observation cycle.
sl2_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average enhanced water surface slope with respect to the geoid
	short_name	slope2_avg
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	coordinates	p_lon p_lat
	comment	Average enhanced slope in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, produced using smoothed node wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
sl2_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the average enhanced water surface slope
	short_name	slope2_avg_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average enhanced slope in the observation cycle.
dxa_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average change in cross-sectional area
	short_name	change_in_cross_sectional_area_avg
	units	m^2
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Average cross-sectional area change in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
dxa_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty of the average change in the cross-sectional area
	short_name	change_in_cross_sectional_area_avg_uncert
	units	m^2
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average cross-sectional area change in the observation cycle.
are_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average water surface area
	short_name	area_avg
	units	m^2
	quality_flag	reach_q
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Average water surface area in the observation cycle. This variable is corrected for the potential presence of dark water.
are_avg_u		
	type	float

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	fill_value	-999999999999.0
	long_name	uncertainty in the average total water surface area
	short_name	area_avg_uncert
	units	m^2
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average water surface area in the observation cycle.
dsg_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average discharge
	short_name	discharge_avg
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Average unconstrained discharge in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
dsg_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	dsg_avg_u
	short_name	discharge_avg_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the average unconstrained discharge in the observation cycle.
dsc_avg		
	type	float
	fill_value	-999999999999.0
	long_name	average gauge-constrained discharge
	short_name	discharge_gauge_constr_avg
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Average gauge-constrained discharge in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
dsc_avg_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in average gauge-constrained discharge
	short_name	discharge_gauge_constr_avg_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	comment	Estimated uncertainty in the average gauge-constrained discharge in the observation cycle.
t_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	UTC time at the minimum water surface elevation
	standard_name	time
	short_name	time_at_min_wse
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	units	seconds since 2000-01-01 00:00:00.000
	coordinates	p_lon p_lat
	comment	UTC time corresponding to the minimum water surface elevation (hmin) in the observation cycle. Time of measurement in seconds in the UTC time scale since 1 Jan 2000 00:00:00 UTC. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
t_tai_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	TAI time at the minimum water surface elevation
	standard_name	time
	short_name	time_tai_at_min_wse
	calendar	gregorian
	units	seconds since 2000-01-01 00:00:00.000
	comment	TAI time corresponding to the minimum water surface elevation (hmin) in the observation cycle. Time of measurement in seconds in the TAI time scale since 1 Jan 2000 00:00:00 TAI. This time scale contains no leap seconds. The time difference (in seconds) between TAI and UTC at the time of the first measurement of the dataset is given by the metadata [tu_minwse:tai_utc_difference].
t_str_hmin		
	type	text
	fill_value	no_data
	long_name	UTC time at the minimum water surface elevation
	standard_name	time
	short_name	time_str_at_min_wse
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	comment	UTC time corresponding to the minimum water surface elevation (hmin) in the observation cycle. The format is YYYY-MM-DDThh:mm:ssZ, where the Z suffix indicates UTC time. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
wse_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	minimum water surface elevation with respect to the geoid
	short_name	wse_min
	units	m
	quality_flag	reach_q

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	valid_min	-1500
	valid_max	150000
	comment	Minimum water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
wse_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the minimum water surface elevation
	short_name	wse_min_uncert
	units	m
	valid_min	0
	valid_max	999999
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the minimum water surface elevation observed in the cycle.
wid_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	reach width at the minimum water surface elevation
	short_name	width_at_min_wse
	units	m
	quality_flag	reach_q
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Measured width during the pass containing the minimum water surface elevation in the observation cycle.
wid_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the reach width at minimum water surface elevation
	short_name	width_at_min_wse_uncert
	units	m
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Estimated width uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
slp_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	water surface slope with respect to the geoid at minimum water surface elevation
	short_name	slope_at_min_wse
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Slope measured during the pass containing the minimum water surface elevation in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

slp_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the water surface slope at minimum water surface elevation
	short_name	slope_at_min_wse_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated slope uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
sl2_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	enhanced water surface slope with respect to the geoid at minimum water surface elevation
	short_name	slope2_at_min_wse
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Enhanced slope measured during the pass containing the minimum water surface elevation in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, produced using smoothed node wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
sl2_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the enhanced water surface slope at minimum water surface elevation
	short_name	slope2_at_min_wse_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated enhanced slope uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
dxa_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	change in cross-sectional area at minimum water surface elevation
	short_name	change_in_cross_sectional_area_at_min_wse
	units	m^2
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Cross-sectional area change during the pass containing the minimum water surface elevation in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
dxa_hmin_u		
	type	float

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	fill_value	-999999999999.0
	long_name	uncertainty of the change in the cross-sectional area at minimum water surface elevation
	short_name	change_in_cross_sectional_area_at_min_wse_uncert
	units	m^2
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated cross-sectional area change uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
are_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	water surface area at minimum water surface elevation
	short_name	area_at_min_wse
	units	m^2
	quality_flag	reach_q
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Water surface area measured during the pass containing the minimum water surface elevation in the observation cycle. This variable is corrected for the potential presence of dark water.
are_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the water surface area at minimum water surface elevation
	short_name	area_at_min_wse_uncert
	units	m^2
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Estimated water surface area uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
dsg_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	discharge at minimum water surface elevation
	short_name	discharge_at_min_wse
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Unconstrained discharge estimated during the pass containing the minimum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
dsg_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the discharge at minimum water surface elevation
	short_name	discharge_at_min_wse_uncert
	units	m^3/s

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated unconstrained discharge uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
dsc_hmin		
	type	float
	fill_value	-999999999999.0
	long_name	gauge-constrained discharge at minimum water surface elevation
	short_name	discharge_gauge_constr_at_min_wse
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Gauge-constrained discharge estimated during the pass containing the minimum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
dsc_hmin_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in gauge-constrained discharge at minimum water surface elevation
	short_name	discharge_gauge_constr_at_min_wse_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated gauge-constrained discharge uncertainty during the pass containing the minimum water surface elevation in the observation cycle.
t_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	UTC time at the median water surface elevation
	standard_name	time
	short_name	time_at_med_wse
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	units	seconds since 2000-01-01 00:00:00.000
	coordinates	p_lon p_lat
	comment	UTC time corresponding to the median water surface elevation (hmed) in the observation cycle. Time of measurement in seconds in the UTC time scale since 1 Jan 2000 00:00:00 UTC. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
t_tai_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	TAI time at the median water surface elevation
	standard_name	time
	short_name	time_tai_at_med_wse

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	calendar	gregorian
	units	seconds since 2000-01-01 00:00:00.000
	comment	TAI time corresponding to the median water surface elevation (hmed) in the observation cycle. Time of measurement in seconds in the TAI time scale since 1 Jan 2000 00:00:00 TAI. This time scale contains no leap seconds. The time difference (in seconds) between TAI and UTC at the time of the first measurement of the dataset is given by the metadata [tu_medwse:tai_utc_difference].
t_str_hmed		
	type	text
	fill_value	no_data
	long_name	UTC time at the median water surface elevation
	standard_name	time
	short_name	time_str_at_med_wse
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	comment	UTC time corresponding to the median water surface elevation (hmin) in the observation cycle. The format is YYYY-MM-DDThh:mm:ssZ, where the Z suffix indicates UTC time. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
wse_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	median water surface elevation with respect to the geoid
	short_name	wse_med
	units	m^3/s
	quality_flag	reach_q
	valid_min	-1500
	valid_max	150000
	coordinates	p_lon p_lat
	comment	Median water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
wse_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the median water surface elevation
	short_name	wse_med_uncert
	units	m^3/s
	valid_min	0
	valid_max	999999
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the median water surface elevation observed in the cycle.
wid_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	reach width at the median water surface elevation
	short_name	width_at_med_wse
	units	m
	quality_flag	reach_q
	valid_min	0

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	valid_max	100000
	coordinates	p_lon p_lat
	comment	Measured width during the pass containing the median water surface elevation in the observation cycle.
wid_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the reach width at median water surface elevation
	short_name	width_at_med_wse_uncert
	units	m
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Estimated width uncertainty during the pass containing the median water surface elevation in the observation cycle.
slp_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	water surface slope with respect to the geoid at median water surface elevation
	short_name	slope_at_med_wse
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Slope measured during the pass containing the median water surface elevation in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
slp_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the water surface slope at median water surface elevation
	short_name	slope_at_med_wse_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated slope uncertainty during the pass containing the median water surface elevation in the observation cycle.
sl2_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	enhanced water surface slope with respect to the geoid at median water surface elevation
	short_name	slope2_at_med_wse
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	comment	Enhanced slope measured during the pass containing the median water surface elevation in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, produced using smoothed node wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
sl2_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the enhanced water surface slope at median water surface elevation
	short_name	slope2_at_med_wse_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated n slope uncertainty during the pass containing the median water surface elevation in the observation cycle.
dxa_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	change in cross-sectional area at median water surface elevation
	short_name	change_in_cross_sectional_area_at_med_wse
	units	m^2
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Cross-sectional area change during the pass containing the median water surface elevation in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
dxa_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty of the change in the cross-sectional area at median water surface elevation
	short_name	change_in_cross_sectional_area_at_med_wse_uncert
	units	m^2
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated cross-sectional area change uncertainty during the pass containing the median water surface elevation in the observation cycle.
are_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	water surface area at median water surface elevation
	short_name	area_at_med_wse
	units	m^2
	quality_flag	reach_q
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	comment	Water surface area measured during the pass containing the median water surface elevation in the observation cycle. This variable is corrected for the potential presence of dark water.
are_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the water surface area at median water surface elevation
	short_name	area_at_med_wse_uncert
	units	m^2
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Estimated water surface area uncertainty during the pass containing the median water surface elevation in the observation cycle.
dsg_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	discharge at median water surface elevation
	short_name	discharge_at_med_wse
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Unconstrained discharge estimated during the pass containing the median water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
dsg_hmed_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the discharge at median water surface elevation
	short_name	discharge_at_med_wse_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated unconstrained discharge uncertainty during the pass containing the median water surface elevation in the observation cycle.
dsc_hmed		
	type	float
	fill_value	-999999999999.0
	long_name	gauge-constrained discharge at median water surface elevation
	short_name	discharge_gauge_constr_at_med_wse
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Gauge-constrained discharge estimated during the pass containing the median water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
dsc_hmed_u		

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in gauge-constrained discharge at median water surface elevation
	short_name	discharge_gauge_constr_at_med_wse_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated gauge-constrained discharge uncertainty during the pass containing the median water surface elevation in the observation cycle.
t_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	UTC time at the maximum water surface elevation
	standard_name	time
	short_name	time_at_max_wse
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	units	seconds since 2000-01-01 00:00:00.000
	coordinates	p_lon p_lat
	comment	UTC time corresponding to the maximum water surface elevation (hmed) in the observation cycle. Time of measurement in seconds in the UTC time scale since 1 Jan 2000 00:00:00 UTC. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
t_tai_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	TAI time at the maximum water surface elevation
	standard_name	time
	short_name	time_tai_at_max_wse
	calendar	gregorian
	units	seconds since 2000-01-01 00:00:00.000
	comment	TAI time corresponding to the maximum water surface elevation (hmed) in the observation cycle. Time of measurement in seconds in the TAI time scale since 1 Jan 2000 00:00:00 TAI. This time scale contains no leap seconds. The time difference (in seconds) between TAI and UTC at the time of the first measurement of the dataset is given by the metadata [tu_medwse:tai_utc_difference].
t_str_hmax		
	type	text
	fill_value	no_data
	long_name	UTC time at the maximum water surface elevation
	standard_name	time
	short_name	time_str_at_max_wse
	calendar	gregorian
	tai_utc_difference	[value of TAI-UTC at time of first record]
	leap_second	YYYY-MM-DDThh:mm:ssZ
	comment	UTC time corresponding to the maximum water surface elevation (hmin) in the observation cycle. The format is YYYY-MM-DDThh:mm:ssZ, where the Z suffix indicates UTC time. [tai_utc_difference] is the difference between TAI and UTC reference time (seconds) for the first measurement of the data set. If a leap second

© 2023 California Institute of Technology. Government sponsorship acknowledged.

		occurs within the data set, the metadata leap_second is set to the UTC time at which the leap second occurs.
wse_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	maximum water surface elevation with respect to the geoid
	short_name	wse_max
	units	m^3/s
	quality_flag	reach_q
	valid_min	-1500
	valid_max	150000
	coordinates	p_lon p_lat
	comment	maximum water surface elevation in the observation cycle relative to the provided model of the geoid, with corrections for media delays (wet and dry troposphere, and ionosphere), crossover correction, and tidal effects applied.
wse_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the maximum water surface elevation
	short_name	wse_max_uncert
	units	m^3/s
	valid_min	0
	valid_max	999999
	coordinates	p_lon p_lat
	comment	Estimated uncertainty in the maximum water surface elevation observed in the cycle.
wid_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	reach width at the maximum water surface elevation
	short_name	width_at_max_wse
	units	m
	quality_flag	reach_q
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Measured width during the pass containing the maximum water surface elevation in the observation cycle.
wid_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the reach width at maximum water surface elevation
	short_name	width_at_max_wse_uncert
	units	m
	valid_min	0
	valid_max	100000
	coordinates	p_lon p_lat
	comment	Estimated width uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
slp_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	water surface slope with respect to the geoid at maximum water surface elevation
	short_name	slope_at_max_wse

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Slope measured during the pass containing the maximum water surface elevation in the observation cycle. The water surface slope is reported with respect to the geoid, and with the same corrections and geophysical fields applied as wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
slp_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the water surface slope at maximum water surface elevation
	short_name	slope_at_max_wse_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated slope uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
sl2_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	enhanced water surface slope with respect to the geoid at maximum water surface elevation
	short_name	slope2_at_max_wse
	units	m/m
	quality_flag	reach_q
	valid_min	-0.001
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Enhanced slope measured during the pass containing the maximum water surface elevation in the observation cycle. The enhanced water surface slope is reported with respect to the geoid, produced using smoothed node wse. The upstream or downstream direction is defined by the prior river database. A positive slope means that the downstream WSE is lower.
sl2_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the enhanced water surface slope at maximum water surface elevation
	short_name	slope2_at_max_wse_uncert
	units	m/m
	valid_min	0
	valid_max	0.1
	coordinates	p_lon p_lat
	comment	Estimated enhanced slope uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
dxa_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	change in cross-sectional area at maximum water surface elevation
	short_name	change_in_cross_sectional_area_at_max_wse

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	units	m^2
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Cross-sectional area change during the pass containing the maximum water surface elevation in the observation cycle. The change in channel cross sectional area is computed with respect to the value reported in the prior river database.
dxa_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty of the change in the cross-sectional area at maximum water surface elevation
	short_name	change_in_cross_sectional_area_at_max_wse_uncert
	units	m^2
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated cross-sectional area change uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
are_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	water surface area at maximum water surface elevation
	short_name	area_at_max_wse
	units	m^2
	quality_flag	reach_q
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Water surface area measured during the pass containing the maximum water surface elevation in the observation cycle. This variable is corrected for the potential presence of dark water.
are_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the water surface area at maximum water surface elevation
	short_name	area_at_max_wse_uncert
	units	m^2
	valid_min	0
	valid_max	2000000000
	coordinates	p_lon p_lat
	comment	Estimated water surface area uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
dsg_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	discharge at maximum water surface elevation
	short_name	discharge_at_max_wse
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000

© 2023 California Institute of Technology. Government sponsorship acknowledged.

	coordinates	p_lon p_lat
	comment	Unconstrained discharge estimated during the pass containing the maximum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms.
dsg_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in the discharge at maximum water surface elevation
	short_name	discharge_at_max_wse_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated unconstrained discharge uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
dsc_hmax		
	type	float
	fill_value	-999999999999.0
	long_name	gauge-constrained discharge at maximum water surface elevation
	short_name	discharge_gauge_constr_at_max_wse
	units	m^3/s
	quality_flag	reach_q
	valid_min	-10000000
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Gauge-constrained discharge estimated during the pass containing the maximum water surface elevation in the observation cycle. The discharge featured in this field is based on a consensus derived from the individual discharge algorithms and constrained by available in-situ observations.
dsc_hmax_u		
	type	float
	fill_value	-999999999999.0
	long_name	uncertainty in gauge-constrained discharge at maximum water surface elevation
	short_name	discharge_gauge_constr_at_max_wse_uncert
	units	m^3/s
	valid_min	0
	valid_max	10000000
	coordinates	p_lon p_lat
	comment	Estimated gauge-constrained discharge uncertainty during the pass containing the maximum water surface elevation in the observation cycle.
reach_q		
	type	int4
	fill_value	-999
	long_name	summary quality indicator for the averaged reaches
	standard_name	status_flag
	short_name	reach_qual
	flag_meanings	good suspect degraded bad
	flag_values	0 1 2 3
	valid_min	0
	valid_max	3
	comment	Summary quality indicator for the reach measurements. A value of 0 indicates a nominal measurement, 1 indicates a suspect measurement, 2 indicates a degraded measurement, and 3 indicates a bad measurement.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

reach_q_b		
	type	int9
	fill_value	-99999999
	long_name	bitwise quality indicator for the averaged reach
	standard_name	status_flag
	short_name	reach_qual_bitwise
	flag_meanings	classification_qual_suspect geolocation_qual_suspect water_fraction_suspect bright_land few_area_observations few_wse_observations far_range_suspect near_range_suspect partially_observed classification_qual_degraded geolocation_qual_degraded no_observations
	flag_masks	2 4 8 128 1024 2048 8192 16384 32768 262144 524288 268435456
	valid_min	0
	valid_max	269282446
	comment	Bitwise quality indicator for the reach measurements. If this word is interpreted as an unsigned integer, a value of 0 indicates good data, values greater than 0 but less than 262144 represent suspect data, values greater than or equal to 262144 but less than 8388608 represent degraded data, and values greater than or equal to 8388608 represent bad data.
geoid_hght		
	type	float
	fill_value	-999999999999.0
	long_name	geoid height
	standard_name	geoid_height_above_reference_ellipsoid
	short_name	geoid_height
	source	EGM2008 (Pavlis et al., 2012)
	institution	GSFC
	units	m
	valid_min	-150
	valid_max	150
	coordinates	p_lon p_lat
	comment	Geoid height above the reference ellipsoid with a correction to refer the value to the mean tide system i.e., includes the permanent tide (zero frequency).
geoid_slop		
	type	float
	fill_value	-999999999999.0
	long_name	geoid slope
	short_name	geoid_slope
	source	EGM2008 (Pavlis et al., 2012)
	institution	GSFC
	units	m/m
	valid_min	-0.001
	valid_max	0.01
	coordinates	p_lon p_lat
	comment	Geoid slope in the along-stream direction, based upon a least-square linear fit along the reach. A positive slope means that the downstream geoid model height is lower.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

6 References

[1] E. H. Altenau, T. Pavelsky, M. T. Durand, X. Yang, R. P. d. M. Frasson and L. Bendezu, "The Surface Water and Ocean Topography (SWOT) mission River Database (SWORD): A global river network for satellite data products," *Water Resources Research,* no. WRCR25408, 2021, https://doi.org/10.1029/2021WR030054.

[2] C. Pottier and C. Stuurman, "Surface Water and Ocean Topography Product Description: Level 2 KaRIn high rate lake average vector product (L2_HR_LakeAvg)," CNES, TN-CDM-0676-CN, 2021.

[3] Esri, "ESRI Shapefile Technical Description, an ESRI White Paper," 1998. [Online]. Available: http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf.

[4] JPL D-79084, Revision A, "SWOT mission performance and error budget," Jet Propulsion Laboratory Internal Document, 2017.

[5] R. P. d. M. Frasson, R. Wei, M. T. Durand, J. T. Minear, A. Domeneghetti, G. Schumann, B. A. Williams, E. Rodriguez, C. Picamilh, C. Lion, T. Pavelsky and P.-A. Garambois, "Automated river reach definition strategies: Application for the Surface Water and Ocean Topography Mission," *Water Resources Research,* vol. 53, no. 10, pp. 8164-8186, 2017, https://doi.org/10.1002/2017WR020887.

[6] A. S. Montazem, P.-A. Garambois, S. Calmant, P. Finaud-Guyot, J. Monnier, D. M. Moreira, J. T. Minear and S. Biancamaria, "Wavelet-based river segmentation using hydraulic control-preserving water surface elevation profile properties," *Geophysical Research Letters,* vol. 46, no. 12, pp. 6534-6543, 2019, https://doi.org/10.1029/2019GL082986.

[7] S. Biancamaria, D. P. Lettenmaier and T. M. Pavelsky, "The SWOT Mission and its capabilities for land hydrology," *Surveys in Geophysics,* vol. 37, no. 2, pp. 307-337, 2016, https://doi.org/10.1007/s10712-015-9346-y.

[8] R. P. d. M. Frasson, G. J.-P. Schumann, A. J. Kettner, G. R. Brakenridge and W. F. Krajewski, "Will the Surface Water and Ocean Topography (SWOT) satellite mission observe floods?," *Geophysical Research Letters,* vol. 46, no. 17-18, pp. 10435-10445, 2019, https://doi.org/10.1029/2019GL084686.

[9] JPL D-56413, "Surface Water and Ocean Topography Product Description: Level 2 KaRIn high rate river single pass vector product (L2_HR_RiverSP)," Jet Propulsion Laboratory Internal Document, 2023.

[10] JPL D-105505, "Surface Water and Ocean Topography Algorithm Theoretical Basis Document: Level 2 KaRIn high rate river single pass product," Jet Propulsion Laboratory Internal Document, 2023.

[11] K. L. Verdin and J. P. Verdin, "A topological system for delineation and codification of the Earth’s river basins," *Journal of Hydrology,* vol. 218, no. 1-2, pp. 1-12, 1999, https://doi.org/10.1016/S0022-1694(99)00011-6.

[12] B. Lehner and G. Grill, "Global river hydrography and network routing: baseline data and new approaches to study the world’s large river systems," *Hydrological Processes,* vol. 27, no. 15, pp. 2171-2186, 2013, https://doi.org/10.1002/hyp.9740.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

[13] Open Street Map, "Copyright and License," 2021. [Online]. Available: https://www.openstreetmap.org/copyright.

[14] N. K. Pavlis, S. A. Holmes, S. C. Kenyon and J. K. Factor, "The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)," *Journal of Geophysical Research: Solid Earth,* vol. 117, no. B4, pp. 1978-2012, 2012, https://doi.org/10.1029/2011JB008916.

[15] C. J. Gleason and M. T. Durand, "Remote sensing of river discharge: A review and a framing for the discipline," *Remote Sensing,* vol. 12, no. 7, p. 1107, 2020, https://doi.org/10.3390/rs12071107.

[16] C. Gleason, P. A. Garambois and M. T. Durand, "Tracking river flows from space," *EOS,* vol. 98, 2017; https://doi.org/10.1029/2017EO078085.

[17] K. Larnier, J. Monnier, P.-A. Garambois and J. Verley, "River discharge and bathymetry estimation from SWOT altimetry measurements," *Inverse Problems in Science and Engineering,* vol. 29, no. 6, pp. 759-789, 2021, https://doi.org/10.1080/17415977.2020.1803858.

[18] K. M. Andreadis, C. Gleason and C. Brinkerhoff, "Constraining the assimilation of SWOT observations with hydraulic geometry relations," *Water Resources Research,* vol. 56, no. 5, 2020, https://doi.org/10.1029/2019WR026611.

[19] H. Oubanas, I. Gejadze, P.-O. Malaterre, M. T. Durand, R. Wei, R. P. d. M. Frasson and A. Domeneghetti, "Discharge estimation in ungauged basins through variational data assimilation: The potential of the SWOT mission," *Water Resources Research,* vol. 54, no. 3, pp. 2405-2423, 2018, https://doi.org/10.1002/2017WR021735.

[20] R. P. d. M. Frasson, M. T. Durand, K. Larnier, C. Gleason, K. M. Andreadis, M. Hagemann, R. Dudley, D. Bjerklie, H. Oubanas, P.-A. Garambois, P.-O. Malaterre, P. Lin, T. M. Pavelsky, J. Monnier, C. B. Brinkerhoff and C. H. David, "Exploring the factors controlling the error characteristics of the Surface Water and Ocean Topography mission discharge estimates," *Water Resources Research,* vol. 57, no. e2020WR028519, 2021, https://doi.org/10.1029/2020WR028519.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Appendix A. **Acronyms**

AD	Applicable Document
API	Application Interface
ATBD	Algorithm Theoretical Basis Document
BAM	Bayesian AMHG-Manning
CNES	Centre National d'Études Spatiales
ECMWF	European Center for Medium-Range Weather Forecasts
GSFC	Goddard Space Flight Center
HiVDI	Hierarchical Variational Discharge Inference
HR	High Rate
JPL	Jet Propulsion Laboratory
KaRIn	Ka-band Radar Interferometer
LR	Low Rate
L2	Level 2
MetroMan	Metropolis-Manning
MOMMA	MOdified Manning Method Algorithm
PIXC	Pixel Cloud
PRD	Prior River Database
RD	Reference Document
SADS	SWOT Assimilated DiScharge
SDS	Science Data System
SP	Single Pass
SWOT	Surface Water Ocean Topography
TAI	International Atomic Time
TBC	To Be Confirmed
TBD	To Be Determined
UTC	Coordinated Universal Time
WSE	Water Surface Elevation
XML	Extensible Markup Language

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Appendix B. **Description of XML**

In the L2_HR_RiverAvg product, the use of the term "attributes" usually follows the shapefile nomenclature in referring to the variables associated with each feature in the .shp file. Other than in this appendix, this term should not be confused with attributes as typically used in the context of netCDF files. Rather, the L2_HR_RiverAvg product uses the term "attributes" in reference to the contents of the .dbf file, and uses the term "metadata" in reference to characteristics of each attribute of the entire shapefile.

However, the Esri shapefile format adopted for L2_HR_RiverAvg product does not have a standard representation for including such metadata. L2_HR_RiverAvg product therefore includes metadata in an extensible markup language (XML) file that is produced alongside each reach shapefile (Section 3.2). That is, for the L2_HR_RiverAvg product, the reach shp.xml (Table 1) files convey the information provided in Table 7 through Table 9 of this document.

These XML files contain metadata about the entire shapefile (the equivalent of "global attributes" in a netCDF file). The global metadata fields are provided in

Table 8. Examples include the starting and ending times of the data contained in the shapefile, and the geospatial bounding box coordinates encompassing the data represented in the shapefile.

These XML files also contain metadata fields, as listed in Table 7, pertaining to specific attributes in the shapefile (the equivalent of per-variable "attributes" in a netCDF file). The reach XML files effectively reproduce the specific metadata fields pertaining to attributes that are provided in Table 9 of this document. Examples include metadata such as the allowable minimum and maximum values of an attribute, and the associated units.

Note, however, that the XML files use the word "attributes" in element names following netCDF conventions to refer to metadata fields, not to variables in the shapefile .dbf file. This mix of nomenclature should be clear in context, as variables and metadata fields are named explicitly in the XML file.

These XML files are organized as follows. Following a standard XML declaration, a single top-level XML element *swot_product* always contains exactly two elements *global_attributes* and *attributes*. The *global_attributes* element gives metadata that apply to the entire shapefile, whereas the *attributes* element gives metadata for each shapefile attribute. Child elements of the *global_attributes* element represent individual global metadata fields, with the metadata values as the XML contents between start- and end-tags that define the name of the global metadata field. The *attributes* element has a child element for each attribute of the corresponding shapefile being described; the start- and end-tags of each of these per-attribute elements correspond to the name of the attribute. Each per-attribute element has child elements that give the metadata fields applicable to that attribute, with the metadata values as the contents between start- and end-tags that define the name of the per-attribute metadata field. Not all attributes are associated with the same set of metadata fields. Children of a given element are always unique. While most metadata values will always be the same across different granules of the L2_HR_RiverAvg product, some fields do vary between granules (e.g., those involving leap seconds).

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Examples are shown below for several XML elements of the .reach.shp.xml file. Note that the XML comments in the example below are included here for descriptive purposes but would not exist in the actual XML file.

```
<swot_product>
    <global_attributes>

<!-- Global metadata listed in
        Table 8 here -->
        <!-- Example entries: -->

        <title>Level 2 KaRIn High Rate River Single Pass Vector
        Product</title>
        <continent>EU</continent>. <!-- From Table 2 -->

        <!-- Other global metadata -->

     <!-- End of global metadata -->


    </global_attributes>

    <attributes>

        <!-- Individual entries for each attribute in Table 9-->
        <!-- Each attribute uses metadata fields from Table 7 -->
        <!-- Example entries for the reach XML file: -->

        <reach_id>
            <type>text</type>
            <long_name> reach ID from prior river database
            </long_name>
            <coordinates>p_lon p_lat</coordinates>
            <comment> Unique reach identifier from the prior river
            database. The format of the identifier is CBBBBBRRRRT,
            where C=continent, B=basin, R=reach, T=type.</comment>
        </reach_id>

        <wse>
            <type>float</type>
            <fill_value>-999999999999</fill_value>
            <long_name> water surface elevation with respect to the
            geoid</long_name>
            <units>m</units>
            <valid_min>-1000</valid_min>
            <valid_max>100000</valid_max>
            <coordinates>p_lon p_lat</coordinates>
            <comment>Fitted reach water surface elevation, relative
            to the provided model of the geoid (geoid_hght), with
```

© 2023 California Institute of Technology. Government sponsorship acknowledged.

```
            corrections for media delays (wet and dry troposphere,
            and ionosphere), crossover correction, and tidal effects
            (solid_tide, load_tidef, and pole_tide) applied</comment>
        </wse>

        <!-- Metadata fields for other attributes in Table 9-->

      <!-- End of attributes from Table 9  -->
    </attributes>
</swot_product>
```

There are a variety of options to display the XML content. For example, many browsers can display XML content directly. Another option is to use XSLT (eXtensible Stylesheet Language Transformations) to transform XML into Hypertext Markup Language (HTML) for a more convenient visualization of the XML content within a browser. To perform this conversion with XLST, there is a tool named “xsltproc” (e.g., http://www.xmlsoft.org/XSLT/xsltproc.html) that can be used to convert the XML files into HTML. For example, to convert the reach XML file on a Linux platform with this tool use the command line:

xsltproc reach.shp.xsl reach.shp.xml > reach.shape.html,

where reach.shp.xsl is an XSLT style sheet of the user’s choosing. An example of a reach.shp.xsl style sheet that a user might choose to use is provided below.

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:variable name="prodtitle" select="swot_product/title"/>
<xsl:template match="swot_product">
  <html>
  <head>
    <title><xsl:value-of select="$prodtitle"/></title>
    <style type='text/css'>
    caption {
      font-weight: bold;
      text-align: center;
    }
    h1 {
      text-align: center;
    }
    th.headcolor {
      background-color: #A9D0F5;
    }
    td.attrcolor {
      background-color: #A9D0F5;
    }
    </style>
  </head>
  <body>
```

```
  <br>
  </br>
  <h1><xsl:value-of select="$prodtitle"/></h1>
  <br>
  </br>
  <xsl:for-each select="global_attributes">
    <table border="1" width="100%" bgcolor="#ffffff" cellspacing="0"
cellpadding="2">
    <caption>Global Metadata of <xsl:value-of
select="$prodtitle"/></caption>
    <tbody>
        <tr>
           <th class="headcolor">Item</th>
           <th class="headcolor">Value</th>
        </tr>
        <xsl:for-each select="*">
          <tr>
            <td>
              <xsl:value-of select="name()"/>
            </td>
            <td>
              <xsl:value-of select="node()"/>
            </td>
          </tr>
        </xsl:for-each>
    </tbody>
    </table>
  </xsl:for-each>
  <br>
  </br>
  <br>
  </br>
  <xsl:for-each select="attributes">
    <table border="1" width="100%" bgcolor="#ffffff" cellspacing="0"
cellpadding
="2">
    <caption>Attributes of <xsl:value-of
select="$prodtitle"/></caption>
    <tbody>
        <xsl:for-each select="*">
          <tr>
            <td colspan="3" class="attrcolor">
              <xsl:value-of select="name()"/>
            </td>
          </tr>
          <xsl:for-each select="*">
          <tr>
            <td width="40">
            </td>
            <td>
```

© 2023 California Institute of Technology. Government sponsorship acknowledged.

```
              <xsl:value-of select="name()"/>
            </td>
            <td>
              <xsl:value-of select="node()"/>
            </td>
          </tr>
          </xsl:for-each>
        </xsl:for-each>
    </tbody>
    </table>
   </xsl:for-each>
  </body>
  </html>
</xsl:template>
</xsl:stylesheet>
```

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Appendix C. **Quality Flag Bit Definitions**

Quality flags in SWOT products are sometimes represented as bit flags such that the information from multiple individual conditions is captured in a single flag variable. This is accomplished by defining the flag variable as an unsigned integer whose bits in a binary (base-2 number system) representation reflect the states (true or false) of the individual conditions captured by the flag.

For example, a bit-flag variable q might capture information from three independent binary conditions C_3, C_2, and C_1, each of which might be true or false, in its three least significant bits (LSBs). The value of the variable q would then give the states of C_3, C_2, and C_1 per the table below:

Table 10. Bit Flag Example

Value of *q*	State of C_3	State of C_2	State of C_1
0	False	False	False
1	False	False	True
2	False	True	False
3	False	True	True
4	True	False	False
5	True	False	True
6	True	True	False
7	True	True	True

Equivalently, the value of the bit-flag variable q is defined mathematically as

$$q = \sum_{k=0}^{n-1} 2^k C_k$$

where n is the number of bits and C_k (whose value is either 0 or 1 to represent the false and true states, respectively) is the condition associated with bit k.

The bit meanings of the *reach_q_b* flag is given in Table 11.

For each row of the table, the decimal and hexadecimal values represent the value of the flag variable if the bit of that row were 1 and all other bits were 0. All of the information in this table is captured by the *flag_masks* and *flag_meanings* metadata fields of a given bit-flag variable. Where no condition is specified in the table, the bit is unassigned (not used) and should never be 1. It is possible that these bits will become assigned in future versions of the product, however. The color shading of the table gives a rough, qualitative indication of how much a nonzero bit value for each row would be expected to reduce confidence in the measurement, with redder hues indicating greater degradation.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

Table 11. Measurement Quality Flag Bit Definitions

Bit (from LSB)	Decimal	Hex	reach_q_b
0	1	1	
1	2	2	classification_qual_suspect
2	4	4	geolocation_qual_suspect
3	8	8	water_fraction_suspect
4	16	10	
5	32	20	
6	64	40	
7	128	80	bright_land
8	256	100	
9	512	200	
10	1024	400	few_area_observations
11	2048	800	few_wse_observations
12	4096	1000	
13	8192	2000	far_range_suspect
14	16384	4000	near_range_suspect
15	32768	8000	partially_observed
16	65536	10000	
17	131072	20000	
18	262144	40000	classification_qual_degraded
19	524288	80000	geolocation_qual_degraded
20	1048576	100000	
21	2097152	200000	
22	4194304	400000	
23	8388608	800000	
24	16777216	1000000	
25	33554432	2000000	
26	67108864	4000000	
27	134217728	8000000	
28	268435456	10000000	no_observations

The meanings of the different conditions specified by Table 11 are described below:

- *classification_qual_suspect*: The measurement includes reach observations with pixels whose land/water classification information is marked suspect in the pixel-cloud inputs to the river processing.
- *geolocation_qual_suspect*: The measurement includes reach observations with pixels whose geolocation information is marked suspect in the pixel-cloud inputs to the river processing.
- *water_fraction_suspect*: The measurement includes reach observations that contain pixels whose water-fraction information is suspiciously large.
- *bright_land*: The measurement uses reach observations that are flagged as bright land and are therefore suspect.

© 2023 California Institute of Technology. Government sponsorship acknowledged.

- *few_area_observations*: The measured reach area is based on at least one reach observation that has very few pixels with valid area.
- *few_wse_observations*: The measured reach WSE is based on at least one reach observation that has very few pixels with valid WSE.
- *far_range_suspect*: The measurement includes at least one node that is located at greater than 60 km cross track.
- *near_range_suspect*: The measurement includes at least one node that is located at less than 10 km cross track.
- *partially_observed*: The measurement includes at least one reach observation in which the fraction of nodes observed is below a threshold.
- *classification_qual_degraded*: The measurement includes reach observations with pixels whose land/water classification information is marked degraded in the pixel-cloud inputs to the river processing.
- *geolocation_qual_degraded*: The measurement includes reach observations with pixels whose geolocation information is marked degraded in the pixel-cloud inputs to the river processing.
- *no_observations:* No valid reach observations are available to average.

© 2023 California Institute of Technology. Government sponsorship acknowledged.