Joint Research Centre

Global Surface Water - Data Users Guide (v2)

Background

This document provides a detailed technical description of the datasets that are being made available as part of the Joint Research Centre's Global Surface Water Dataset. For each of the datasets it provides a purpose, description, bands and symbology so that users can understand each of the datasets and use them efficiently and appropriately. Much of the information in this document is taken directly from the following paper which provides a detailed description of the method that was used to create these datasets:

Jean-Francois Pekel, Andrew Cottam, Noel Gorelick, Alan S. Belward, High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418-422 (2016). (doi:10.1038/nature20584)

The technical descriptions given here relate to both the data that is available in Google Earth Engine and to the data that is available for download from the data access section of the website. The data access section also has ancillary files available which will help when working with the data, such as symbologies, metadata files and other files. The data that are available in Google Earth Engine are available in a number of different multi-band assets: one for the mapped products, one for the water history and one for the metadata datasets.

The datasets that are available are intended to show different facets of the spatial and temporal distribution of surface water over the last 35 years. Some of those datasets are intended to be mapped (e.g. the seasonality layer) and some are intended to show the temporal change at specific locations (i.e. the water history). In addition, there are datasets that provide metadata on the number of observations and valid observations that were used and these can be used to estimate the confidence levels in the data. The following sections describe the map products, water history and metadata datasets.

Map products

Water Occurrence

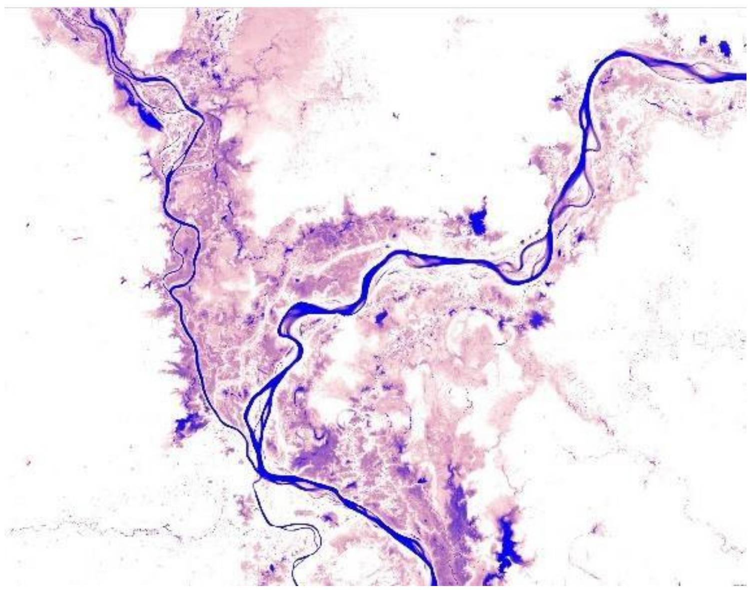
Purpose:

The Water Occurrence shows where surface water occurred between 1984 and 2018 and provides information concerning overall water dynamics. This product captures both the intra and inter-annual variability and changes.

Description (*from the paper*)

The frequency with which water was present on the surface from March 1984 to December 2018 was captured in a single product called surface water occurrence (SWO). To compute SWO, the water detections (WD) and valid observations (VO) from the same months are summed, that is, water detections and valid observations from March 1984 are added to water detections and valid observations from March 1985 and so on, such that SWOmonth = ∑WD month / ∑VO month Averaging the results of all monthly SWOmonth calculations gives the long-term overall surface water occurrence. The month-by-month time step normalizes occurrence against seasonal variation in the number of valid observations across the year. Typically, more cloud-free observations (and thus valid observations) are available during dry seasons than wet. Without monthly weighting, the overall water occurrence (that is, computed over the full period) would be biased by temporal distribution of the valid observations (that is, giving more weight to the dry season than to the wet season).

An example of the Water Occurrence dataset is shown below for Northern Bangladesh. The permanent water surfaces (100% occurrence over 35 years) are represented in blue, and areas where water sometimes occurs are shown in pink through to purple. The paler shades are areas where the water occurs less frequently.



Bands

The occurrence dataset contains the following band:

Name	Data type	Description
occurrence	unsigned int8	The occurrence value expressed as a percentage

In addition, statistics on the overall number of observations and valid observations that were used to compute the occurrence (and the other datasets) are provided in the metadata datasets (see the Metadata datasets section).

Symbology

The occurrence dataset has the following values and symbology. The values from 1 to 100 are discrete.

Value	Symbol	Colour	Label
0		#FFFFFF	Not water
1		#FF0000 (1% opacity)	1% occurrence
100		#0000FF (100% opacity)	100% occurrence
255		#CCCCCC	No data

Occurrence Change Intensity

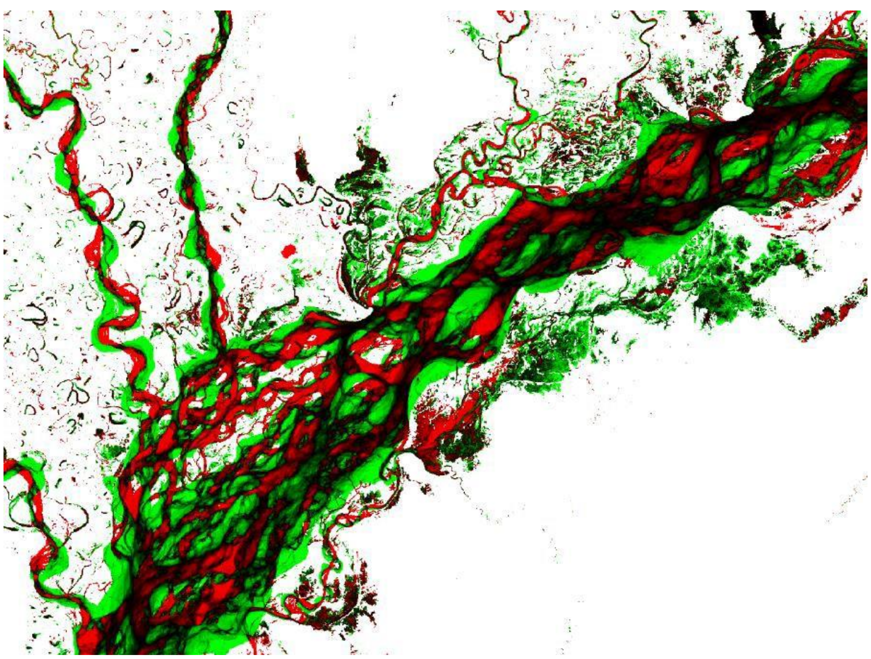
Purpose

The Occurrence Change Intensity map provides information on where surface water occurrence increased, decreased or remained the same between 1984-1999 and 2000-2018. Both the direction of change and its intensity are documented.

Description (*from the paper*)

Change in water occurrence intensity between two epochs (16 March 1984 to 31 December 1999, and 1 January 2000 to 10 December 2018) was also produced. This is derived from homologous pairs of months (that is, the same months contain valid observations in both epochs). The occurrence difference between epochs was computed for each pair and differences between all homologous pairs of months were then averaged to create the surface water occurrence change intensity map. Areas where there are no pairs of homologue months could not be mapped. The averaging of the monthly processing mitigates variations in data distribution over time (that is, both seasonal variation in the distribution of valid observations, temporal depth and frequency of observations through the archive) and provides a consistent estimation of the water occurrence change.

An example of the Occurrence Change Intensity dataset is shown below for the Brahmaputra River. Increases in water occurrence are shown in green and decreases are shown in red. Black areas are those areas where there is no significant change in the water occurrence during the 1984 -2018 period. The intensity of the color represents the degree of change (as a percentage). For example, bright red areas show greater loss of water than light red areas. Some areas appear grey in the maps, these are locations where there is insufficient data to compute meaningful change statistics.



Bands

The Occurrence Change Intensity dataset contains the following bands:

Name	Data type	Description
change_abs	unsigned int8	The absolute difference in the mean occurrence value between the two epochs for homologous months
change_norr	unsigned int8	The normalised difference in the mean occurrence value between the two epochs for homologous months epoch1-epoch2 /epoch1+epoch2

Symbology

The Occurrence Change Intensity dataset has the following values and symbology for the change_norm band. The values from 0 to 200 are discrete.

Value (TIFF)	Value (GEE)	Symbol	Colour	Label
0	-100		#FF0000	-100% loss of occurrence
100	0		#000000	No change
200	100		#00FF00	100% increase in occurrence
253	masked		#FFFFFF	Not water
254	-128		#888888	Unable to calculate a value due to no homologous months
255	127		#CCCCCC	No data

A palette for users with deuteranopia colour-blindness is also available on the Data Access section of the website.

Seasonality

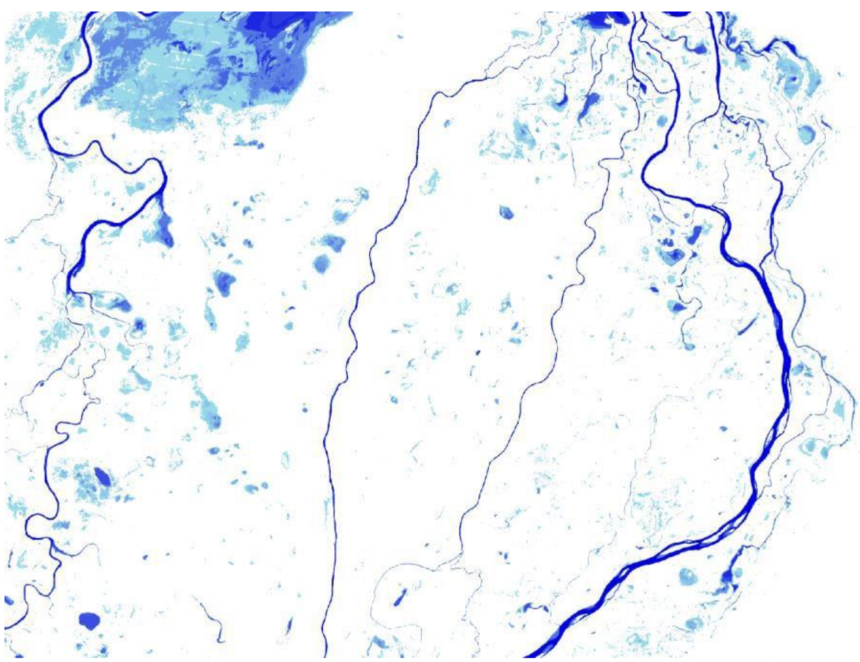
Purpose

The Seasonality map provides information concerning the intra-annual behaviour of water surfaces for a single year (2018) and shows permanent and seasonal water and the number of months water was present.

Description (*from the paper*)

A permanent water surface is underwater throughout the year, while a seasonal water surface is underwater for less than 12 months of the year. In some places we do not have observations for all 12 months of the year (for example, because of the polar night in winter) and in these cases water is considered to be seasonal if the number of months where water is present is less than the number of months where valid observations were acquired. A second consideration is that some lakes freeze for part of the year. However, during the frozen period water is still present under the ice layer, both for lakes and the sea. In the methodology ice is treated as a non-valid observation, so the observation period corresponds only to the unfrozen months. If water is present throughout the observation period (that is, the unfrozen period), the lake is considered to be a permanent water surface. If the area of the lake contracts during the unfrozen period, then the pixels along the borders of the lake no longer represent water, and those pixels will be considered to represent seasonal water surface. Seasonality is computed for every year. A single data set for the contemporary period (2018) is made available.

An example of the seasonality map is shown below for the Mopti area of Mali. The permanent water is represented in dark blue and areas of seasonal water are shown in lighter blue.



Bands

The Seasonality dataset contains the following band:

Name	Data type	Description
seasonality	unsigned int8	The number of months that water was present from January 2018 to December 2018

Symbology

The Seasonality dataset has the following values and symbology. The values from 1 to 12 are discrete.

Value	Symbol	Colour	Label
0		#FFFFFF	Not water
1		#99D9EA	1 month of water
12		#0000AA	12 months of water (permanent water)
255		#CCCCCC	No data

Recurrence

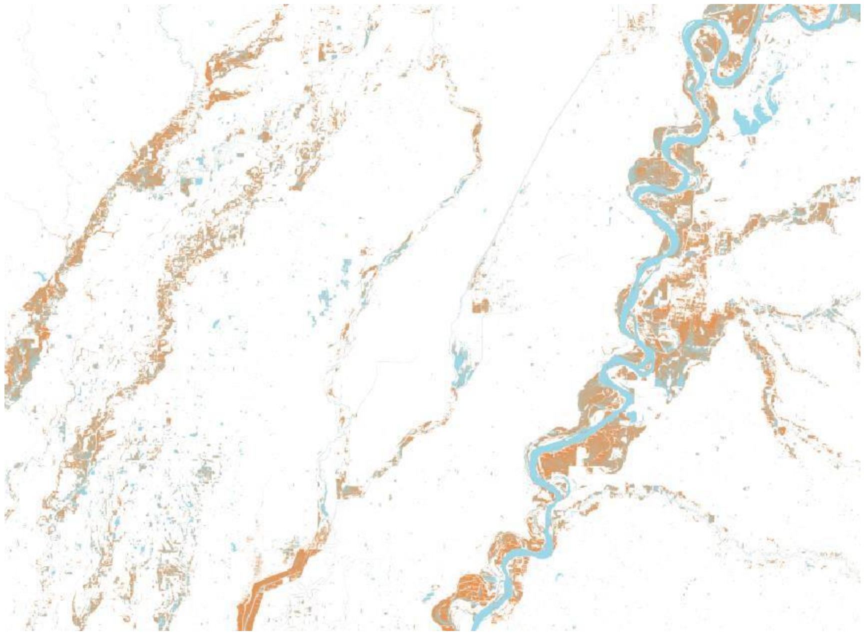
Purpose

The Recurrence map provides information concerning the inter-annual behavior of water surfaces and captures the frequency with which water returns from year to year.

Description (*from the paper*)

Water recurrence is a measurement of the degree of inter-annual variability in the presence of water. This describes how frequently water returned from one year to another (expressed as a percentage). Recurrence refers specifically to the temporal behavior of water surfaces; unlike occurrence, recurrence is not systematically computed over the full span of the archive, because water may not have been present from the beginning to the end of the archive. Thus, we first have to define a 'water period'—that is, that part of the archive where water was present at least from time to time; the recurrence in fact quantifies this 'time to time'. The water period is established individually for each pixel. The water period runs from the first month in the first year in which water is observed to the last month of the last year in which water is observed of the entire 35-year period. In addition to defining the water period we also need to define a 'water season' (not equivalent to a 'wet season'). The water season is identified from the monthly water recurrence and is defined as those months of the year that from time to time have water. A 'water year' is a year with at least one water observation, while an 'observation year' is a year with at least one valid observation within the water season. Water recurrence is then calculated as the ratio of the number of water years to observation years. The count of the number of years starts with the year in which water was first observed and ends with the most recent year in which water was observed. Years that contain only observations outside the water season are not counted; we have no way of knowing whether water might have occurred in the water season because we have no observations.

An example of the Recurrence map is shown below for North-East Arkansas in the USA with areas of water that are inundated regularly (whether seasonal or permanent) shown in blue and areas that are flooded on an episodic basis shown in orange.



Bands

The Recurrence dataset contains the following band:

Name	Data type	Description
recurrenc	unsigned int8	The frequency with which water returns from years to year expressed as a percentage

Symbology

The Recurrence dataset has the following values and symbology. The values from 1 to 100 are discrete.

Value	Symbol	Colour	Label
0		#FFFFFF	Not water
1		#FF7F27	1% recurrence
100		#99D9EA	100% recurrence
255		#CCCCCC	No data

Transitions

Purpose

The Transitions map provides information on the change in seasonality between the first and last years and captures changes between the three classes of not water, seasonal water and permanent water.

Description (*from the paper*)

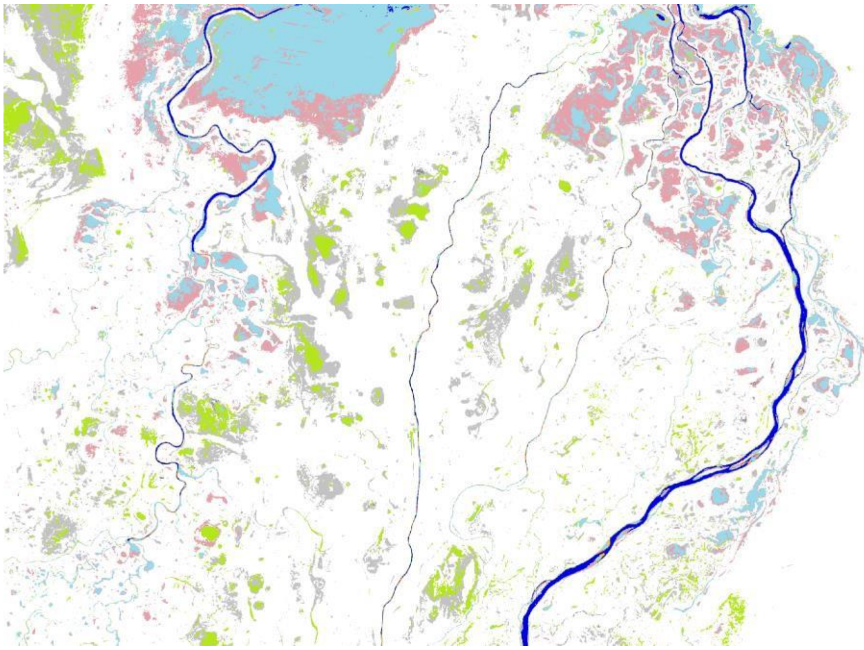
The thematic maps and temporal profiles were used to identify a set of water classes that characterize transitions between the first year in which representative observations were acquired and the last year of observation. Representative years are identified by comparing each year in turn with the annual pattern of monthly recurrence from the temporal profiles. These profiles identify months in which water was observed, and indicate the percentage of valid observations classified as water in any given month. A year is flagged as representative if it contains sufficient valid observations from any combination of months to bring confidence to the determination of the presence or absence of water. The overall level of confidence is determined by the annual sum of the monthly long-term recurrences of observed months (per year). The rationale is that the likelihood of a real absence of water for a year is higher if the water is absent for months showing a high long-term water recurrence than from one showing small rates of recurrence. In the latter case the absence of water may be explained by a seasonal shift, and does not confer enough confidence to conclude that water was not present later.

Therefore, we considered that if the sum of the recurrence of the observed months is greater than 100, the absence of water observation brings enough confidence to consider that water was actually not present. Conversely, a single water presence is enough to demonstrate water presence. The water class in that representative year is then fixed as the 'first' year. The last year's water class is always the class assigned to the last year of observation (2018) because we have enough observations available within a year during this period.

The following transitions were mapped: unchanging permanent water surfaces; new permanent water surfaces (conversion of land into permanent water); lost permanent water surfaces (conversion of permanent water into land); unchanging seasonal water surfaces; new seasonal water surfaces (conversion of land into seasonal water); lost seasonal water surfaces (conversion of a seasonal water into land); conversion of permanent water into seasonal water; and the conversion of seasonal water into permanent water.

These conversions refer to changes in state from the beginning and end of the time series; they do not describe what happened in the intervening years, so an unchanging water surface means that the seasonality at that particular point was the same in the first and last year it was observed, and not necessarily that it was stable throughout. Stability must be checked at the pixel scale by using the long-term water history described by the temporal profiles plus the recurrence and occurrence maps. There are instances where water is not present at the beginning or the end of the observation record but is present in some of the intervening years. By tracking the inter-annual patterns of such 'ephemeral' events and their intra-annual characteristics, each such pixel can be classified as either ephemeral permanent water (land replaced by permanent water that subsequently disappears) or ephemeral seasonal water (land replaced by seasonal water that subsequently disappears), depending on the majority of the observed seasonality during the period of water presence.

An example of the transitions map is shown below for the Mopti region of Mali.



Bands

The transitions dataset has the following band:

Name	Data type	Description
transition	unsigned int8	The type of transition between the first and last year

Symbology

The Transitions dataset has the following values and symbology. The values from 1 to 10 are discrete.

Value	Symbol	Colour	Label
0		#FFFFFF	Not water
1		#0000FF	Permanent
2		#22B14C	New permanent
3		#D1102D	Lost permanent
4		#99D9EA	Seasonal
5		#B5E61D	New seasonal
6		#E6A1AA	Lost seasonal
7		#FF7F27	Seasonal to permanent
8		#FFC90E	Permanent to seasonal
9		#7F7F7F	Ephemeral permanent
10		#C3C3C3	Ephemeral seasonal
255		#CCCCCC	No data

Maximum water extent

Purpose/Description

The Maximum Water Extent provides information on all the locations ever detected as water over the 35-year period. It is the union of all of the other datasets. An example is shown below:

Bands

The maximum water exent dataset has the following band:

Name	Data type	Description
max_extent	unsigned int8	A simple flag indicating if water was detected or not

Symbology

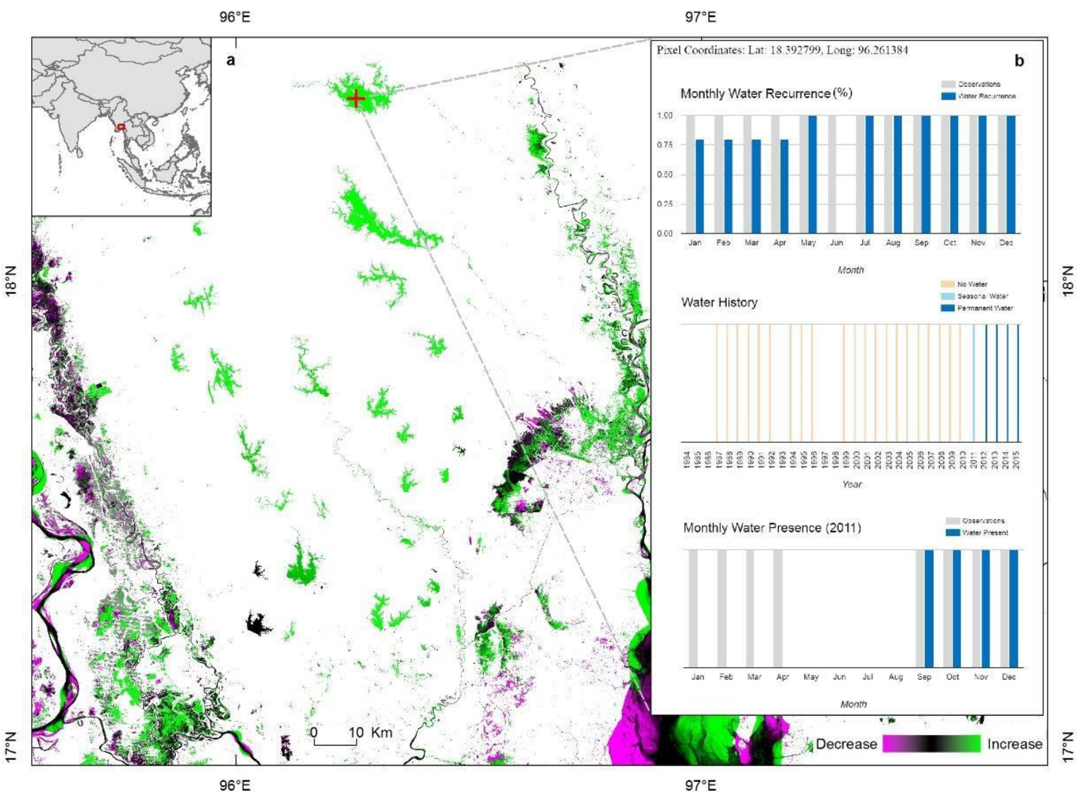
The Transitions dataset has the following values and symbology. The values from 1 to 10 are discrete.

Value	Symbol	Colour	Label
0		#FFFFFF	Not water
1		# 6666FF	Water detected
255		#CCCCCC	No data

Water History

To help understand how water is distributed temporally throughout the 1984-2018 period and throughout the year, there are three datasets that capture the temporal distribution of water: the monthly recurrence, yearly water history and the monthly water history. These histories are shown in the Global Surface Water Explorer and users can click anywhere to see the full water history of a particular place. The history datasets are currently accessible in Google Earth Engine but will be made available for download in the near future.

An example of the three history datasets from the website are given below (from Extended Data Figure 6 in the paper).



a Examples of increasing surface water occurrence in Myanmar.

b Pixel based temporal profiles showing recurrence by month over 35 years (top), water history, by seasonality class, by year over 35 years (middle) and monthly water presence for each year in the water seasonality record, in this case 2007 (bottom). Collectively the graphs show that at this location (Lat: 18.3928º, Long; 96.2633º) there are no observations in 1984-'86, '93, '97 or '98 (the gaps in the middle graph), that prior to 2011 this was dry land, that the dam formed in 2011 and this point was flooded sometime between April and September (bottom graph), but since then it has been permanent water (centre graph), and that in the 35 years of observation water has not been detected in June (no observations have been made in June since the dam filled

Monthly Recurrence

Purpose
The Monthly Recurrence dataset provides information on the intra-annual distribution of the water and characterizes seasonality. Its main purpose is to see how water distribution varies over the year for a particular place.

Description
The monthly recurrence is calculated in the same way as the annual occurrence but at the monthly level. It is therefore a measure of how often water occurs between years for a particular month. There are a total of 12 datasets for the monthly recurrence (one for each month) and each dataset has the same bands and symbology.

The bands include a flag to indicate if the month in question has observations or not. Without this flag it would not be clear from the data if there was no monthly recurrence whether this was because there was no water detected or simply because there were no observations for that month. This may be the case in high latitudes for example, because of the polar night in winter when there are no observations.

Bands
The monthly recurrence dataset is available in Google Earth Engine as an image collection with 12 images. Each image contains the following bands:

Name	Data type	Description
monthly_recurrence	signed int8	The recurrence value expressed as a percentage for this month
has_observations	signed int8	A flag to indicate if the month has observations

Each image also has a metadata property to indicate the month of the year.

Yearly History

Purpose

The Yearly History provides information on the seasonality of the water over the 35-year period. It contains the same information as the Seasonality dataset but for every year where observations are available within the period 1984 to 2018.

Description

The description of a single year within the yearly seasonality classification is the same as for the seasonality dataset.

Bands

The yearly history dataset is available in Google Earth Engine as an image collection with 35 images. Each image contains the following band:

Name	Data type	Description
waterClas	unsigned int8	Classification of the seasonality of water throughout the year.

The values are as follows:

Value	Description
0	No observations
1	Not water
2	Seasonal water
3	Permanent water

Monthly Water History

Purpose

The water detection history provides information on all the water detections at the monthly level from March 1984 to December 2018. Its main purpose is to see when particular events occurred, for example, to see when a dam was constructed in a particular place.

Description

The water detection history is the dataset from which many of the other datasets are derived (e.g. the yearly seasonality classifications) and provides information on whether or not water was detected in a particular month.

Bands

The monthly water history dataset is available in Google Earth Engine as an image collection with 380 images. Each image contains the following band:

Name	Data type	Description
water	int	Water detection for the month

The values are as follows:

Value	Description
0	No observations
1	Not water
2	Water detected

Each image also has metadata properties to indicate the month and the year.

Metadata datasets

The Global Surface Water datasets also provide some 'metadata' information which are statistics on the overall number of water detections, number of observations and valid observations that are present in the 1984-2018 period. These statistics can be used as a proxy measure of confidence as the precision of the metrics at any location improves as the number of valid observations increases.

Bands

The occurrence metadata dataset is available in Google Earth Engine and has the following bands:

Name	Data type	Description
detections	signed int16	The total number of water detections
valid_obs	signed int16	The total number of valid observations
total_obs	signed int16	The total number of observations (i.e. scenes)

Using Symbology Files

To use the *.qml file in QGIS 2.16.3:

- Double click on the layer to open the Layer properties dialog
- At the bottom of the dialog, click Style | Load Style…
- Navigate to the location of the *.qml file and click Open

To use the *.lyr file in ArcGIS 10.4:

- Double click on the layer to open the Layer properties dialog
- Click on the Symbology tab and change the rendered to Unique Values
- Click 'Yes' when prompted to build an attribute table
- Click on the Import Symbology button in the top right
- Navigate to the location of the *.lyr file and click Add

Procedures for earlier versions of the software are similar. For more information see the online documentation for QGIS here and for ArcGIS here.

Using the Metadata Files

The metadata files contain minimal information on the associated datasets. Much more information on the method, derivation of data and technical specification is given in the Nature paper, this user guide and the FAQ on the website.

To use the *.xml metadatafiles in QGIS 2.16.3:

- Install the Metatools Plugin
- Select the dataset you want to import the metadata into
- Click Plugins | Metatools | Import Metadata
- Select the *.xml metadata file

To use the *.xml metadatafiles in ArcGIS 10.4:

- Open ArcCatalog and click on the dataset that you want to import metadata into
- Click on the Description tab and then Import
- In the Source Metadata box, enter the location of the *.xml metadata file
- Set the Import Type as FROM_ARCGIS and click OK

Using Web Map Tiled Services (WMTS)

These mapping services provide the best option for simply displaying the Global Surface Water datasets as they use the cached images that are used in the Global Surface Water Explorer. If you do not need to analyse the water data but simply map it, then you should use the WMTS services as they will be the fastest to load and use and you will not need to download any *.TIFF files. To use these services in Desktop GIS:

In ArcGIS for Desktop:

- In the ArcCatalog Window, click on GIS Servers and then double click on Add WMTS Servers
- In the URL box, enter: https://storage.googleapis.com/global-surface-water/downloads_ancillary/WMTS_Global_Surface_WaterV2.xml and click OK
- Expand the 'Global Surface Water on storage.googleapis.com' item and drag layers onto the map

In QGIS:

- In the Manage Layers toolbar, click on Add WMS/WMTS Layer
- Click New and enter a name (Global Surface Water) and URL (https://storage.googleapis.com/global-surface-water/downloads_ancillary/WMTS_Global_Surface_WaterV2.xml)
- Click OK and click Connect - select a layer to add to the map

If you have any errors due to SSL security certificates then click Ignore and Save.

License

All data here is produced under the Copernicus Programme and is provided free of charge, without restriction of use. For the full license information see the Copernicus Regulation.

Publications, models and data products that make use of these datasets must include proper acknowledgement, including citing datasets and the journal article as in the following citation.

Citation

Jean-Francois Pekel, Andrew Cottam, Noel Gorelick, Alan S. Belward, High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418-422 (2016). (doi:10.1038/nature20584)

If you are using the data as a layer in a published map, please include the following attribution text: 'Source: EC JRC/Google'

Contact

For more information and feedback please contact:

jrc-surfacewater@ec.europa.eu