Comparison of visible and multi-satellite global inundation datasets at high-spatial resolution

Filipe Aires, Catherine Prigent, Etienne Fluet-Chouinard, Dai Yamazaki, Fabrice Papa, Bernhard Lehner

Comparison of visible and multi-satellite global inundation datasets at high-spatial resolution

Filipe Aires[a,b,c,*], Catherine Prigent[a,b,c], Etienne Fluet-Chouinard[d], Dai Yamazaki[e], Fabrice Papa[f,g], Bernhard Lehner[h]

[a] *LERMA / Observatoire de Paris, UPMC, CNRS, Paris France*
[b] *Estellus, Paris, France*
[c] *Water Center / Columbia University, New York, USA*
[d] *Center for Limnology, University of Wisconsin-Madison, Madison, WI, USA*
[e] *Department of Integrated Climate Change Projection Research, Japan Agency for Marine-Earth Science and Technology, Yokohama, 236-0001, Japan*
[f] *LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France*
[g] *Indo-French Cell for Water Sciences, IRD-IISc-NIO-IITM Joint Intern. Lab., IIS, India*
[h] *Department of Geography, McGill University, Montreal, QC, Canada*

Abstract

Several new satellite-derived and long-term surface water datasets at high-spatial resolution have recently become available at the global scale, showing different characteristics and abilities. They are either based on visible imagery from Landsat - the Global 3-second Water Body Map (G3WBM) and the Global Surface Water Explorer (GSWE) - or on the merging of passive/active microwave and visible observations - Global Inundation Extent from Multi-Satellite (GIEMS-D3) - that has been downscaled from a native resolution of 25 km×25 km to the 90 m×90 m resolution. The objective of this paper is to perform a thorough comparison of the different water surface

*Corresponding author
Email address: filipe.aires@obspm.fr (Filipe Aires)
URL: https://vm-wordpress-lerma01.obspm.fr/faires/ (Filipe Aires)

estimates in order to identify the advantages and disadvantages of the two approaches and propose a strategy for future developments of high-resolution surface water databases. Results show that due to their very high spatial resolution (30 m) the Landsat-based datasets are well suited to retrieve open water surfaces, even at very small size. GIEMS-D3 has a better ability to detect water under vegetation and during the cloudy season, and it shows larger seasonal dynamics. However, its current version overestimates surface water extent on water-saturated soils, and due to its low original (i.e. before downscaling) spatial resolution, it is under-performing at detecting small water bodies. The permanent waters for G3WBM, GSWE, GIEMS-D3 and GLWD represent respectively: 2.76, 2.05, 3.28, and 3.04 million km^2. The transitory waters shows larger discrepancies: 0.48, 3.72, 10.39 and 8.81 million km^2.

Synthetic Aperture Radar (SAR) data (from ENVIronment SATellite (ENVISAT), Sentinel and soon the Surface Water Ocean Topography (SWOT)) would be a good complementary information because they have a high nominal spatial resolution and are less sensitive to clouds than visible measurements. However, global SAR datasets are still not available due to difficulties in developing a retrieval scheme adequate at the global scale. In order to improve our estimates of global wetland extents at high resolution and over long-term records, three interim lines of action are proposed: (1) extend the temporal record of GIEMS-D3 to exploit the full time series of microwave observations (from 1978 to present), (2) develop an approach to fuse the GSWE and GIEMS-D3 datasets leveraging the strengths of both, and (3) prepare for the release of SAR global datasets.

Keywords: Wetlands and Inundation, Remote sensing, Landsat, Passive microwaves

2010 MSC: 00-01, 99-00

1. Introduction

The distribution and dynamics of surface water, i.e. permanently and temporarily inundated areas including lakes, rivers, and wetlands, are important because of their interaction with climate, ecology and human wellbeing. For instance, nearly 30% of global methane emissions (Bousquet et al., 2006) originate from wetland areas, risk management responds to inundation patterns (Winsemius et al., 2015), and food security and rice paddy cultivation relies, in certain regions of the world, on surface waters. In return, surface water ecosystems are affected by human activity, land use, hydrologic alterations, and climate change. The complex feedback mechanisms between surface water and climate are difficult to assess and can potentially exacerbate the sensitivity and vulnerability of these regions to changes in precipitation, evapotranspiration, and flow regimes (Gleick, 1989; Chahine, 1992) putting lakes, rivers, and wetlands at risk of rapid deterioration in quantity and quality. Among the many topics about wetlands and climate change, sea level rise and carbon sequestration are major issues. But wetlands are also threatened by land use change and invasive species.

Global distribution and dynamics of surface waters at high-spatial resolution (around 100 m) are still not available, in particular over densely vegetated areas, to satisfy all the needs of the large community of potential users including hydrologists, water and disaster managers, or climate

scientists. Indeed, the global, long-term, frequent, and high-resolution characterisation of all surface water types is beyond the capabilities of current satellite observations.

Visible satellite observations are a primary candidate for the detection of surface waters from space. Moderate Resolution Imaging Spectroradiometer (MODIS) observations have been used to derive global products every two days (http://oas.gsfc.nasa.gov/floodmap/) but visible/infrared observations suffer from the presence of clouds (about 70% of Earth surface at any time) (Wilson and Jetz, 2016) and vegetation. Despite the limitations from vegetation canopy and cloud cover, this type of data is of great value to the community to detect open water. Yamazaki et al. (2015) introduced the Global 3 arc-second Water Body Map (G3WBM) at a pixel resolution of 3 arc-seconds (approximately 90 m at the equator) based on Landsat imagery. This dataset exploits multi-temporal acquisitions in order to distinguish permanent from temporal open water areas. However, no full dynamics of the wetland map are provided. Other datasets have been built from Landsat imagery: (Feng et al., 2014) is global for the year 2000, (Mueller et al., 2016) focused on Australia, (Tulbure et al., 2016) created a three decade dataset over a semi-arid region, and (Verpoorter et al., 2014) mapped an inventory of global lakes. Pekel et al. (2016) recently produced a new Global Surface Water Explorer (GSWE) dataset also from Landsat imagery but using the full 32-year record, allowing for a better description of the trends of surface waters and their occurrence.

Synthetic Aperture Radar (SAR) data has the potential to retrieve surface waters at high-spatial resolution ($\simeq$10 m) as well as capture sub-canopy inun-

dation (L-band) as demonstrated by Santoro et al. (2010) using ENVISAT-ASAR, or more recently using the Sentinel 1 mission (Pham-Duc et al., 2017). Although existing SAR retrievals from a number of sensors cover a large extent of the globe, their use for mapping surface inundation has been protracted due to the local calibration needed for accuracy. The past or current availability of the data has not yet allowed for producing a full global high-spatial resolution surface water dataset from SAR data, although such initiatives have been suggested in the past, e.g. (Westerhoff et al., 2013). There is clearly a need to invest more time in retrieval algorithms and potentially perform data fusion in order to obtain a global, long-term, reliable, and high-resolution dataset of water extent from this type of observations.

The NASA/CNES Surface Water and Ocean Topography (SWOT) mission, planned for launch in 2021, is specifically designed to provide high-spatial resolution ($\simeq$ 10 m) and good temporal sampling (22 days repeat) of the extent (and altitude) of continental surface waters (Prigent et al., 2016; Biancamaria et al., 2016) thanks to an interferometric Ka-band radar (Rodriguez, 2015). Although the SWOT data is expected to deliver a new generation of global water surface extents at unprecedented quality and resolution, the availability of this product is still years in the future. Meanwhile, alternative efforts should be pursued to provide the community with the best possible information about the spatial and temporal variations of global surface water extents. Such efforts would also allow for the extension of the SWOT temporal record backward in time, with existing past imagery; this will be a crucial step in assembling multi-decadal measurements of surface water variation.

A possible approach in this direction it to use the synergy from multiple satellite observations. Following this idea, the Global Inundation Extent from Multi-Satellites (GIEMS) database has been developed through a retrieval scheme that combines satellite observations in the visible, near-infrared, and passive/active microwaves (Prigent et al., 2007, 2012; Papa et al., 2010). GIEMS provides a monthly-mean water surface extent at a low spatial resolution ($0.25^\circ \times 0.25^\circ$ equal-area grid) over a 15-year period (1993-2007). In order to obtain global inundation estimate at high resolution, downscaling techniques have been developed on GIEMS (Aires et al., 2013, 2014). In particular, Fluet-Chouinard et al. (2015) exploited topographic and hydrographic information derived from the Shuttle Radar Topography Mission (SRTM) and trained on a global land cover map to produce a 15 arc-second ($\sim$500 m) resolution map of the minimum and maximum inundation extents at global scale (GIEMS-D15) (http://www.estellus.fr/index.php?static13/giems-d15). In (Aires et al., 2017), an evolution of this downscaling methodology was proposed to obtain a global and dynamic inundation dataset GIEMS-D3, at even higher spatial resolution of 3 arc-second ($\sim$90 m), over 15 years with a monthly time step.

Given the variety of remote sensing approaches, a cross comparison of existing products is needed to explore differences and combined uses of the resulting data. The objective of this paper is to compare and contrast two Landsat-based products (G3WBM and GSWE) and one multi-satellite-based product (GIEMS-D3). Section 2 presents the databases used in this work. A global-scale comparison of the three inundation datasets is performed in section 3, and section 4 presents regional comparisons. To evaluate inherent

uncertainties, we contrast the differences among surface water databases with tree and cloud cover data. The advantages and disadvantages of each type of data are then investigated. Finally, section 5 summarises the conclusions of this comparative study. We discuss ways forward to improve estimates of global, high-spatial resolution extents and long-term dynamics of surface waters of multiple types.

As the descriptions of original data sources are not always using identical vocabulary, this paper explicitly defines the following terminology: The expression "inundation" refers to all surfaces that are detected as water pixels in the original datasets, including lakes, rivers, temporarily inundated land, but also (by error or not) saturated soils. The expression "transient" is used throughout this paper for non-permanent inundation, including temporal, seasonal, intermittent, and ephemeral inundation, as well as spatio-temporal transitions such as moving river meanders or newly constructed reservoirs. The expression "wetland" is used only in descriptive terms without specific definition and may include all types of inundated areas (such as lakes, reservoirs, and rivers) but also wet soils or non-inundated areas such as peatlands.

2. Datasets

2.1. G3WBM

The Global 3 arc-second Water Body Map (G3WBM) uses an automated algorithm to process multi-temporal Landsat images from the Global Land Survey (GLS) database (Yamazaki et al., 2015). Over 33,000 scenes were used from four GLS snapshots at 5-year intervals between 1990 and 2010 in order to delineate a seamless water body map, without cloud and ice/snow gaps.

Permanent water bodies were distinguished from transitory water-covered areas by calculating the frequency of water body existence from overlapping, multi-temporal Landsat scenes. By analysing the frequency of water body existence at 3 arc-second resolution, the G3WBM separates river channels and floodplains (http://hydro.iis.u-tokyo.ac.jp/ yamadai/G3WBM/). Despite distinguishing between permanent and transitory water surfaces, only permanent waterbodies are believed to be comprehensively mapped, while not all transitory waterbodies are captured by the four used used scenes (Yamazaki et al., 2015). The seven different G3WBM surface types are described in Tab. 1 and classes 1-4 represent transitory types.

2.2. GSWE

The Global Surface Water Occurrence (GSWE) dataset (Pekel et al., 2016) uses three million Landsat satellite images to quantify inundation over 32 years (from 1984 to 2015) at a 30 m spatial resolution. The GSWE dataset is freely available at: https://global-surface-water.appspot.com/. Each Landsat pixel has been classified as open water, land, or non-valid observation using an expert system. Open water is defined as any feature of water larger than 30 m $\times$ 30 m open to the sky, including fresh and saltwater. Validation against Google Earth images (i.e. from satellite datasets covering mostly the last three years from visible instruments such as Landsat) and labelling from experts shows that the classifier produces less than 1% of false water detections, and misses less than 5% of water. About 52% of the maximum inundation extent (over the 32 years) is found above 44°N. In 2015, permanent bodies of water represented about 2.78 million km^2, with 86% of that area present throughout the 32-year period (Pekel et al., 2016). A rigorous

Class	**G3WBM** (Yamazaki et al., 2015)	**GLWD** (Lehner and Döll, 2004)
1	Snow	Lake
2	Wet Soil, Wet Vegetation, Lava	Reservoir
3	Salt Marsh	River
4	Temporally Flooded Area	Freshwater Marsh, Floodplain
5	Permanent Water	Swamp Forest, Flooded Forest
6	Permanent Water (from SWBD)	Coastal Wetland
7	Ocean (external land/sea mask)	Pan, Brackish/Saline Wetland
8		Bog, Fen Mire (Peatland)
9		Intermittent Wetland/Lake
10		50–100% Wetland
11		25–50% Wetland
12		0–25% Wetland

Table 1: The seven surface classes of the G3WBM dataset, and the twelve classes of GLWD (classes 1-3 corresponds to open water, 4-12 to wetlands).

validation has been performed based on over 40,000 Landsat-derived control points spanning all three Landsats operational life spans dating back to 1985.

When using visible observations, instantaneous estimates can be contaminated by error, for instance due to the presence of clouds (only clear pixels are used to compute the occurrence, but this cloud detection can be erroneous). Other potential errors and issues need also to be covered, i.e. omission errors, algorithmic issues, and permanent water that was not present at all times during the 32-year period, such as newly built reservoirs or river meanders that have moved. Therefore, a threshold in occurrence of inundation is needed in order to distinguish permanent from transitory pixels. The GSWO (Global Surface Water Occurence) product provides the surface open water occurrence "O(surface water)" over the 32-year Landsat record. This occurrence was obtained using only cloud-free pixels. This clear-sky sampling aliasing increases the importance of the dry seasons and decrease that of the wet seasons. To alleviate this difficult, the occurrence in GSWE was normalised by month of the year. Note that the sampling during the day should still exclude boreal winter in the GSWO occurrence. *In the following, we will refer to GSWO for the occurence product of the GSWE dataset.*

For the purpose of conducting baseline comparisons as conducted in this paper, we use:

- The threshold of O(surface water) $\geq 90\%$ for permanent waters.
- and O(surface water) $< 90\%$ to identify transitory waters over the full 32-year record in order to accommodate some uncertainties.

This definition of permanency differs significantly from (Pekel et al., 2016)

who produced permanent water estimates only for shorter time periods (permanent waters are defined at the yearly scale, not at the 32-year scale). The 0.9 threshold appears as a good pragmatic compromise: With a too low threshold, seasonal waters would be classified as permanent, and with a too high threshold, erroneous classifications would suppress truly permanent pixels. Erroneous classifications can be related for instance to omission errors, vegetation and algorithmic issues, and to include permanent water that was not present at all times during the 32 year periods, such as newly built reservoirs or river meanders that have moved. Note that by using a 90% threshold on the occurence, if a dam is constructed for less time than the 10% of the full record, it will not be a permanent water body but a transitory. If we had chosen a 100% threshold, then even less such structures would be kept as permanent waters. It should be noted however that the selection of this threshold, which only represents our best guess in the absence of conclusive and consistent evidence, will affect the results presented in this study. However, the maximum water extent (permanent plus transitory) is independent of this threshold. Other threshold or no threshold could be considered instead.

2.3. GIEMS and GIEMS-D3 databases

To minimise limitations and uncertainties related to measurements by individual instruments, a multi-sensor technique has been developed to estimate surface water extent and dynamics at global scale (Prigent et al., 2007, 2012; Papa et al., 2010). The method exploits the complementary sensitivities of different satellite observations to surface characteristics (e.g., water, vegetation, soil). Passive microwave observations are particularly sensitive

to the presence of surface water, even under vegetation canopy. However, additional observations have to be used to subtract the contribution of confounding factors such as vegetation from the signal and to avoid confusion with other surface types such as dry sand. The following satellite observations were used to generate GIEMS: (1) passive microwaves from the Special Sensor Microwave/ Imager (SSM/I) measurements between 19 and 85 GHz; (2) active microwave backscattering coefficients at 5.25 GHz from scatterometers; and (3) visible and near-infrared reflectances and the derived NDVI (Normalized Difference Vegetation Index). The methodology is described in details in (Prigent et al., 2001). The use of multiple satellite sources is a true challenge, e.g. it is required to obtain a well calibrated dataset for each individual source. But it was shown that the GIEMS approach is able to benefit from the synergy of these satellite sources to obtain a general algorithm able to perform well in most environments. GIEMS data have been used for modelling of surface water dynamics and biogeochemical fluxes and have been thoroughly evaluated. Note that GIEMS estimates include all surface waters such as rivers, floodplains or lakes indiscriminately. The inundation is expressed as the fractional inundation within each 773 km^2 pixel of an equal-area grid with 0.25° resolution at the equator. Regional quality assessments of the GIEMS database using SAR data indicate that the approach captures realistically complex wetlands. However, it can underestimate small surface waters comprising less than 10% fractional coverage of a grid cell ($\leq$80 km^2) due to its coarse spatial resolution and it can also overestimate large surface waters comprising more than 90% of fractional coverage due to water-saturated soils (Prigent et al., 2012). GIEMS is available at

(http://lerma.obspm.fr/spip.php?article91lang=en).

Downscaling methods have recently been developed to reduce the spatial resolution of GIEMS estimates from 25 km to 500 m (15 arc-second, GIEMS-D15) (Fluet-Chouinard et al., 2015) (http://www.estellus.fr/index.php?static13/giems-d15) and 90 m (3 arc-second, GIEMS-D3) (Aires et al., 2017). The methodology uses a floodability index which is predicted for each pixel based on topography and hydrography information from the HydroSHEDS database (Lehner et al., 2008). This floodability index at high-spatial resolution is used to distribute the water fraction from coarse resolution into the high-resolution pixels. A smoothing procedure is applied during the downscaling to reduce transition artefacts at the edges of the low-resolution boxes from GIEMS. GIEMS-D3 has been assessed by analysing its spatial and temporal variability, and evaluated by comparisons to other independent satellite observations (Aires et al., 2017). Topographic information is useful in predicting natural inundation where hydrology is controlled by elevation, but is more limited in human-modified areas (e.g. artificial reservoirs or rice paddies). The probability of surface water "P(surface water)" is defined in GIEMS-D3 as the temporal recurrence (in percent per high-resolution pixel) of observed water in the available monthly time series (1993-2007).

2.4. Auxiliary datasets

2.4.1. GLWD

The Global Lakes and Wetlands Database (GLWD) represents a comprehensive dataset of global surface water area, including small and large lakes, reservoirs, rivers, and wetlands (Lehner and Döll, 2004). GLWD was generated through a compilation and assimilation of existing analog and digital

maps and cartographic products, and due to the historic perspective of the majority of data included, it is assumed to represent maximum open water and wetland extents. In comparison to water and wetland extents from land cover maps, it is the most extensive water mask of its kind (Nakaegawa, 2012). The "level 3" dataset of GLWD that is used here provides a global 30 arc-second resolution grid describing twelve different surface water types (see Tab. 1). GLWD is static and offers little information on seasonality.

2.4.2. Tree density

In order to assess the surface water retrieval from visible observations when vegetation is present, the global percent tree cover map by (Hansen et al., 2009) is used. The MODIS vegetation continuous fields algorithm and a supervised regression algorithm were applied to estimate the percent tree cover over 500 m pixels. Results show that MODIS data yield greater spatial detail in characterising the tree cover compared to past efforts using for instance AVHRR data. Furthermore, validation efforts have shown a reasonable agreement between the MODIS-estimated and the observed tree cover over validation sites.

2.4.3. Cloud fraction

A cloud fraction dataset can be used over land to investigate where visible observations (from Landsat) can actually retrieve surface parameters or not. Cloud cover dynamics are captured over a large extent yet at fine spatial grain twice-daily by the MODIS satellite images. A near-global, fine-resolution ($\simeq$1 km) monthly cloud frequency dataset is presented in (Wilson and Jetz, 2016) (http://www.earthenv.org/cloud), spanning a temporal

range over 15 years.

Note that the tree density and the cloud fraction datasets are derived from the same instrument, MODIS, but their processing are independent from each other.

3. Global comparison

3.1. Surface water extents and probabilities

Fig. 1 presents a comparison of the G3WBM and GSWE Landsat-derived datasets, along with the GIEMS-D3 multi-satellite estimate, at the global scale at 90 m spatial resolution. The three estimates share the major hydrological structures, with large surface water extents in Canada, Bangladesh, or Eastern Argentina.

The two Landsat-derived products provide similar results over the boreal region with permanent surface water in large areas in Canada, over Scandinavia and in Siberia. The surface water occurrence (for GSWO) and surface water probability (GIEMS-D3) (as defined in the data section) in these regions is generally lower with GIEMS-D3. The differences can be partly related to the reference time periods to calculate the inundation statistics. Landsat only provides information under daylight conditions. As explained in the data section, the GSWO dataset provides statistics done at the monthly level so each month has the same weight in the overall statistics, independent of the number of valid observations in each month. However, during the boreal winter, no observation is available in the Landsat products due to lack of daylight which does not allow for the calculation of inundation occurrence. On the other hand, the GIEMS-D3 product is essentially based

on microwave observations that are available regardless of light conditions. The reference time period for the inundation probability includes the winter months when snow and ice cover the boreal region (this is not considered to be inundated in GIEMS-D3). As a consequence, the surface water probability is systematically lower with GIEMS-D3 than the GSWO occurrence in these regions.

Compared to the GIEMS-D3 estimates, the Landsat-derived water estimates are more uniformly spread, with small surface water detected in most environments all over the globe. The GIEMS-D3 data is based on the low spatial resolution GIEMS product which has known difficulties in detecting water surfaces that cover less than 10% of the original $0.25^\circ \times 0.25^\circ$ pixels. The downscaling methodology partly compensates for these omissions through fusion with GLWD data (Aires et al., 2017), yet in general these omissions are propagated into the higher resolution version.

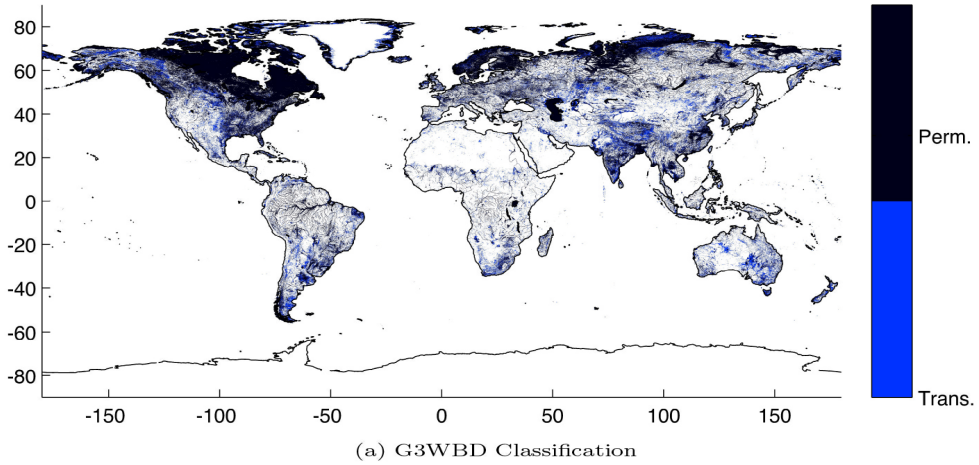
In contrast, GIEMS-D3 detects larger areas of surface water than the Landsat estimates in equatorial forests, e.g. around the Amazon or the Congo Rivers. The microwave observations used to derive the GIEMS product can partly penetrate through the dense canopy whereas the Landsat measurements are blocked by the vegetation.

In Australia and in South Africa, significant differences are observed between the similar Landsat products and GIEMS-D3. More inundation is detected in South Africa with Landsat than with GIEMS-D3, while in the north of the Western Territories in Australia the opposite prevails with larger inundation detected by GIEMS-D3. Significant differences are also observed between the G3WBM and the GSWO products around Lake Mackay. This is

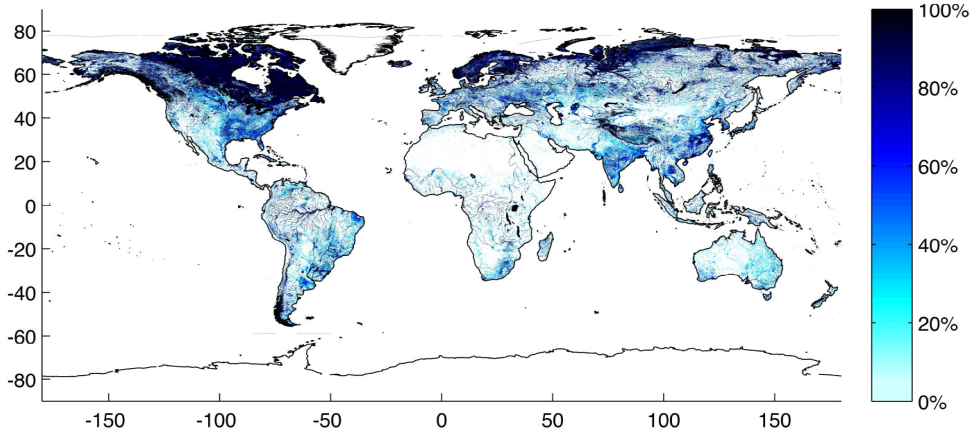
likely related to the difference in the processing record for these two products (32 years for GSWO and only four observations for G3WBM).

3.2. Permanent and transitory surface water

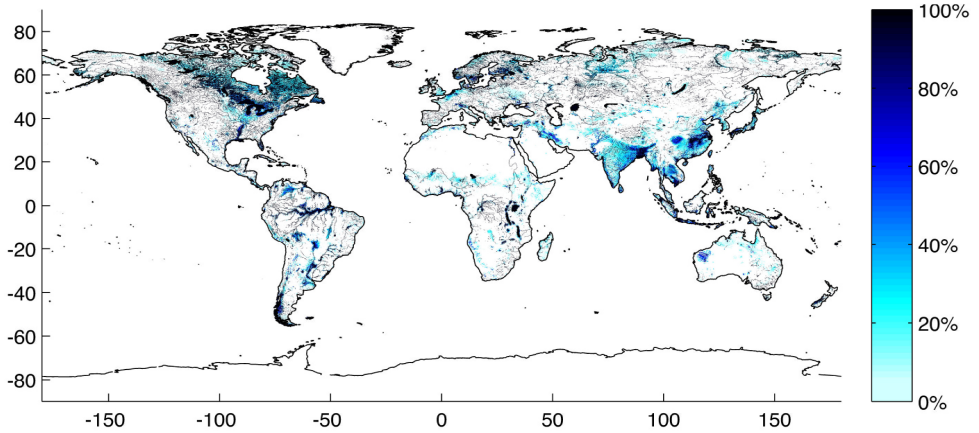
For further comparisons, Fig. 2 shows only the permanent water bodies for G3WBM, GSWO (O(surface water) > 0.9), and GIEMS-D3 (P(surface water) $= 1$), along with the difference between GSWO and GIEMS-D3. To improve the visual interpretation, the fraction of permanent inundation is aggregated into each $0.1° \times 0.1°$ pixel. Furthermore, the same HydroSHEDS land/sea mask has been applied to all datasets to avoid discrepancies along the marine coastlines and to exclude the Caspian Sea from all calculations. Overall, the three maps show generally good spatial agreement, with a total permanent surface water extent of 2.76 million km^2 for G3WBM, 2.05 million km^2 for GSWO, and 3.28 million km^2 for GIEMS-D3 (Tab. 2). The estimate of GSWO increases to 2.78 million km^2 when using a single year (i.e. 2015) and an O-threshold of 100% instead of the 32-year record with an O-threshold of 90%. All results are reasonably close to the reference data of GLWD (3.04 million km^2). It can be seen in the difference map (D) that in the tropical regions, GIEMS has higher water surfaces (blue colour) probably because it has the ability to better detect water through the presence of clouds or vegetation. Conversely, in higher latitudes, GSWO tends to show more water surfaces (red colour), probably due to spatial resolution limitations in GIEMS-D3. These hypotheses will be further discussed in the following. It should be noted, however, that the interpretation of permanent water is not identical in the three datasets as GIEMS-D3 includes inundation that is not visible from the sky (i.e. beneath vegetation).



(a) G3WBD Classification



(b) GSWO Occurrence



(c) GIEMS-D3 Probability

Figure 1: (a) G3WBM classification into permanent and transitory inundation (0-1), (b) the GSWO inundation occurrence (0-100%), and (c) the GIEMS-D3 inundation probability (0-100%). The rendering of this figure is difficult at very high spatial resolution (90 m), the maximum water extents are respectively: 3.74, 5.77 and 13.67 million km^2 (Tab. 2).

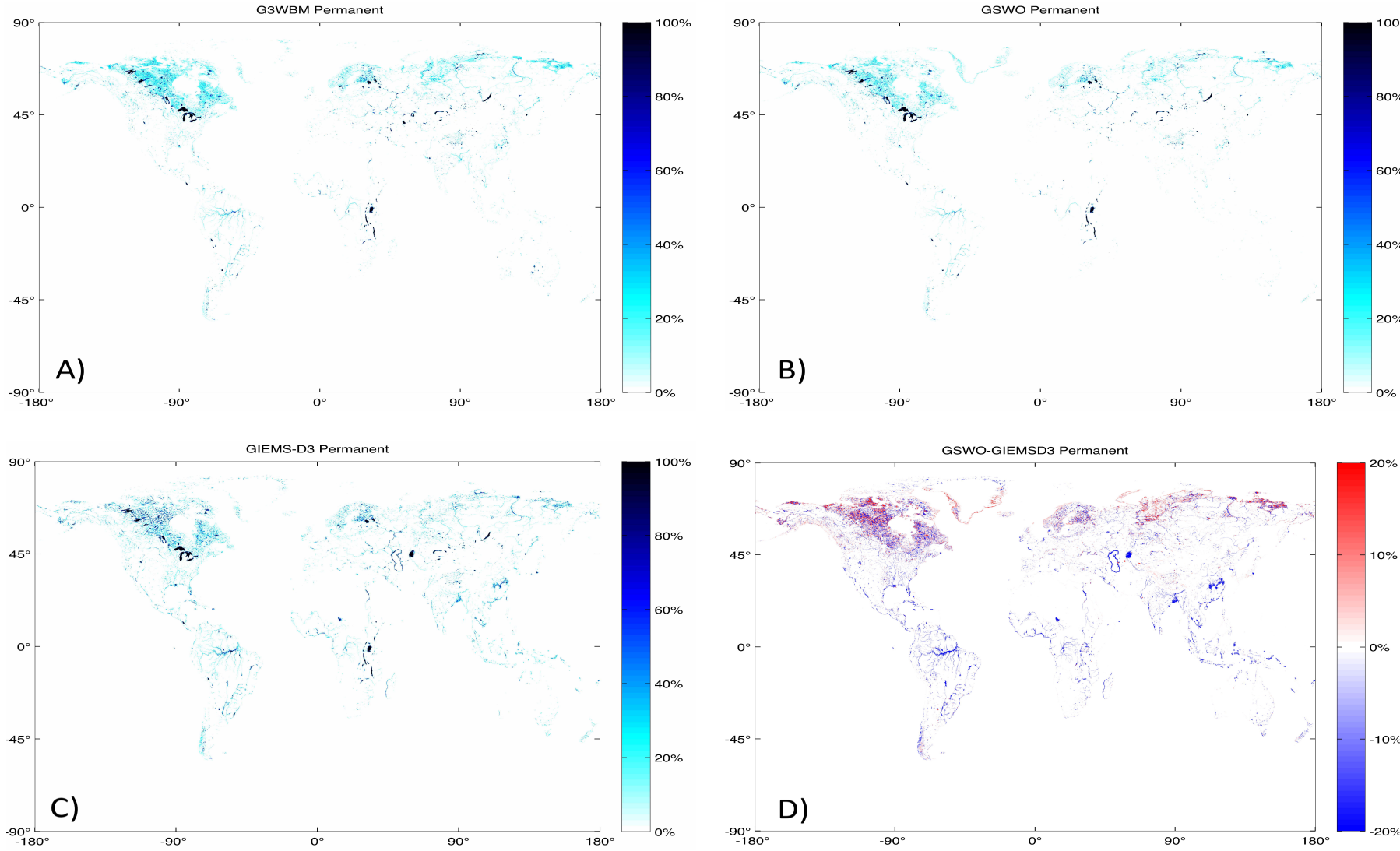


Figure 2: Area-fraction (in %) of permanent inundation states over 0.1°× 0.1° pixels for G3WBM (A), GSWO (B), and GIEMS-D3 (D). The last figure represents the area differences (in %) of GSWO minus GIEMS-D3 (D).

		G3WBM	GSWE		GIEMS-D3	GLWD
Reference		(Yamazaki et al., 2015)	(Pekel et al., 2016)		(Aires et al., 2017)	(Lehner and Döll, 2004)
Observations		Landsat	Landsat		downscaled / multi-sensor	compiled maps & charts
Period		1990, 1995, 2000, 2001	1984-2015	2015	1993-2007	Pre-1990s
Resolution		3 arc-sec	1 arc-sec		3 arc-sec	30 arc-sec
Detect		Open water bodies	Open water bodies		All surface inundation	Waterbodies & natural wetlands
Product(s)		Surface type (see Tab. 1)	Occurence, water occurrence change intensity, seasonality, annual water recurrence, water transitions, maximum water extent		Inundation probability, 15-years time record of monthly binary inundation state	Surface type (see Tab. 1)
Validation		Visual quality control Comparison to National Hydrography Dataset	40,000 Landsat control points dating back to 1985 for the 3 Landsats		Comparison to other inst. (SAR, MODIS, Landsat) & other products (River height, river discharge)	Against documented databases
Permanent	Open-water	2.76	2.05	2.78	3.28	3.04
	Inundation					
Transitory	Open-water	0.48	3.72	0.81	10.39	8.81
	Natural wetlands					
	Inundation (w. rice)					
Total		3.24	5.77		13.67	11.85

Table 2: Estimates of permanent and transitory water extents (in million km^2) according to the three investigated datasets, and GLWD for reference. For GSWE, results are provided for the full 32-year record (using $O > 90\%$ as threshold for permanent waters) and for 2015 only (using $O = 100\%$). Permanent water of GLWD is derived from classes 1-3 only (see Tab. 1). Some numbers are different from the original papers due to inland-water/ocean classification or because of the effects of the threshold selected here. Areas

Fig. 3 presents the transitory water bodies: classes 1-4 from G3WBM, O(surface water)$\leq$0.9 for GSWO, and P(surface water)$<$1 for GIEMS-D3. Here, much larger differences can be observed between the three products. The total transitory surface water is 0.48 million km^2 for G3WBM, 3.72 million km^2 for GSWO, and 10.39 million km^2 for GIEMS-D3, compared to 8.81 million km^2 for GLWD (Tab. 2). Despite the common use of Landsat imagery, G3WBM detects much less transitory surface water than GSWO, which can be attributed both to differences in the observation period and the detection algorithms. The more extensive temporal coverage of GSWO is more likely to capture seasonal variations than the fewer composite snapshots used for G3WBM.

Disagreements between GSWO and GIEMS-D3 transitory water bodies have different origins, depending on the environment. First, all over the globe, small transitory water bodies are detected by GSWO and are likely missed by GIEMS-D3 due to the low detection limit of the original GIEMS data. Second, in Asia, GSWO detects much less transitory water surfaces than GIEMS-D3. Local analyses in the Ganges-Brahmaputra region (Papa et al., 2015) or in the Mekong Delta (see estimates from Sakamoto et al. (2007)) show that GIEMS tends to over-estimate the surface water in areas with saturated soils and irrigated rice culture. However, in these regions, visible and near-infrared observations can also have difficulties in detecting and correctly classifying mixed surfaces with some vegetation coverage (Sakamoto et al., 2007; Crétaux et al., 2016). Hence in the GSWO dataset, some inundated pixels that are partly covered by vegetation may not be detected. In addition, these regions are under very persistent cloud cover, especially dur-

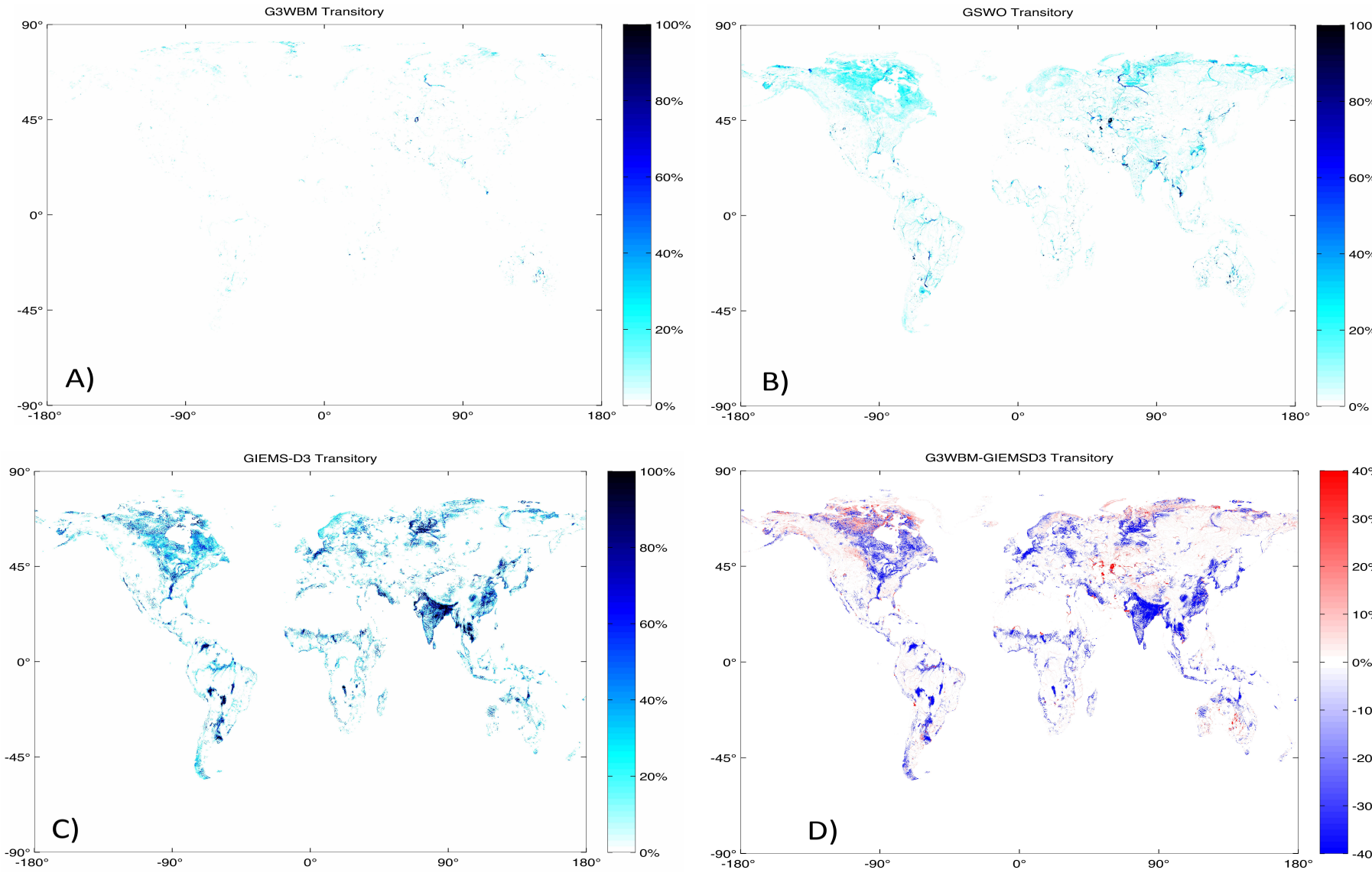


Figure 3: Area-fraction (in %) of transitory inundation states over 0.1°× 0.1° pixels for G3WBM (A), GSWO (B), and GIEMS-D3 (C). The last figure represents the area differences (in %) of GSWO minus GIEMS-D3 (D).

ing the wet season. Landsat observations cannot be used efficiently in these conditions to detect the presence of water, limiting drastically the possible detection of transitory surface water. Third, around extensive hydrological land-water systems (e.g., the Mississippi, Orinoco or Amazon Rivers and floodplains, or the Pantanal wetlands), the GSWO transitory water bodies occupy less area than in GIEMS-D3. With Landsat observations hampered by vegetation and clouds, an underestimation of the transitory water bodies can be expected in particular for densely vegetated areas prone to large cloud cover. Finally, at high latitudes, differences between GSWO and GIEMS-D3 likely result from the combination of two factors: the lack of spatial resolution in the original GIEMS data to detect very small lake fractions and the difference in the reference period to calculate the temporal statistics of surface water (with or without the boreal winter).

The longitudinal and latitudinal distribution of surface water extents are illustrated in Fig. 4 for permanent and transitory surface waters of all three estimates (G3WBM, GSWO, and GIEMS-D3). The agreement of the permanent surface water extents is reasonable, with slightly higher values for GSWO in the tropical regions, especially in South America. The slight underestimation around the equator of Landsat estimates with respect to GIEMS-D3 is expected to be related to the rain forest or clouds that block the visible and near infrared observations. Transitory water bodies, on the other hand, show much weaker agreement among sources, with substantially more transitory water surfaces from GIEMS-D3 in comparison to G3WBM and GSWO, especially over higher latitudes in North America and Eurasia. These spatial differences can also be observed in Figs. 2(D) and 3(D). More detailed

explanations on these differences will be given in the following sections.

Fig. 5 presents histograms of the probability distribution function for permanent and transitory surface water extents over $0.1^\circ \times 0.1^\circ$ pixels, at different latitudinal bands. G3WBM and GSWO estimates reveal strong agreement for permanent water bodies, regardless of the environment. GIEMS-D3 shows slightly less pixels with a smaller percentage of permanent water coverage (below 5%), across all latitudinal sections, yet slightly more with a higher percentage (above 10%). The increase in the probability for the surface water percentage close to 100% for all products is due to water bodies (mostly lakes) larger than the $0.1^\circ \times 0.1^\circ$ grid used for the calculation. For the transitory water surfaces, the pattern of under-detected small water fractions and exceeding large water fractions by GIEMS-D3 compared to the other products is more prominent.

A $0.1^\circ \times 0.1^\circ$ pixel-to-pixel comparison of the GSWO and GIEMS-D3 datasets (from Fig.s 2 and 3) is provided in Fig. 6. The first row represents the GSWO/GIEMS-D3 confusion matrices for permanent (left) and transitory (right) water extents (in percentage between 0 and 1). The colorbar represents the logarithm of the number of $0.1^\circ \times 0.1^\circ$ samples in each 0.1×0.1 bin of the plot. We observe that for permanent water surfaces, this matrix is quite diagonal. This means that there is no notable difference between the two datasets, except for the differences that can result from random errors in both datasets. For the transitory pixels, GSWO has a reduced range of water extent percentages (x-axis), whereas GIEMS-D3 shows a much broader range and increasing frequencies of high water surface percentages. This can be explained by the fact that vegetated fractions in the $0.1^\circ \times 0.1^\circ$ pixels can

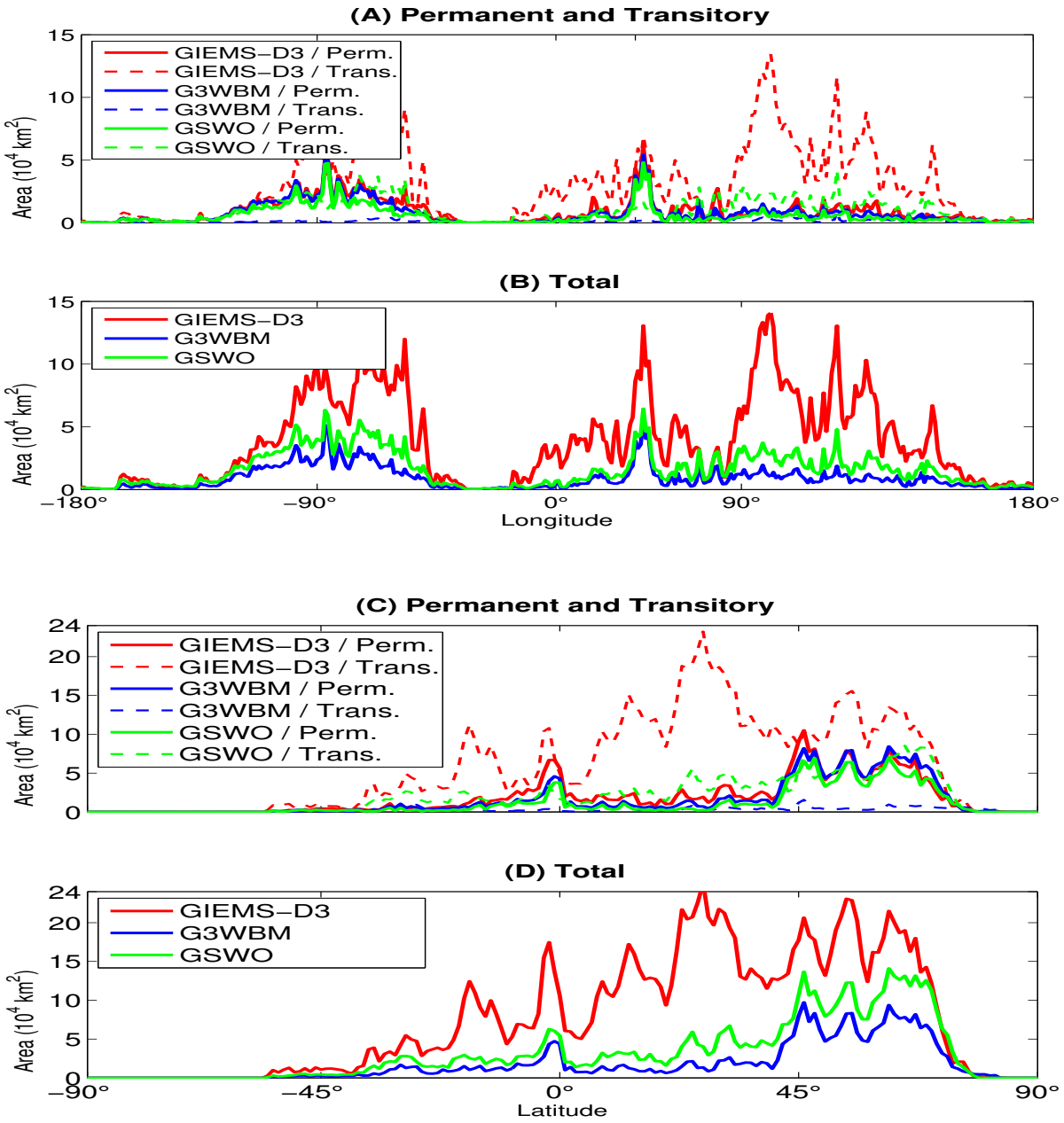


Figure 4: Averaged longitudinal (A and B) and latitudinal (C and D) surface water extents (in 10^4 km^2 per 1° bin), for G3WBM, GSWO, and GIEMS-D3 estimates. Distributions are plotted for permanent and transitory surface water extents, and for their sum.

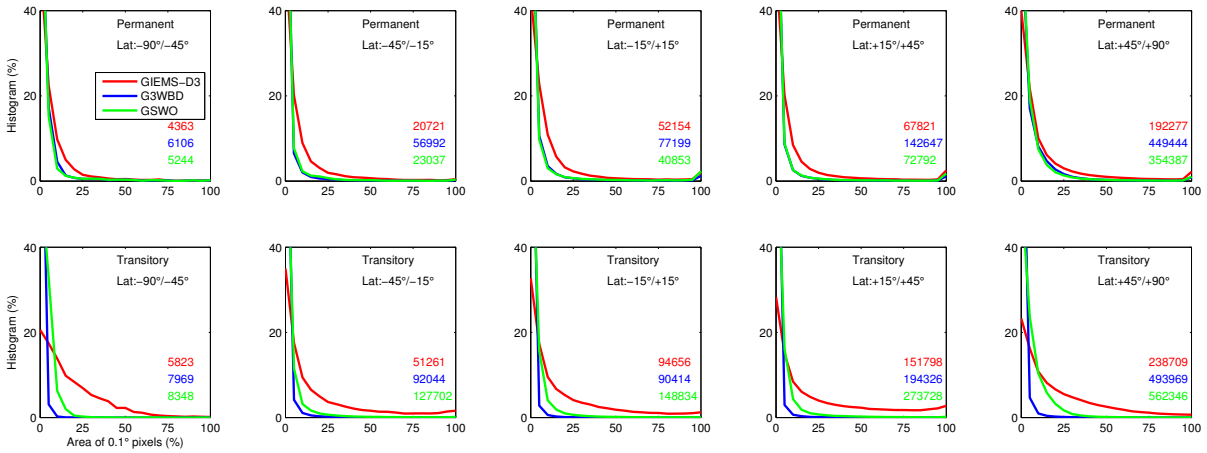


Figure 5: Histograms of the permanent (top row) and transitory (bottom row) surface water percentage over 0.1°× 0.1° pixels, for GIEMS-D3 (red), G3WBM (blue), and GSWO (green). From left to right: for -90°/-45°, -45°/-15°, -15°/+15°, +15°/+45°, and +45°/+90° latitudinal bands. The number of 0.1°× 0.1° pixels for each dataset is also indicated.

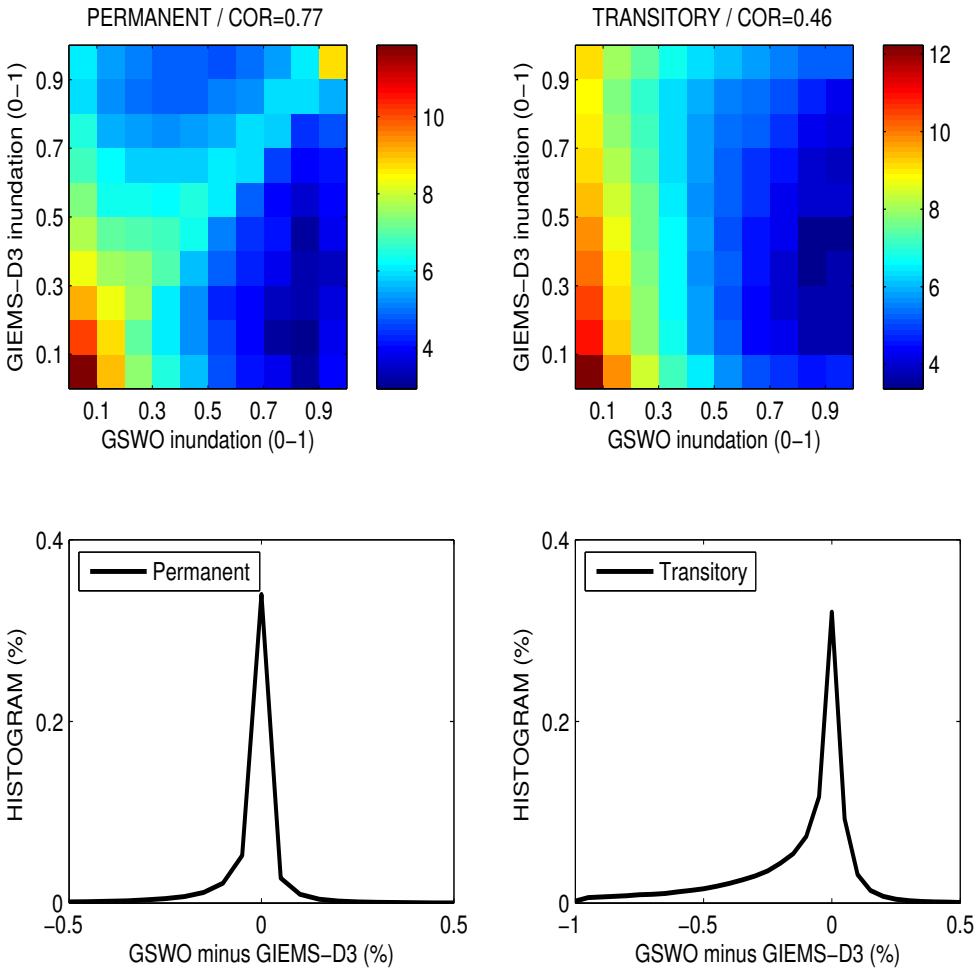


Figure 6: Pixel-to-pixel ($0.1° \times 0.1°$) comparison of the GSWO and GIEMS-D3 datasets, for permanent (left) and transitory (right) surface waters. First row: GSWO/GIEMS-D3 confusion matrices for permanent (left) and transitory (right) water extents (in percentage between 0 and 1). The colorbar represents the logarithm of the number of $0.1° \times 0.1°$ samples in each 0.1×0.1 bin of the plot. Second row: corresponding histogram of differences between the GSWO and GIEMS-D3 water extent estimates (percentages are expressed with the 0-1 range.

be detected in GIEMS-D3, but not in GSWO. This will be examined in more detail in the following sections. The lower row represents the corresponding histograms of differences between the GSWO and GIEMS-D3 water extent estimates. For the permanent waters, this distribution is centred around zero, and the distribution is symmetric meaning that no dataset is over-estimating the water extent compared to the other. For the transitory water extents, the distribution is skewed with a long tail to the left meaning that GIEMS-D3 has more pixels with high transitory water extents.

4. Regional evaluation

In order to investigate the advantages and disadvantages of the various inundation datasets, a regional analysis is performed in this section over several contrasted environments.

4.1. The Amazon River in a tropical forest environment

Fig. 7 presents a comparison between Landsat estimates from G3WBM and GSWO, the GIEMS-D3 probabilities, and the GLWD classification, over the Amazon Basin. Tree cover density is also shown to support interpretation. Overall, this figure supports the general findings of the global-scale comparison (Fig. 1). The spatial structure of the inundation pattern is quite similar for all three satellite datasets and they exhibit very realistic distributions of major rivers and tributaries (Amazon, Solimoes, Negro, Tapajos, Tocantins) as well as major associated inundated areas and wetlands. This finding aligns with several previous studies where GIEMS was intensively evaluated over the Amazon Basin (Papa et al., 2008, 2013; Frappart et al.,

2012; Getirana et al., 2012). Nevertheless, on a finer spatial scale, clear differences can be observed in the morphological patterns, especially over the flooded zones of the Rio Negro Basin (2°S-2°N; 64°W-61°W), the Madeira Basin (15°S-7°S; 69°W-62°W), or the floodplains adjacent to the main channel of the Amazon River. These regions are characterised by extensive savanna or forest floodplains with dense vegetation canopies that are inundated for four to six months each year (Fig. 7e).

To better illustrate the correlation between Landsat retrieval and vegetation, Fig. 8 compares the GSWO and GIEMS-D3 inundation probabilities for ten tree density ranges (0-10% to 90-100%), over a 5°× 5° cell in the Amazon region (5°S-0°S; 70°W-65°W). The fraction of pixels classified as inundated is estimated for each tree density bin. It can be seen in Fig. 7e that the tree density fall generally within the 90-100% or the 0-10% bins; however enough points (several thousand) are present in each intermediate tree density bin so that the inundation probabilities are robust and reliable, which is also corroborated by the fact that the lines in Fig 8 are smooth rather than randomly spiking through the transition zone. The sum of permanent and transitory inundation probability for low vegetation (0-10% bin) is high for both GIEMS-D3 and GSWO datasets, at 97% and 89%, respectively. However, the distribution of permanent and transitory inundation is quite different: 82% permanent and 15% transitory for GIEMS-D3, and 40% permanent and 49% transitory for GSWO. For the GSWO dataset, permanently inundated pixels can be found only in the 0-10% bin while transitory pixels decrease linearly from 50% to 0% with increasing tree cover. This behaviour may, at least in part, be influenced by the thicker forest density in upland

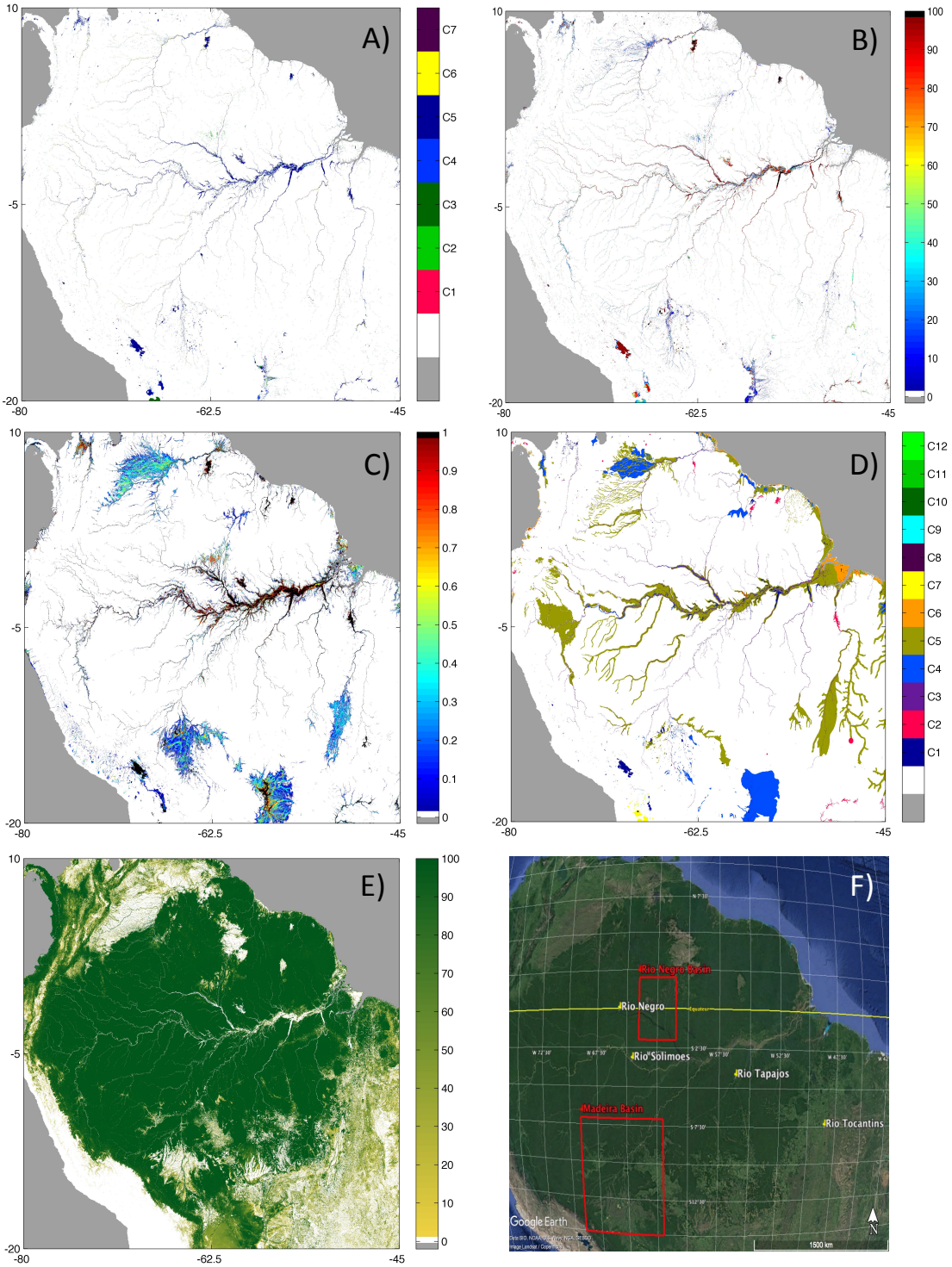


Figure 7: Comparison of the different surface water dataset over the Amazon: G3WBM (A), GSWO (B), GIEMS-D3 (C), GLWD (D), tree density from MODIS for interpretation (E), and indication of the two commented basins and four rivers (F). See Tab. 1 for the definition of the G3WBM and GLWD classes.

areas compounded with the limitations of Landsat detection.

For GIEMS-D3, the transitory inundation probability in tree density bins is fairly constant from 10 to 90%, then experiences a steep decline in the 90-100% bin. A peak occurs at 80-90% which might be related to the overestimation of transitory inundation for water-saturated soils. With GIEMS-D3, still 13% of transitory inundation is observed for the highest tree density bin, confirming that flooding can be detected by GIEMS-D3 even under densely vegetated areas.

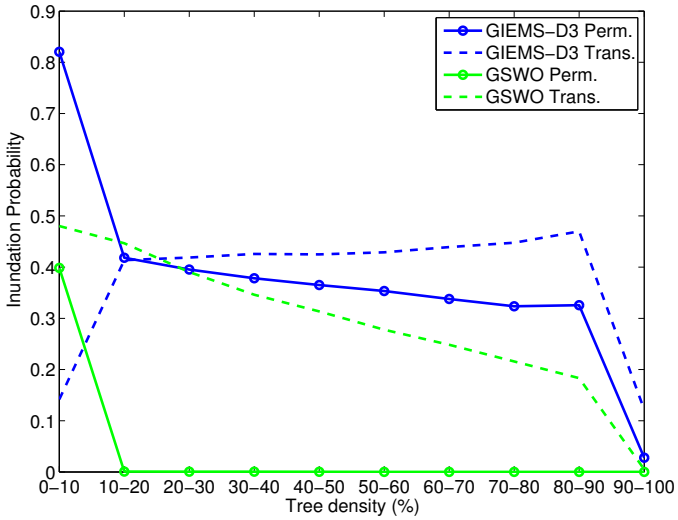
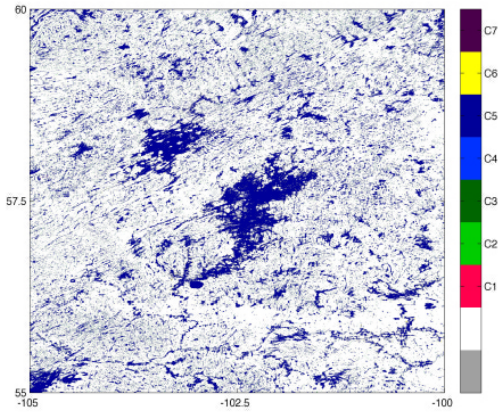


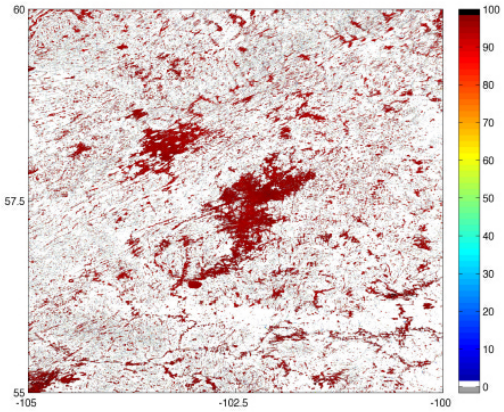
Figure 8: Probability of inundation for ten tree density ranges (0-10% to 90-100%), over a 5°× 5° cell in the Amazon region (5°S-0°S; 70°W-65°W). The probability of being inundated (permanently in continuous lines, or transitory in dashed lines) is estimated for each tree density bin, for GIEMS-D3 (blue) and GSWO (green).

4.2. Small water bodies at higher latitudes

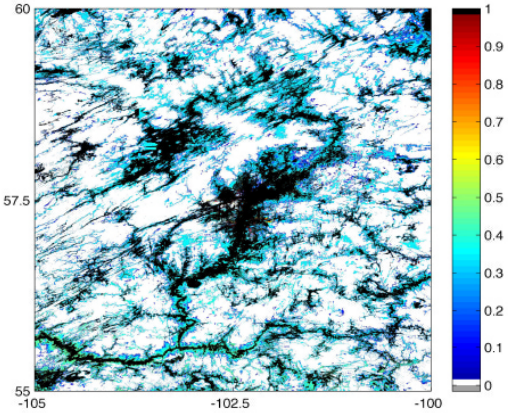
Fig. 9 illustrates the surface water estimates of the G3WBM, GSWO, GIEMS-D3, and GLWD datasets over a high-latitude area (56°N-60°N; 105W°-100°W) in Canada and supports that the larger hydrological features are very similar from one dataset to another. The reference data of GLWD shows an extensive homogeneous area of a wetland complex (classified as "25-50% wetland"), likely representing a generalised peatland region. This "biome class" is not replicated in the other datasets that convey a water classification information. Landsat estimates (G3WBM and GSWO) show no transitory surface waters but only permanent ones. This pattern is likely related to an inundation definition issue or a time sampling problem causing confusion with snow/ice cover, as inundation variation occurs during snowmelt periods at these latitudes. Another possible explanation is also the presence of vegetation since Landsat cannot sense water beneath vegetation. On the other hand, GIEMS-D3 shows a gradient of inundation probability, from the center of the hydrological objects to the borders influenced likely by some confusion with snow/ice cover and the limited ability to capture small lake features in the landscape. The problem may be amplified by the downscaling scheme relying on global-scale topographic and hydrographic information that is too uncertain to adequately represent small variations in elevation within flat regions. This regional case study over the northern latitudes highlights the sampling issues in Landsat, and the difficulty for GIEMS-D3 to retrieve small water bodies.



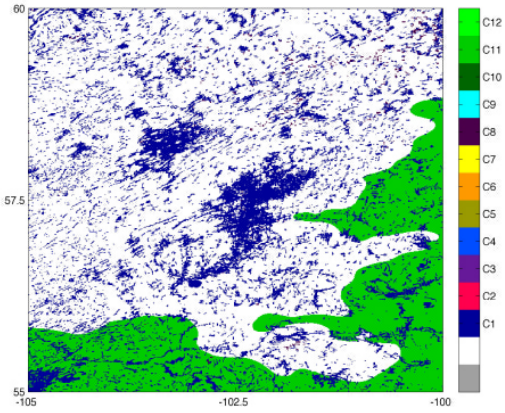
(a) G3WBM Classif.



(b) GSWO Occur.



(c) GIEMS-D3 Prob.



(d) GLWD Classif.

Figure 9: Inundation estimates from G3WBM (a), GSWO (b), GIEMS-D3 (c) and GLWD (d), over a 5 °× 5° cell in Canada. See Tab. 1 for the definition of the G3WBM and GLWD classes.

4.3. The Ganges-Brahmaputra region and related cloud issues

Characterising the distribution and variability of surface freshwater at high resolution is critical for the Indian sub-continent and surrounding areas where the availability of water resources is highly influenced by anthropogenic pressure. Fig. 10 compares, over the Ganges-Brahmaputra River system and the Bengal Delta, the various datasets. To support the interpretations, Fig. 10 also shows tree density (E), the mean climatological cloud fraction from MODIS in August over the region (F), the seasonality of the cloud cover (G), and the number of months in which water is present during 2014 from GSWO (H).

For reference, the GLWD map coarsely depicts three larger homogeneous wetland areas classified as freshwater marsh and floodplains (C4), coastal wetlands (C6), and 0-25% wetlands (C12). The three satellite-derived products agree in their representation of the permanent major rivers (Ganges-Brahmaputra-Meghna River systems) and their large tributaries, as well as some of the associated large inundated areas or wetlands of the Meghna River and the Bengal Delta. However, this region is characterised by the presence of complex areas with extensive transitory and seasonal flooding, which are better captured by the more complete temporal coverage of GSWO compared to G3WBM. For instance, the flooding region along the confluence of the Ganges and the Kosi River in India as well as the floodplain of the Meghna River are captured in GSWO, yet they are not visible in G3WBM. The overall low coverage of surface water in both Landsat estimates is in high contrast to GIEMS-D3 estimates, which are characterised by extensive and in some cases overestimated flooded areas. The issue of overestimation in

the original GIEMS data over the Indian subcontinent has been discussed in several publications (Papa et al., 2006, 2008, 2015; Salameh et al., 2017) and while GIEMS is broadly able to capture the distributions and variations of surface freshwater in the Ganges-Brahmaputra Basin, some analysis suggests that the method encounters difficulties in accurately discriminating between very saturated/moist soil and standing open water which can lead to potential overestimations of actual surface water extents, especially for saturated soil in pixels with high flood coverage such as the delta region (Papa et al., 2010).

Unlike in the Amazon Basin, low surface water detection in GSWO and G3WBM over the Ganges-Brahmaputra region are not associated with dense tree cover, see Fig. 10(E). On the other hand, Fig. 10(F) reveals that in August the entire region is covered by clouds, in particular the areas around the lower Ganges-Brahmaputra and delta which are covered by clouds for 90% of the month. August is also the month when excessive rainfall caused by the South-West Indian monsoon produces maximum river discharge in the Ganges-Brahmaputra system (Papa et al., 2012) along with associated large flooding in the basin and in the lower delta. The persistent presence of cloud cover during the monsoon season from May to September, see Fig. 10(G), suggests that GSWO and G3WBM might encounter difficulties to properly retrieve the maximum extent of inundation and seasonal surface water bodies during their peak season. This might explain why GSWO detects most of the surface water in April-May or October-December as shown in Fig. 10(G).

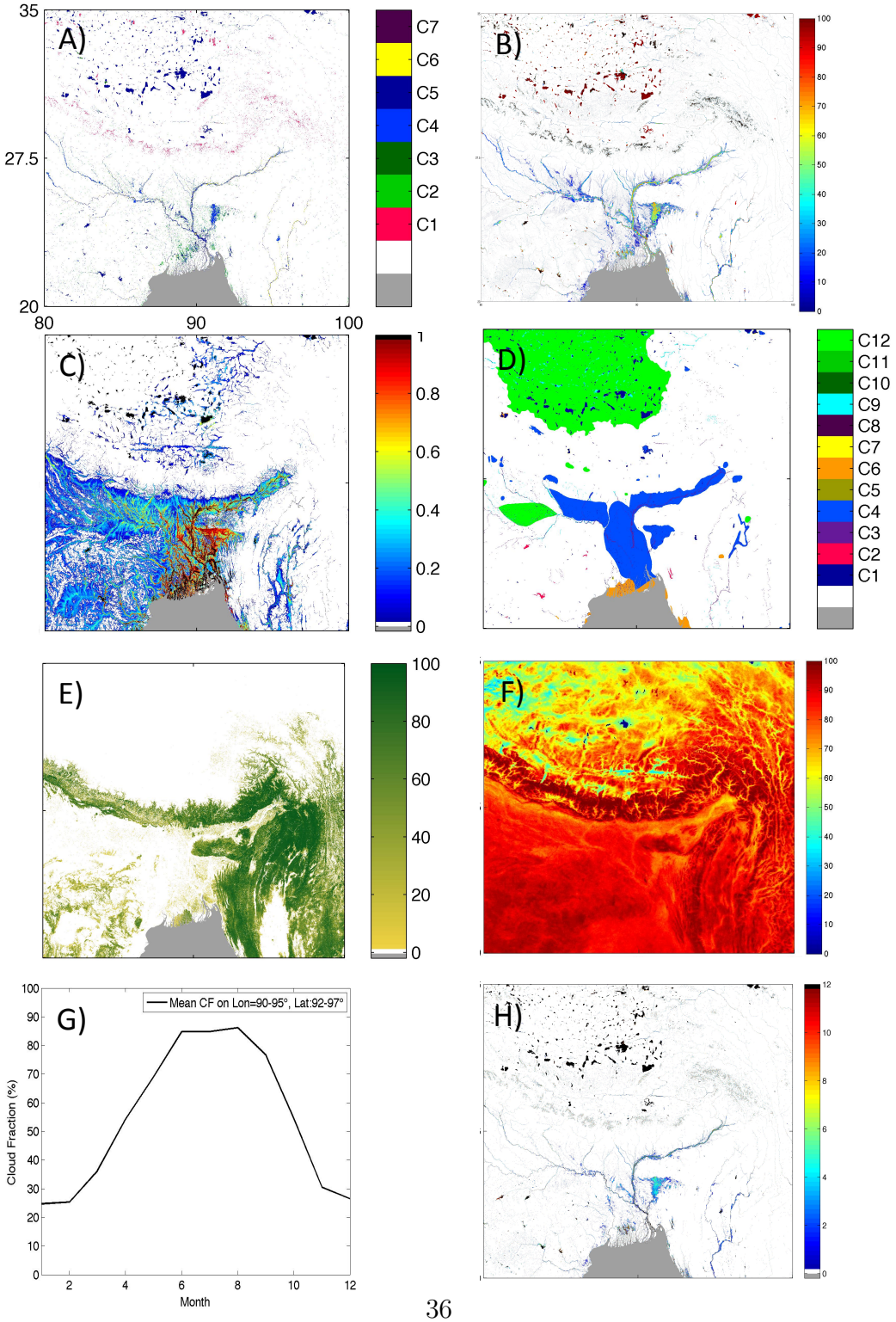


Figure 10: Comparison of surface water datasets for the Ganges-Brahmaputra region: G3WBM (A), GSWO (B), GIEMS-D3 (C), and GLWD (D). Tree density from MODIS (E), cloud fraction in August (F), seasonal cloud fraction (G) and GSWO season duration (in months) (H) are also represented for interpretation purposes. See Tab. 1 for the definition of G3WBM and GLWD classes.

5. Conclusion and perspectives

5.1. Conclusion

A comparison of Landsat-based (G3WBM and GSWO) and multi-satellite (GIEMS-D3) datasets was conducted at global and regional scales to identify advantages and disadvantages in terms of detecting water surface areas for both types of approaches. Overall, good agreement in permanent water bodies was found for all datasets. Due to differences in observation methods, discrepancies are apparent in particular for transitory (i.e. not permanently inundated) zones. Landsat-based datasets have a very high nominal spatial resolution and can detect even small water bodies at high accuracy, such as those ubiquitous in northern latitudes. However, Landsat observations are affected by cloud cover, snow/ice and vegetation. As a result, Landsat products are most apt at detecting open water bodies and variation within them, but provide incomplete representation of seasonal wetlands or inundation areas where the spectral signature is confounded by vegetation.

The multiple sensors behind GIEMS-D3 can penetrate vegetation and clouds and are thus capable of retrieving non-open water wetland types and of detecting temporal dynamics even if seasonal cloud cover is prevalent during the periods of the year when inundation occurs. Current implementations of the GIEMS-D3 downscaling method, however, have used topographic information which does not provide spatial accuracy comparable to direct observations from Landsat sensors. Furthermore, the current version of GIEMS probably over-estimates inundation over water-saturated soils, and due to its low nominal spatial resolution, it is not able to retrieve small water bodies.

SAR data (from ENVISAT, Sentinel and soon SWOT) would be a good

supplement because they warrant high-spatial resolution and are less sensitive to clouds than the visible observations. However, long-term and global mapping of surface waters has remained challenging due to localised parameterisation necessary for obtaining locally accurate maps. Regardless, SAR remains among the most promising approaches for global baseline mapping of inundation in vegetated areas, even if it cannot provide a long-term historic time record such as available from GIEMS or Landsat.

In summary, the data sources compared possess complimentary strengths which could be leveraged through assimilation and combination. The extensive archive of Landsat imagery provides reliable estimates of open water bodies and variation, while downscaled GIEMS maps such as GIEMS-D3 provide a more complete, if less locally accurate, extent of inundation under canopy and cloud cover.

5.2. Perspectives

In order to obtain the best estimate of global wetland extents at high spatial resolution and with a long-term record, several ways forward should be investigated which are listed in the following.

Improvement of GIEMS and GIEMS-D3 - It was shown that GIEMS and GIEMS-D3 have some good complementary information to Landsat observations. However, GIEMS and GIEMS-D3 also need to be further enhanced. Three main ways are suggested to improve the original GIEMS dataset: (1) The retrieval algorithm could be corrected to reduce the over-estimation of inundated areas over saturated soils. (2) The temporal resolution could be increased, from the monthly to a 10-day scale. (3) Most importantly, the time

record of GIEMS (currently 1993-2007) could be extended on both ends. The time period covered by GIEMS is limited by auxiliary information (e.g. temporally coherent surface temperatures, cloud flags) required for the GIEMS retrievals, whereas the microwave observations are available for a longer period (1972-present). This information was obtained form the ISCCP (International Satellite Cloud Climatology Project)(Rossow and Schiffer, 1999). Alternative ways could be investigated to perform the retrieval without these auxiliary information, which would allow for building a GIEMS dataset from 1978 to present (about 40 years).

GIEMS-D3 could also be improved in several ways, for instance by using a better downscaling scheme: (1) The floodability index used in the downscaling could be trained on Landsat data instead of GLWD. (2) A floodability index could be developed with regional tuning adapted to different wetland types (currently, the same model is used for arid, tropical, temperate or polar areas) (Aires et al., 2017). (3) Other Digital Elevation Models (DEMs) could be used instead of, or in combination with, SRTM/HydroSHEDS. For instance, the newer version of SRTM at 30 m resolution could be tested. New DEMs (Yamazaki et al., 2017) could also correct some artefacts such as vegetation noise or tree height contamination (e.g., (Simard et al., 2011)).

Data fusion of several sources of information - Since both Landsat-based data and GIEMS-D3 carry complementary information, it would be valuable to combine them to obtain a better inundation product. There are several avenues to achieve this: (1) The permanent water body area from Landsat could be used to correct the minimum values of GIEMS-D3 which currently

tend to be underestimated due to the limited retrieval capability of GIEMS in areas of low water fractions. (2) The Landsat occurrence could also replace the topography-based floodability index that is used to perform the downscaling of GIEMS-D3. In this way, an updated version of GIEMS-D3 would be more compatible to Landsat products because it would follow similar spatial patterns. The difficulty with of this approach is that the maximum GIEMS inundation estimate is higher than that from Landsat, hence a Landsat-based probability map is not covering the required extent. This shortcoming could be solved by fusing the Landsat-based and topography-based probability indices.

Aside from water and inundation masks, categorising surface water types could benefit from developments of databases specific to discriminated waterbody types. For instance, lakes and reservoirs (HydroLAKES) (Messager et al., 2016) or the Global River Widths from Landsat (GRWL) (Allen and Pavelsky, 2015) could be used to improve downscaling of GIEMS or for post-hoc classification efforts.

Combination with SAR data - Although global SAR coverage may eventually provide the most comprehensive estimates of surface waters on its own, steps improving global SAR retrieval would benefit from the use of GIEMS or Landsat datasets. For instance, the GIEMS/Landsat estimates could help in the development of a global SAR retrieval algorithm by serving either as an *a priori* mask or as a first guess. GIEMS/Landsat combination would also be a good tool to select the sites of validation campaigns for the SWOT mission that will be launched in 2021 (Rodriguez, 2015; Prigent et al., 2016).

Finally, it could help in assessing the sensitivity of hydrological models to this kind of high resolution data (e.g. assimilation experiments to estimate river discharges), feed the SWOT simulator, or measure signal-to-noise ratio constraints.

Arguably the most important reason for a continued improvement of surface water datasets based on Landsat or multi-sensor products like GIEMS is that they are our most promising tools to preserve the best possible inundation record of the past. Once reliable global SAR estimates become available (from ENVISAT, Sentinel-1 or SWOT), they can be combined with historic GIEMS/Landsat estimates to construct an instantaneous record of inundation areas back to 1978, which then will be carried into the future by the new and advanced SAR measurements.

6. Acknowledgements

The authors would like to thank NASA for their support at Columbia University for the project entitled "Downscaling of flooded fraction derived from low-resolution microwave measurements" (Contract number NNH13CH27C) lead by John Galantowicz at Atmospheric and Environmental Research (AER, Inc.). We would also like to thank the French spatial agency CNES ("Centre National d'Etudes Spatiales"), and in particular Selma Cherchali for funding in 2012/2013 a study related to this work (Aires et al., 2013, 2014), and Nicolas Picot for funding a "Research and Technology" project on the space interpolation of wetland maps in the framework of the SWOT mission. We would like to thank the European Union for funding the Earth2Obs project that allowed us to perform part of the evaluation of the GIEMS-D3 prod-

uct (Aires et al., 2017). Images from the Global Surface Water Occurrence (Landsat) are courtesy of the European Commission Joint Research Centre, and we would like to thank Jean-François Pekel and colleagues for making them available.

References

Aires, F., Miolane, L., Prigent, C., Pham, B., Fluet-Chouinard, E., Lehner, B., Papa, F., 2017. A Global Dynamic Long-Term Inundation Extent Dataset at High Spatial Resolution Derived through Downscaling of Satellite Observations. Journal of Hydrometeorology 18, 1305–1325. URL: http://journals.ametsoc.org/doi/10.1175/JHM-D-16-0155.1, doi:10.1175/JHM-D-16-0155.1.

Aires, F., Papa, F., Prigent, C., 2013. A Long-Term, High-Resolution Wetland Dataset over the Amazon Basin, Downscaled from a Multi-wavelength Retrieval Using SAR Data. Journal of Hydrometeorology 14, 594–607. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2013JHyMe..14..594A&link_type=EJOURNAL, doi:10.1175/JHM-D-12-093.1.

Aires, F., Papa, F., Prigent, C., Crétaux, J.F., Bergé-Nguyen, M., 2014. Characterization and Space–Time Downscaling of the Inundation Extent over the Inner Niger Delta Using GIEMS and MODIS Data. Journal of Hydrometeorology 15, 171–192. URL: http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-13-032.1, doi:10.1175/JHM-D-13-032.1.

692 Allen, G.H., Pavelsky, T.M., 2015. Patterns of river width and sur-
693 face area revealed by the satellite-derived North American River Width
694 data set. Geophysical Research Letters 42, 395–402. URL: http://
695 onlinelibrary.wiley.com/doi/10.1002/2014GL062764/full, doi:10.
696 1002/2014GL062764.

697 Biancamaria, S., Lettenmaier, D.P., Pavelsky, T.M., 2016. The SWOT Mis-
698 sion and Its Capabilities for Land Hydrology. Surveys in Geophysics 37,
699 307–337. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_
700 query?bibcode=2016SGeo...37..307B&link_type=EJOURNAL,
701 doi:10.1007/s10712-015-9346-y.

702 Bousquet, P., Ciais, P., Miller, J.B., Dlugokencky, E.J., Hauglustaine, D.A.,
703 Prigent, C., Van der Werf, G.R., Peylin, P., Brunke, E.G., Carouge, C.,
704 Langenfelds, R.L., Lathière, J., Papa, F., Ramonet, M., Schmidt, M.,
705 Steele, L.P., Tyler, S.C., White, J., 2006. Contribution of anthropogenic
706 and natural sources to atmospheric methane variability. Nature 443, 439–
707 443. URL: http://www.nature.com/doifinder/10.1038/nature05132,
708 doi:10.1038/nature05132.

709 Chahine, M.T., 1992. The hydrological cycle and its influence on cli-
710 mate. Nature 359, 373–380. URL: http://adsabs.harvard.edu/
711 cgi-bin/nph-data_query?bibcode=1992Natur.359..373C&link_type=
712 EJOURNAL, doi:10.1038/359373a0.

713 Crétaux, J.F., Abarca-del Río, R., Berge-Nguyen, M., Arsen, A.,
714 Drolon, V., Clos, G., Maisongrande, P., 2016. Lake Volume Moni-
715 toring from Space. Surveys in Geophysics 37, 269–305. URL: http:

716 //adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2016SGeo.
717 ..37..269C&link_type=EJOURNAL, doi:10.1007/s10712-016-9362-6.

718 Feng, M., Sexton, J.O., Channan, S., Townshend, J.R., 2014. A global,
719 high-resolution (30-m) inland water body dataset for 2000: first results
720 of a topographic–spectral classification algorithm. International Journal
721 of Digital Earth 9, 113–133. URL: http://www.tandfonline.com/doi/
722 full/10.1080/17538947.2015.1026420, doi:10.1080/17538947.2015.
723 1026420.

724 Fluet-Chouinard, E., Lehner, B., Rebelo, L.M., Papa, F., Hamilton, S.K.,
725 2015. Development of a global inundation map at high spatial resolution
726 from topographic downscaling of coarse-scale remote sensing data. Re-
727 mote Sensing of Environment 158, 348–361. URL: http://linkinghub.
728 elsevier.com/retrieve/pii/S0034425714004258, doi:10.1016/j.rse.
729 2014.10.015.

730 Frappart, F., Papa, F., da Silva, J.S., Ramillien, G., Prigent, C.,
731 Seyler, F., Calmant, S., 2012. Surface freshwater storage and dynam-
732 ics in the Amazon basin during the 2005 exceptional drought. Env-
733 iron. Res. Lett. 7. URL: http://stacks.iop.org/1748-9326/7/i=4/
734 a=044010?key=crossref.366806a20b43202797338f61d0a23969, doi:10.
735 1088/1748-9326/7/4/044010.

736 Getirana, A.C.V., Boone, A., Yamazaki, D., Decharme, B., Papa, F.,
737 Mognard, N., 2012. The Hydrological Modeling and Analysis Platform
738 (HyMAP): Evaluation in the Amazon Basin. Journal of Hydrometeorology
739 13, 1641–1665. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_

query?bibcode=2012JHyMe..13.1641G&link_type=EJOURNAL, doi:10.1175/JHM-D-12-021.1.

Gleick, P.H., 1989. Climate change, hydrology, and water resources. Reviews of Geophysics 27, 329–344. URL: http://doi.wiley.com/10.1029/RG027i003p00329, doi:10.1029/RG027i003p00329.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Carroll, M., Dimiceli, C., Sohlberg, R.A., Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Carroll, M., Dimiceli, C., Sohlberg, R.A., 2009. Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm 7, 1–15. URL: http://journals.ametsoc.org/doi/abs/10.1175/1087-3562(2003)007%3C0001%3AGPTCAA%3E2.0.CO%3B2, doi:10.1175/1087-3562(2003)007<0001:GPTCAA>2.0.CO;2.

Lehner, B., Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and wetlands. Journal of Hydrology 296, 1–22. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2004JHyd..296....1L&link_type=EJOURNAL, doi:10.1016/j.jhydrol.2004.03.028.

Lehner, B., Verdin, K., Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union 89, 93–94. URL: http://doi.wiley.com/10.1029/eost2008EO10, doi:10.1029/eost2008EO10.

Messager, M.L., Lehner, B., Grill, G., Nedeva, I., Schmitt, O.,

2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nature Communications 7, 13603. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2016NatCo...713603M&link_type=EJOURNAL, doi:10.1038/ncomms13603.

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., Ip, A., 2016. Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia. Remote Sensing of Environment 174, 341–352. URL: http://dx.doi.org/10.1016/j.rse.2015.11.003, doi:10.1016/j.rse.2015.11.003.

Nakaegawa, T., 2012. Comparison of Water-Related Land Cover Types in Six 1-km Global Land Cover Datasets. Journal of Hydrometeorology 13, 649–664. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2012JHyMe..13..649N&link_type=EJOURNAL, doi:10.1175/JHM-D-10-05036.1.

Papa, F., Bala, S.K., Pandey, R.K., Durand, F., Gopalakrishna, V.V., Rahman, A., Rossow, W.B., 2012. Ganga-Brahmaputra river discharge from Jason-2 radar altimetry: An update to the long-term satellite-derived estimates of continental freshwater forcing flux into the Bay of Bengal. J. Geophys. Res. Oceans 117, C11021. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2012JGRC..11711021P&link_type=EJOURNAL, doi:10.1029/2012JC008158.

Papa, F., Frappart, F., Güntner, A., Prigent, C., Aires, F., Getirana, A.C.V., Maurer, R., 2013. Surface freshwater storage and variability in the Amazon basin from multi-satellite observations, 1993–2007. Journal of Geophysical Research: Atmospheres 118, 11,951–11,965. URL: http://onlinelibrary.wiley.com/doi/10.1002/2013JD020500/full, doi:10.1002/2013JD020500.

Papa, F., Frappart, F., Malbeteau, Y., Shamsudduha, M., Vuruputur, V., Sekhar, M., Ramillien, G., Prigent, C., Aires, F., Pandey, R.K., Bala, S., Calmant, S., 2015. Satellite-derived surface and sub-surface water storage in the Ganges–Brahmaputra River Basin. Journal of Hydrology: Regional Studies 4, 15–35. URL: http://dx.doi.org/10.1016/j.ejrh.2015.03.004, doi:10.1016/j.ejrh.2015.03.004.

Papa, F., Guntner, A., Frappart, F., Prigent, C., Rossow, W.B., 2008. Variations of surface water extent and water storage in large river basins: A comparison of different global data sources. Geophysical Research Letters 35, L11401. URL: http://doi.wiley.com/10.1029/2008GL033857, doi:10.1029/2008GL033857.

Papa, F., Prigent, C., Aires, F., Jimenez, C., Rossow, W.B., Matthews, E., 2010. Interannual variability of surface water extent at the global scale, 1993–2004. Journal of Geophysical Research: Atmospheres (1984–2012) 115. URL: http://onlinelibrary.wiley.com.ezproxy.cul.columbia.edu/doi/10.1029/2009JD012674/full, doi:10.1029/2009JD012674.

Papa, F., Prigent, C., Rossow, W.B., 2006. Inundated wetland dynamics over boreal regions from remote sensing: The use of Topex-Poseidon dual-

frequency radar altimeter observations 27, 4847–4866. URL: http://www.tandfonline.com/doi/abs/10.1080/01431160600675887, doi:10.1080/01431160600675887.

Pekel, J.F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2016Natur.540..418P&link_type=EJOURNAL, doi:10.1038/nature20584.

Pham-Duc, B., Prigent, C., Aires, F., 2017. Surface Water Monitoring within Cambodia and the Vietnamese Mekong Delta over a Year, with Sentinel-1 SAR Observations. Water 9, 366–21. URL: http://www.mdpi.com/2073-4441/9/6/366, doi:10.3390/w9060366.

Prigent, C., Lettenmaier, D.P., Aires, F., Papa, F., 2016. Toward a High-Resolution Monitoring of Continental Surface Water Extent and Dynamics, at Global Scale: from GIEMS (Global Inundation Extent from Multi-Satellites) to SWOT (Surface Water Ocean Topography). Surveys in Geophysics 37, 339–355. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2016SGeo...37..339P&link_type=EJOURNAL, doi:10.1007/s10712-015-9339-x.

Prigent, C., Matthews, E., Aires, F., Rossow, W.B., 2001. Remote sensing of global wetland dynamics with multiple satellite data sets. Geophysical Research Letters 28, 4631–4634. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2001GeoRL..28.4631P&link_type=EJOURNAL, doi:10.1029/2001GL013263.

Prigent, C., Papa, F., Aires, F., Jimenez, C., Rossow, W.B., Matthews, E., 2012. Changes in land surface water dynamics since the 1990s and relation to population pressure. Geophysical Research Letters 39, L08403. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2012GeoRL..39.8403P&link_type=EJOURNAL, doi:10.1029/2012GL051276.

Prigent, C., Papa, F., Aires, F., Rossow, W.B., Matthews, E., 2007. Global inundation dynamics inferred from multiple satellite observations, 1993–2000. Journal of Geophysical Research: Atmospheres (1984–2012) 112. URL: http://onlinelibrary.wiley.com.ezproxy.cul.columbia.edu/doi/10.1029/2006JD007847/full, doi:10.1029/2006JD007847.

Rodriguez, E., 2015. Surface Water and Ocean Topography Mission (SWOT). Technical Report JPL D-61923. NASA JPL.

Rossow, W.B., Schiffer, R.A., 1999. Advances in understanding clouds from ISCCP. Bull. Amer. Meteor. Soc. 80, 2261–2287.

Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., Yokozawa, M., 2007. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. Remote Sensing of Environment 109, 295–313. URL: http://linkinghub.elsevier.com/retrieve/pii/S0034425707000466, doi:10.1016/j.rse.2007.01.011.

Salameh, E., Frappart, F., Papa, F., Güntner, A., Venugopal, V., Getirana, A., Prigent, C., Aires, F., Labat, D., Laignel, B., 2017. Fifteen

Years (1993–2007) of Surface Freshwater Storage Variability in the Ganges-Brahmaputra River Basin Using Multi-Satellite Observations. Water 9, 245–19. URL: http://www.mdpi.com/2073-4441/9/4/245, doi:10.3390/w9040245.

Santoro, M., Wegmüller, U., Askne, J.I.H., 2010. Signatures of ERS-Envisat Interferometric SAR Coherence and Phase of Short Vegetation: An Analysis in the Case of Maize Fields. IEEE Transactions on Geoscience and Remote Sensing 48, 1702–1713. URL: http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2010ITGRS..48.1702S&link_type=EJOURNAL, doi:10.1109/TGRS.2009.2034257.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. 116, G04021–12. URL: http://doi.wiley.com/10.1029/2011JG001708, doi:10.1029/2011JG001708.

Tulbure, M.G., Broich, M., Stehman, S.V., Kommareddy, A., 2016. Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. Remote Sensing of Environment 178, 142–157. URL: http://linkinghub.elsevier.com/retrieve/pii/S0034425716300621, doi:10.1016/j.rse.2016.02.034.

Verpoorter, C., Kutser, T., Seekell, D.A., Tranvik, L.J., 2014. A global inventory of lakes based on high-resolution satellite imagery. Geophys. Res. Lett. 41, 6396–6402. URL: http://doi.wiley.com/10.1002/2014GL060641, doi:10.1002/2014GL060641.

Westerhoff, R.S., Kleuskens, M.P.H., Winsemius, H.C., Huizinga, H.J., Brakenridge, G.R., Bishop, C., 2013. Automated global water mapping based on wide-swath orbital synthetic-aperture radar. Hydrol. Earth Syst. Sci. 17, 651–663. URL: http://www.hydrol-earth-syst-sci.net/17/651/2013/, doi:10.5194/hess-17-651-2013.

Wilson, A.M., Jetz, W., 2016. Remotely Sensed High-Resolution Global Cloud Dynamics for Predicting Ecosystem and Biodiversity Distributions. PLoS Biol 14, 1–20. URL: http://dx.plos.org/10.1371/journal.pbio.1002415, doi:10.1371/journal.pbio.1002415.

Winsemius, H., Jongman, B., Veldkamp, T., Hallegatte, S., Bangalore, M., Ward, P., 2015. Disaster risk and poverty: assessing the global exposure of the poor to floods and droughts. EGU General Assembly 2015 17, 3225. URL: http://adsabs.harvard.edu/abs/2015EGUGA..17.3225W.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations. Geophys. Res. Lett. 44, 5844–5853. URL: http://doi.wiley.com/10.1002/2017GL072874, doi:10.1002/2017GL072874.

Yamazaki, D., Trigg, M.A., Ikeshima, D., 2015. Development of a global ~90m water body map using multi-temporal Landsat images. Remote Sensing of Environment 171, 337–351. URL: http://linkinghub.elsevier.com/retrieve/pii/S0034425715301656, doi:10.1016/j.rse.2015.10.014.