Article:

Yamazaki, D, Trigg, MA orcid.org/0000-0002-8412-9332 and Ikeshima, D (2015) Development of a global ~90m water body map using multi-temporal Landsat images. Remote Sensing of Environment, 171. pp. 337-351. ISSN 0034-4257

https://doi.org/10.1016/j.rse.2015.10.014

Development of a global ~90 m water body map using multi-temporal Landsat images

Dai Yamazaki
Department of Integrated Climate Change Projection Research,
Japan Agency for Marine-Earth Science and Technology
3173-25, Showa-machi, Kanazawa-ku, Yokohama, Kanagawa, 236-0001, Japan
+81-45-778-5565
d-yamazaki@jamstec.go.jp

Mark A. Trigg
Willis Research Fellow, School of Geographical Sciences, University of Bristol
University Road, Clifton, Bristol, BS8 1SS, UK
Mark.Trigg@bristol.ac.uk

Daiki Ikeshima
Department of Civil Engineering, Tokyo Institute of Technology
2-12-1-M1-6 O-okayama, Meguro-ku, Tokyo 152-8552, Japan
ikeshima.d.aa@m.titech.ac.jp

Abstract

This paper describes the development of a Global 3 arc-second Water Body Map (G3WBM), using an automated algorithm to process multi-temporal Landsat images from the Global Land Survey (GLS) database. We used 33,890 scenes from 4 GLS epochs in order to delineate a seamless water body map, without cloud and ice/snow gaps. Permanent water bodies were distinguished from temporal water-covered areas by calculating the frequency of water body existence from overlapping, multi-temporal, Landsat scenes. By analyzing the frequency of water body existence at 3 arc-second resolution, the G3WBM separates river channels and floodplains more clearly than previous studies. This suggests that the use of multi-temporal images is as important as analysis at a higher resolution for global water body mapping. The global totals of delineated permanent water body area and temporal water-covered area are 3.25 and 0.49 million km^2 respectively, which highlights the importance of river-floodplain separation using multi-temporal images. The accuracy of the water body classification was

validated in Hokkaido (Japan) and in the contiguous United States using an existing water body databases. There was almost no commission error, and about 70% of lakes >1 km^2 shows relative water area error <25%. Though smaller water bodies (<1 km^2) were underestimated mainly due to omission of shoreline pixels, the overall accuracy of the G3WBM should be adequate for larger scale research in hydrology, biogeochemistry, and climate systems and importantly includes a quantification of the temporal nature of global water bodies.

Keywords

Landsat GLS, water body mapping, global analysis, river, floodplain

1. Introduction

1.1 Background

Terrestrial water in rivers and lakes is essential for both human beings and ecosystems (Oki and Kanae, 2006). River and lakes affect the climate system via land-atmosphere interaction processes such as carbon burial and CO_2 exchange as well as other biogeochemical processes (Cole et al., 2007; Sjögersten et al., 2014). Delineating the spatial and temporal distribution of rivers and lakes is important for understanding the water, energy and carbon cycles, both at local and global scales (Downing et al., 2012, 2014; Allen et al., 2015). Mapping water bodies at a global scale is therefore a fundamental step to understand the role of inland water bodies in climate systems (Palmer et al., 2015).

Until very recently, globally available water body maps have been limited in resolution, but in parallel with recent computational advances in various research fields, a high-resolution, high-accuracy global water body database is required. For example, global-scale water body maps have been used in river width calculation for hydrodynamic modeling (O'Loughlin et al., 2013; Yamazaki et al., 2014a and 2014b; Sampson et al., 2015). In addition, given that biogeochemical processes in small lakes may be more active than large lakes (Downing, 2010), a higher-resolution database is needed to accurately quantify the global carbon cycle.

Many global-scale water body databases have been developed in recent years (see Table 1). The SRTM Water Body Data (SWBD) (NASA/NGA, 2003) accurately captures water bodies at 1 arc-second resolution (about ~30 m at the Equator), but it does not cover the entire globe. Some large rivers in SWBD are disconnected by observational gaps which significantly reduce channel connectivity and therefore the utility of the database for hydrology studies. The Global Land Cover Facility (GLCF) MODerate resolution Imaging Spectroradiometer (MODIS) 250 m water mask (Carroll et al., 2009) has a global coverage, but the 250 m resolution is not adequate to resolve small channels or lakes. The GLCF MODIS water mask is considered to be a "snapshot" of circa-2000, so the temporal change in water bodies (such as potential inundation of floodplains) is not represented. Recently, using Landsat images globally, Feng et al. (2015) developed the GLCF Inland surface Water data (GIW) at 30 m resolution and Verpooter et al. (2014) developed Global Water Body data (GLOWABO) at 0.5 arc-second resolution. However, the temporal change of water bodies was not considered in previous high-resolution water body databases.

Some water body databases do consider temporal change in water extent. The Global Lake and Wetland Database (GLWD) (Lehner and Döll, 2004) used a classification of surface water types (e.g. river, lake, floodplain, wetland) and depicts the global distribution of each surface water type at ~1 km resolution. Prigent et al. (2007) and Papa et al. (2010) developed a 25 km resolution inundated area map (Global Inundation Extent from Multi-Satellite: GIEMS) with monthly temporal variations, though the 25-km resolution is not sufficient to depict individual rivers or lakes. Fluet-Chouinard et al. (2015) downscaled GIEMS to a 15 arc-second resolution. The downscaled product (GIEMS-D15) quantifies global water extent at mean annual minimum, mean annual maximum and long term maximum.

Table 1. Comparison of global surface water database

Product	Resolution	Coverage	Frequency of Water	Reference
SWBD	1 sec (~30 m)	N60-S54	No	NASA/NGA, 2003
GLCF MODIS	7.5 sec (~250 m)	N90-S90	No	Carroll et al., 2009
GLCF GIW	30 m	N81-S81	No	Feng et al., 2015
GLOWABO	0.5 sec (~15m)	N81-S56	No	Verpooter et al., 2014
GLWD	30 sec (~1 km)	N90-S60	Water type classification	Lehner and Doll, 2004
GIEMS	25 km (equal area)	N90-S90	Monthly flood extent	Papa et al., 2010
GIEMS-D15	15 sec (~500 m)	N90-S60 [a]	Mean annual max/min	Fluet-Couinard et al., 2015
G3WBM	3 sec (~90 m)	N81-S60	Multi-scene analysis	This Study

[a] Greenland is not included

A major problem in delineating a high-accuracy, high-resolution, water body map comes from the fact that water extent can change in time and space. Given that rivers, lakes, floodplains and wetlands show different characteristics in hydrodynamics, ecosystems and biogeochemistry, it is obviously better to separate permanent water bodies (e.g. low water river channels, lakes with permanent water coverage) and temporal water-covered areas (e.g. floodplains, wetlands, paddy fields) in water body mapping. For example, accurate delineation of low-water river channels (excluding floodplains) is important for improving flood forecasting by global-scale river models (e.g. Pappenburger et al.., 2012), and information on temporal dynamics of surface waters is valuable to estimate global wetland carbon inventory (e.g. Bridgham et al., 2013). Multi-temporal images are needed to carry out frequency analysis, but this significantly increases the quantity of data to be handled, especially when analysis is done at a high resolution. Due to this difficulty in data handling and processing, previous high-resolution water maps do not consider temporal change of water extent, and frequency of water body existence is only represented in low resolution databases.

1.2 Objective

The objective of this study is to develop a new high-resolution global water body map with information on the frequency of water body existence. An automated algorithm was developed to handle multi-temporal Landsat images at a global scale and to analyze the frequency of water body existence at 3 arc-second resolution (about 90 m at the equator). The algorithm was also designed to exclude observational gaps caused by cloud or ice/snow covers by compositing multiple satellite images. The Global 3 arc-second Water Body Map (G3WBM) was generated by applying the developed algorithm to Landsat images in the GLCF Global Land Survey (GLS) database (Gutman et al., 2013). The main aim is to generate a global permanent water body map (e.g. low water river channels, lakes with permanent water coverage) at 3 arc-second resolution, but as a consequence of defining the permanent water areas, additional information on temporal water-covered areas (e.g. floodplains, paddy fields) is included in the G3WBM.

2. Data

2.1 Landsat GLS database

As a starting point for the water body map delineation, we used the GLCF Landsat Global Land Survey (GLS) database (Gutman et al., 2013). The GLS database attempts to provide one cloudless image acquired at each location in World Reference System (WRS). One set of global-coverage images is prepared for 5 different epochs (i.e. GLS1975, GLS1990, GLS2000, GLS2005 and GLS2010). The GLS1975 consists of Landsat Multi-Spectral Scanner (MSS) images, while the other epoch collections are based on Landsat TM (Thematic Mapper) and ETM+ (Enhanced Thematic Mapper Plus) images. Landsat images with lesser cloud cover were selected for the GLS database, but they are not always perfectly cloud-free. Landsat GLS images can be downloaded freely from the GLCF website (http://glcf.umd.edu/data/gls/).

We used all TM and ETM+ images from the Landsat GLS database. A total of 33,890 scenes were used; 7,375 from GLS1990, 8,756 from GLS2000, 9,365 from GLS2005, and 8,484 from GLS2010. The water body map was calculated by combining information from 4 spectral bands and one thermal band; Band 2 (green: G), Band 3 (red: R), Band 4 (near infra-red: NIR), Band 5 (short wave infra-red: SWIR), and Band 6 (thermal infra-red). Band 1 (blue: B) was also used to generate RGB composites for the post-classification analysis and validation steps (see Sections 4 and 5). Please note that there are overlapping areas between adjacent Landsat paths, so that the number of available observations can be larger than 4. The overlapping areas become wider at higher latitude, so that the number of observations is larger in higher latitude (even though eastern Siberia is missing by GLS1990).

2.2 Digital Elevation Model

A Digital Elevation Model (DEM) was used to generate an ocean mask and also for distinguishing shadows from water bodies. We used the global 3 arc-second DEM downloaded from the Viewfinder Panoramas webpage (http://www.viewfinderpanoramas.org/dem3.html). The Viewfinder Panoramas DEM (hereafter VFP-DEM) was generated mainly using the Shuttle Radar Topography Mission 3 arc-second DEM (SRTM3 DEM) (Farr et al., 2007) for regions below 60N, but for above 60N uses the GeoBase DEM for Canada, and Russian topography maps for the Eurasian continent. Large voids (i.e. blank areas due to no data, usually found over water bodies and in mountain areas) in the original SRTM3 DEM were carefully filled in by the developer of the VFP-DEM using auxiliary topography information such as printed topography maps. However, some small voids remain in the distributed DEM and we fill these remaining voids by interpolation using the inverse square distance weighted method.

An ocean mask was generated by marking 0 m elevation pixels which are connected to outer oceans (i.e. 0 m pixels in inland areas are excluded from the ocean mask). In order to include river pixels with 0 m elevation in the analysis, coastline data from OpenStreetMap (available online from http://openstreetmapdata.com/data/coastlines) was also used to generate the ocean mask. Pixels in the ocean mask were excluded from water body classification because Landsat GLS images often have large amount of clouds over oceans and we are only interested in terrestrial water bodies.

Elevation gradient was calculated from the VFP-DEM in order to distinguish shadows from water bodies. Elevation gradient (in meter per pixel) was calculated as the maximum elevation difference between each pixel and its 8 neighboring pixels. If a pixel is lower than any of its neighbors, the elevation gradient was set to 0 m.

2.3 SWBD water mask

The Shuttle Radar Topography Mission Water Body Data (SWBD) (NASA/NGA, 2003) was additionally used in order to ensure channel connectivity in the delineated water body map (See Section 3.5). The SWBD was a byproduct of the SRTM3 DEM which was generated by C-band radar interferometer. Because the coverage of SWBD is between 60N and 56S, connectivity correction was not performed above 60N.

3. Method

3.1 Landsat image processing

Each Landsat image was converted to 3 arc-sec resolution (about ~90 m at the Equator) in the WGS84 grid coordination system by nearest point resampling. Given that the georeferenced error of GLS images was approximately 25 m (Gutman et al., 2013), the georeferenced error should not be a problem at the 3 arc-second resolution. Then, top of the atmosphere reflectance and brightness temperature were calculated from the digital number (DN) using the conversion method described by Chander and Markham (2003) and Chander et al. (2009). Reflectance of blue band ρ_B, green band ρ_G, red band ρ_R, near infra-red band ρ_{NIR}, and short wave infra-red band ρ_{SWIR} were calculated from the DN of bands 1, 2, 3, 4, and 5, respectively. Brightness temperature Tb (in centigrade) was calculated from the DN of band 6. Then, the Normalized Difference Water Index (NDWI) and the Normalized Difference Vegetation Index (NDVI) were calculated as follows:

$$NDWI = \frac{\rho_G - \rho_{SWIR}}{\rho_G + \rho_{SWIR}} \quad (1),$$

$$NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R} \quad (2).$$

We selected the Modified Normalized Difference Water Index (MNDWI) proposed by Xu (2006) from the many variations of NDWI methods (e.g. McFeeters, 1996; Ji et al., 2009). Landsat-7 images after 31st May 2003 have striping gaps due to the failure of the Scan Line Corrector (SLC) (Maxwell et al. 2007). The SLC gaps were filled by the interpolation method described in Appendix A1. Note that the resolution conversion and grid coordination change were done with the "gdalwarp" function of Geospatial Data Abstraction Library (GDAL) (Warmerdam, 2008), while other steps were calculated using Fortran90 codes originally developed by the authors. A schematic diagram of the developed algorithm is shown in Figure 1.

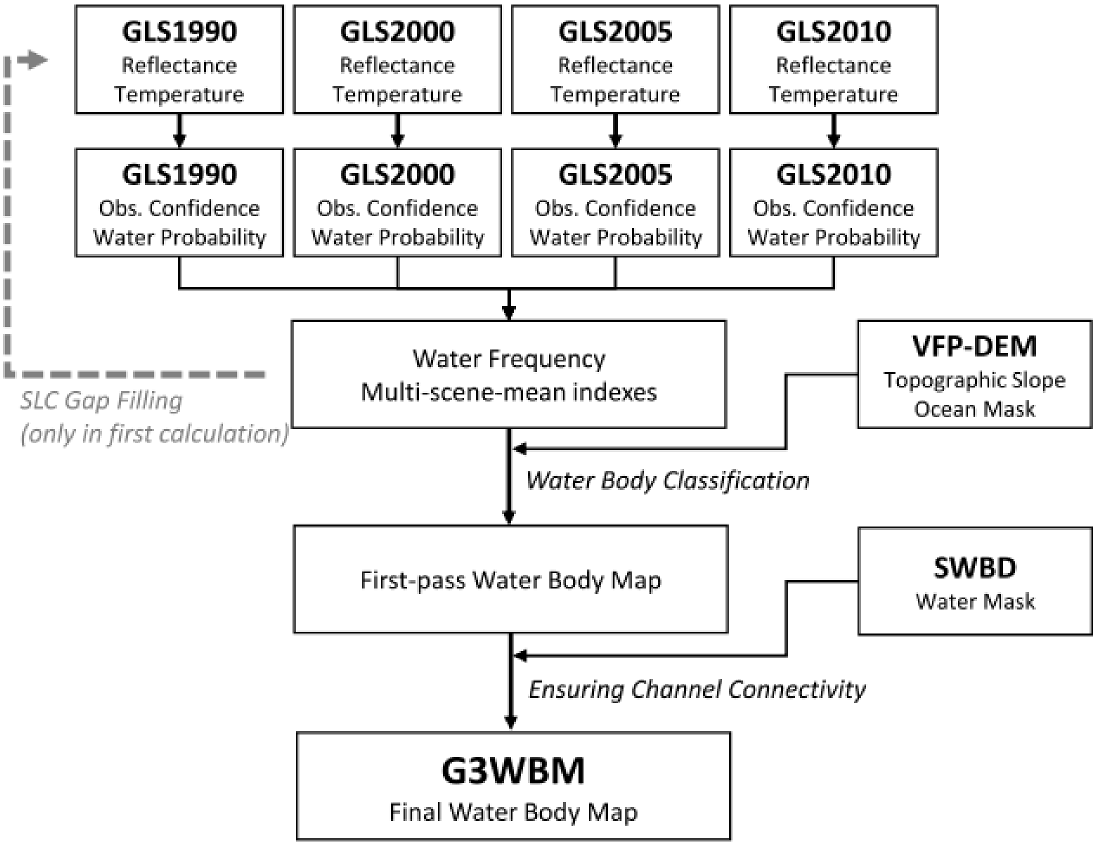

Figure 1: Schematic diagram of the developed algorithm. Note that more than 4 images are used where Landsat scenes are overlapped with adjacent scenes.

3.2 Frequency of Water Body Existence

The index "water frequency" was introduced to separate permanent water bodies, temporal water-covered areas, and land pixels. Water frequency Fw_i of pixel i was defined by the equation (3):

$$Fw_i = \frac{\sum_{j=1}^{N}\left(O_{i,j}W_{i,j}\right)}{\sum_{j=1}^{N} O_{i,j}} \quad (3),$$

where $O_{i,j}$ is observation confidence at pixel i in Landsat scene j, $W_{i,j}$ is water probability at pixel i in Landsat scene j, N is the total number of Landsat scenes available at pixel i. Equation (3) means that water frequency is calculated by the observation-confidence weighted average of water probability. Note that parameters in the following steps (e.g. thresholds and constants in classification conditions) were mainly taken from previous studies (e.g. Irish, 2000; Ji et al., 2009) but are adjusted by trial and error, based on validation.

<Observation Confidence>

Observation confidence $O_{i,j}$ represents the certainty of judging land surface type at pixel i in Landsat scene j. Observation confidence is 1 when land is judged to be perfectly observed without cloud or ice/snow cover, while it becomes smaller when land is not clearly observable (a minimum value was set to 0.001). Observation confidence is defined by Equation (4):

$$O_{i,j} = \max[1 - Pci, 0.001] \qquad (4),$$

where Pci is a probability index of cloud/ice existence at pixel i in Landsat scene j. The probability index Pci ranges from 0 (low cloud/ice probability) to 1 (high cloud/ice probability), and is given by the equation (5):

$$Pci = \frac{\min[\rho_{GRN}, 0.25]}{0.25} f_{NDLI} f_{Tb} \qquad (5),$$

where ρ_{GRN} is minimum reflectance of green, red and near infra-red bands, f_{NDLI} is a correction factor using Normalized Difference Land Index (NDLI, see Appendix A2) and f_{Tb} is a correction factor using brightness temperature. Given that cloud and ice/snow are highly refractive in visible and near infra-red bands, the probability of cloud/ice existence can be mainly judged by the minimum reflectance of the red, green and near infra-red bands ($\rho_{GRN} = \min[\rho_G, \rho_R, \rho_{NIR}]$). The correction factors f_{NDLI} and f_{Tb} were introduced to separate cloud/ice and highly-reflective rock/vegetation. Detailed explanations on these correction factors are outlined in Appendix A2.

<Water probability>

Water probability in Equation (3) was calculated using Equation (6):

$$W_{i,j} = P_{NDWI} f_{NDVI} \quad (6),$$

where P_{NDWI} is a probability index using NDWI and f_{NDVI} is a correction factor using NDVI. The probability index P_{NDWI} was given by the equation (7):

$$P_{NDWI} = \begin{cases} 0 & (NDWI < 0) \\ NDWI/0.3 & (0 \le NDWI \le 0.3) \\ 1 & (0.3 < NDWI) \end{cases} \quad (7).$$

When NDWI is larger than 0.3, the pixel is considered to be water and when smaller than 0, the pixel is considered to be land. For an NDWI between 0 and 0.3, the water body existence is represented by probability. Because shadows sometimes show high NDWI as water, the correlation function using NDVI f_{NDVI} was introduced to distinguish shadow and water. The detailed description of f_{NDVI} is summarized in Appendix A3.

3.3 Multi-scene mean indexes

In addition to water frequency Fw_i, multi-scene mean indexes (i.e. reflectance, NDWI, NDVI, and brightness temperature) are calculated. These multi-scene mean indexes were used for water mask classification in the water body classification step (see Section 3.4). For each index V (e.g. reflectance, NDWI, NDVI), the multi-scene mean was defined by Equation (8):

$$\overline{V_i} = \frac{\sum_{j=1}^{N} \left(O_{i,j} W_{i,j} V_{i,j} \right)}{\sum_{j=1}^{N} \left(O_{i,j} W_{i,j} \right)} \quad (8),$$

where $\overline{V_i}$ is the multi-scene-mean of index V at pixel i. $O_{i,j}$ is the observation confidence, $W_{i,j}$ is the water probability, $V_{i,j}$ is the target index, all at pixel i in Landsat scene j. N is the number of observation scenes available at pixel i. Multi-scene-mean indexes at pixel i are calculated for reflectance of each bands ($\overline{\rho}_{Bi}$, $\overline{\rho}_{Gi}$, $\overline{\rho}_{Ri}$, $\overline{\rho}_{NIRi}$, $\overline{\rho}_{SWIRi}$), minimum reflectance of green, red and near-infra-red bands ($\overline{\rho}_{GRN\,i}$), NDWI, NDVI and brightness temperature ($\overline{WI}_i, \overline{VI}_i, \overline{Tb}_i$).

3.4 Water Body Classification

The main aim of this study is to delineate a permanent water body map. For that purpose, water frequency was used to distinguish permanent water bodies from temporal water-covered areas. However, some land covers (e.g. ice, snow, salt marsh, wet soil, wet vegetation and shadow) show a high NDWI, thus they might mistakenly be classified as water (i.e. commission error). Therefore, other land cover types which showed similar characteristics to water, were excluded before classifying permanent water bodies and temporal water-covered areas. The flowchart of classification steps is shown in Figure 2.

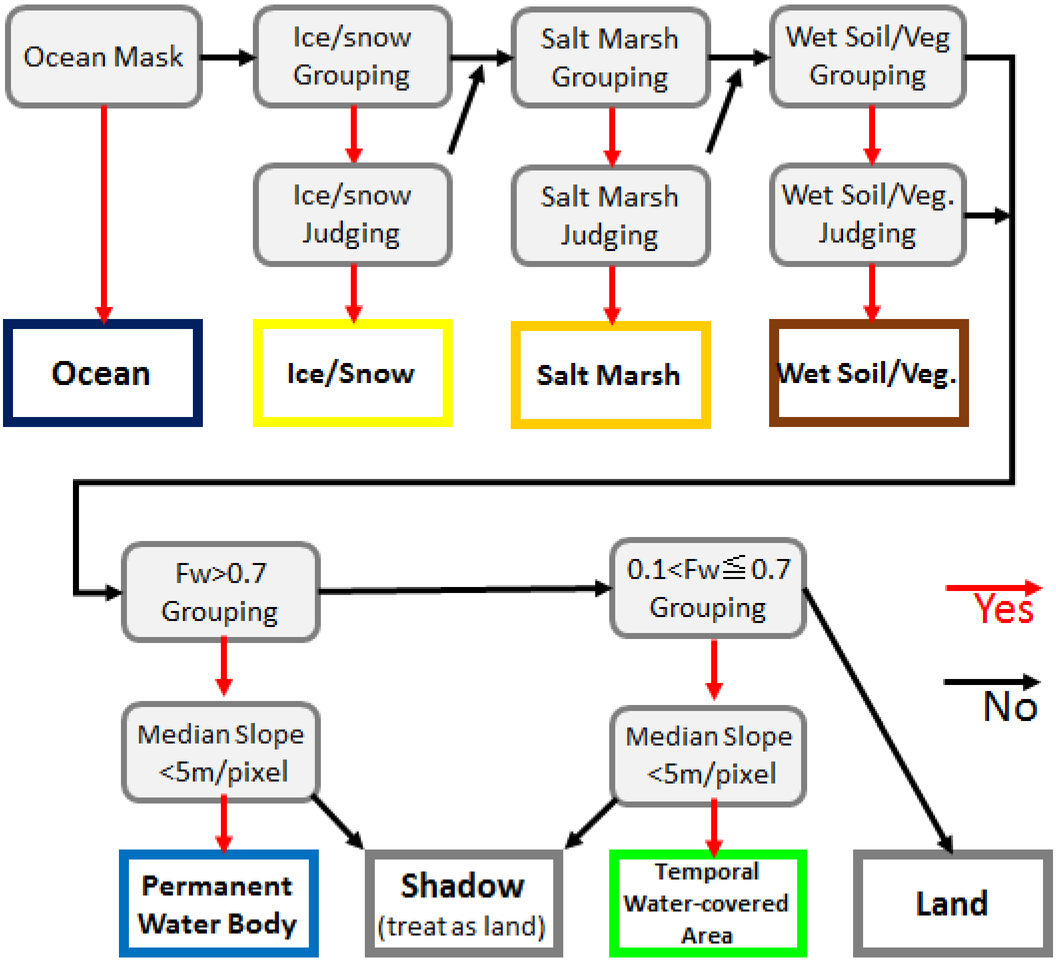

Figure 2. Flowchart of classification steps

<Exclusion of high-NDWI non-water surface>

Ice/snow, salt marsh, and wet soil/vegetation are excluded as high-NDWI non-water surface using the criteria listed in Table 2. In order to avoid generation of a patchy land type classification, adjacent pixels with similar characteristics were grouped using the grouping criteria. Then, one classification type was assigned to pixels in each group by the judging criteria. Group-mean index $\overline{V}^{g}$ was calculated from multi-scene-mean index $\overline{V_i}$ of pixels in each group as follows:

$$\overline{V}^{g} = \frac{\sum_{i=1}^{M} \overline{V_i}}{M} \qquad (9),$$

where M is the number of pixels in each group. Note that the thresholds in Table 2 were determined by trial and error, repeating water body classification and visual checking detailed in Section 3.6.

Table 2. Criteria used for high-NDWI non-water land classification.

	Ice/snow	Salt marsh	Wet soil/vegetation
Grouping Criteria	$Fw_i > 0.3$ $\overline{\rho}_{GRN\,i} > 0.15$ $\overline{Tb}_i < 2$ $\overline{WI}_i > 0.4$ $\overline{VI}_i > -0.2$	$Fw_i > 0.1$ $\overline{\rho}_{GRN\,i} > 0.25$ $\overline{Tb}_i > 0$ $\overline{WI}_i > 0.4$ $\overline{VI}_i > -0.2$	$\overline{\rho}_{GRN\,i} < 0.15$ $0.0 < \overline{WI}_i < 0.5$ $-0.15 < \overline{VI}_i < 0.3$
Judging Criteria	$\overline{\rho}_{GRN}{}^{g} > 0.2$ $\overline{Tb}^{g} < 0$ $\overline{WI}^{g} > 0.6$ $\overline{VI}^{g} > 0.2$	$\overline{\rho}_{GRN}{}^{g} > 0.35$	$\overline{WI}^{g} < 0.4$ $\overline{VI}^{g} > 0.05$

As ice and snow show a high NDWI similar to water, pixels with ice or snow cover were excluded before water body classification. Adjacent pixels whose multi-scene-mean indexes (water frequency Fw_i, minimum reflectance of green, red and near infra-red bands $\overline{\rho}_{GRN\,i}$, brightness temperature $\overline{Tb}_i$, normalized water index $\overline{WI}_i$ and normalized vegetation index $\overline{VI}_i$) satisfy the grouping criteria for ice/snow in Table 2 were grouped as potential ice/snow pixels. If group-mean indexes of the grouped pixels satisfy the judging criteria in Table 2, potential ice/snow pixels within each group were judged to be true ice/snow class. Pixels which were not classified as ice/snow were passed to salt marsh classification.

Salt marsh has a high reflectance in visible bands and has relatively low reflectance in the short wave infra-red band. Therefore, salt marsh shows a high NDWI even when it is not inundated. In order to distinguish dry salt marsh from true water bodies, pixels considered to represent salt marsh were excluded before water mask classification using the grouping and judging criteria in Table 2. Pixels which were not classified as salt marsh were passed to wet soil/vegetation classification.

Wet soil/vegetation sometimes shows a relatively high NDWI, even when the land surface is not inundated. In order to accurately delineate true water bodies, pixels which showed moderate NDWI and moderate NDVI were classified as wet soil/vegetation using the criteria in Table 2. Pixels which were not classified as wet soil/vegetation were passed to the following water body classification.

<Water Body Classification>

After excluding ice/snow, salt marsh, and wet soil/vegetation, the remaining pixels were classified as permanent water bodies, temporal water-covered areas and land, based on water frequency. However, pixels affected by shadows should be distinguished from water bodies because they sometimes show a high NDWI. Here, elevation gradient (calculated from the DEM, defined in Section 2.2) was used to distinguish water bodies from shadows. For permanent water body delineation, adjacent pixels which had $Fw_i > 0.7$ were grouped. In order to exclude mountain shadows, the grouped pixels were classified as permanent water body when more than half the group's pixels had an elevation gradient smaller than 5 m per pixel (i.e. relatively flat areas). Otherwise they were classed as shadow. Remaining pixels were passed to temporal water-covered area classification.

Adjacent pixels with $0.1 < Fw_i \leq 0.7$ were grouped. The grouped pixels were classified as temporal water-covered areas when the group-mean NDWI was larger than 0.5 and when more than half of the pixels had an elevation gradient smaller than 5 m per pixel. All remaining pixels were classified as land.

3.5 Ensuring Channel Connectivity

The proposed frequency analysis is based on the assumption that the river channel location is stable for long periods. Therefore, the developed algorithm is not applicable to river segments where channel position frequently changes with time. In these cases, river channels may be classified as temporal water-covered areas when channel positions are different in different Landsat scenes. In order to ensure flow connectivity of river channels, we overlaid the SWBD water mask onto the delineated water body map. For minimizing the correction amount, we only used SWBD water bodies which were larger than 100 km. Excessive modification of small lakes could be avoided by using this size threshold, and the connectivity correction could be restricted to large rivers. Pixels which were not classified as permanent water but are treated as water bodies in the SWBD, were changed to permanent water body pixels. It is reported that the SWBD also has gaps within water bodies (Carroll et al., 2009), but this did not cause a connectivity problem in this study because we confirmed that locations of water body gaps were not overlapping between the SWBD and the GLS images.

3.6 Visual Checking

We generated a JPEG image of the developed water body map at each 5 degree tile, and visually examined every image to check whether the classifications were appropriate or not. Images were visually checked for consistency and continuity, as well as against other spatial data and images available, such as Google and Bing satellite images. If critical errors were found, we revised the coefficients and thresholds used in the classification step (e.g. numbers in Table 2) and recalculated the entire water body map globally. This visual checking iteration method was repeated more than 10 times until all major misclassifications were eliminated. While much of the processing and analytical testing described in the results section are automated, this human visual step was important for identifying some of the subtle anomalies that can occur when trying to apply an automated method globally using multi-temporal images.

4. Results

Firstly, we demonstrated how the proposed algorithm removes clouds and calculates water frequency. A part of the Congo River (Landsat World Reference System 2 path 182 row 086;

E13.7-E15.9, S5.3-S3.4) was selected as an example because cloudy images were included. The algorithm was applied to four images available at path 182 row 086 (the acquisition dates of the four images are shown in Figure 3). The band 5-4-2 composite image of each GLS scene are shown in Figure 3a. In the band 5-4-2 composite image, vegetation, rock, water, cloud water (warm cloud) and cloud ice (cold cloud) are colored with green, red, dark blue, white and pale blue, respectively. All images contain cloud water and/or cloud ice which disrupt the earth observation. The observation confidence calculated by Equation (4) is shown in Figure 3b. Locations of pixels with low observation confidence (grey) agree well with observable cloud locations. Water probability (Figure 3c) was high for pixels representing open waters. Some pixels under cloud showed a high probability, but their impact in water frequency calculation is low because their observation confidence is low. In the GLS1990 image, some pixels along the Congo mainstem showed a low water probability because of cloud cover. However, water frequency at these pixels was not reduced by the cloud cover because observation confidence was also low when pixels were covered by clouds. Figure 3d illustrates water frequency calculated from the four images. While river and land were clearly distinguished, water frequency was slightly high in some pixels with cloud cover. These pixels were successfully judged to be land after applying the classification algorithm (see Figure 3e).

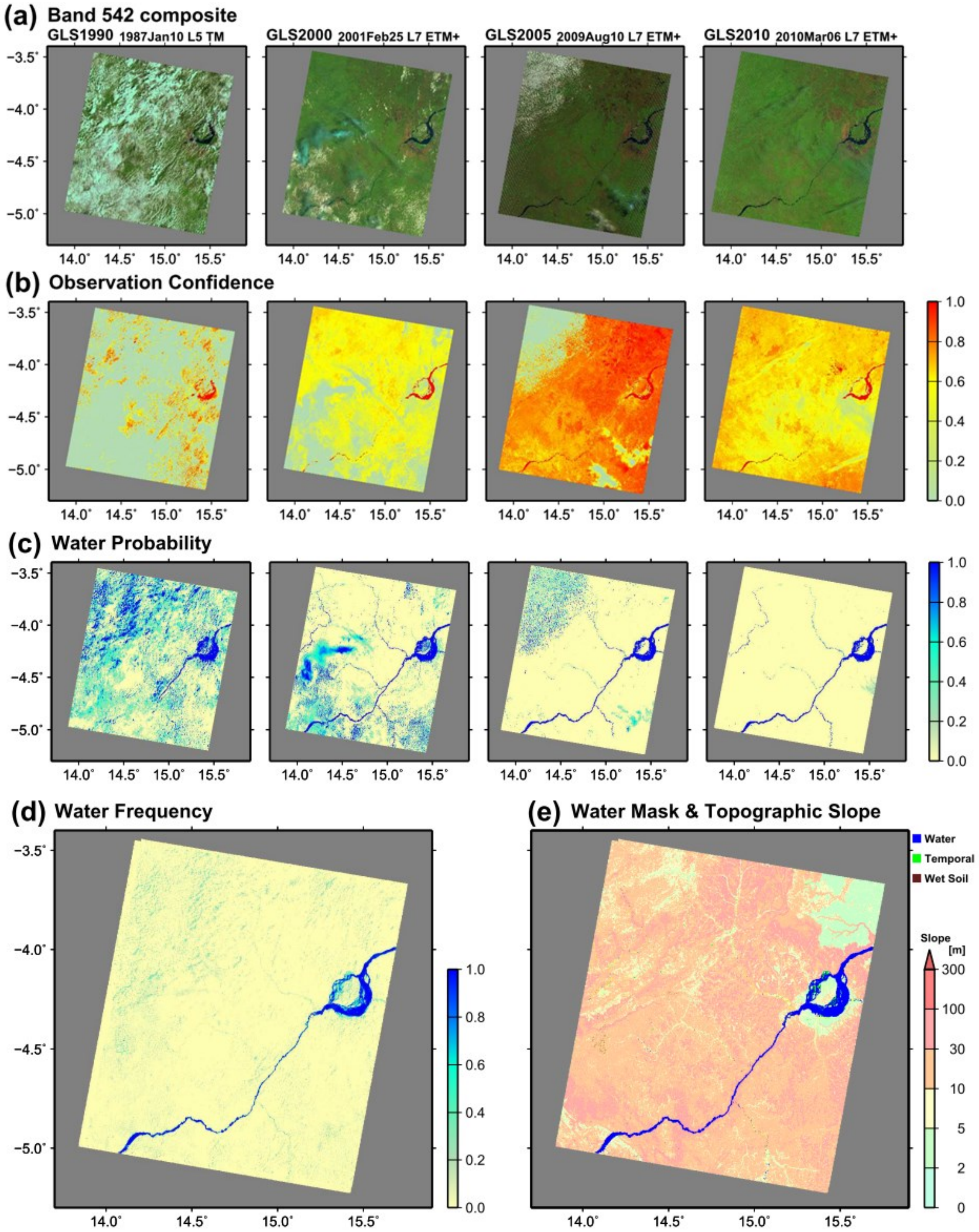

Figure 3: Example of classification procedure, shown for the Congo River basin. (a) Band 5-4-2 composite images, (b) Observation confidence, (c) Water probability, (d) Water frequency calculated from the four images, (e) Water mask classification. In Figure 3e, blue and green represent permanent water bodies and temporal water-covered areas, while background colors represent elevation gradient of land pixels.

From here on, we show the results of applying the developed algorithm to all GLS images at the global scale. Results of selected regions are shown in this paper, but images of other regions can be accessed online (http://hydro.iis.u-tokyo.ac.jp/~yamadai/G3WBM/).

In order to show the differences between the new and previous water body maps and the importance of using multi-temporal images, classification results in floodplains along the Amazon and Ob Rivers are shown in Figure 4. Figure 4a illustrates water bodies of the GLCF GIW in the Amazon floodplain (around W54.6, S2.3). Generally, the GLCF GIW accurately captures water bodies, but some observation gaps due to cloud covers were found. Conversely, no observation gaps were found in the new water body map (Figure 4b) because clouds were removed by overlaying multi-temporal images. Figures 4c and 4d illustrate water bodies in a downstream reach of the Ob River. Because water frequency was not considered, river channels and floodplains were lumped together in the GLCF GIW. Conversely, river channels and floodplains were represented separately in the new water body map. Given that river channels and floodplains show different bathymetry, flow dynamics and ecosystems characteristics, we believe river-floodplain separation in the new database will be useful in various research fields.

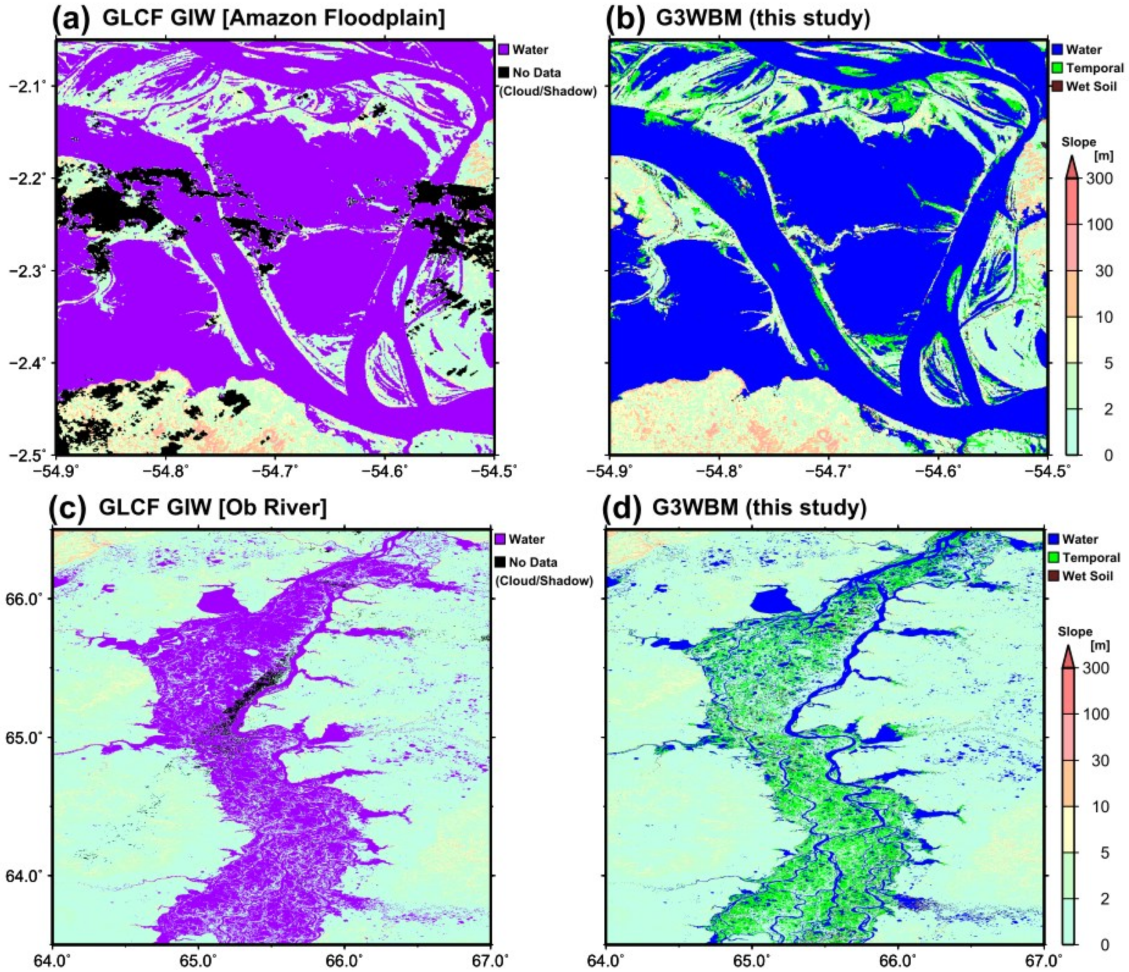

Figure 4: Water bodies represented in GLCF GIW (a, c) and in G3WBM (b, d). In (a, c) observation gaps due to cloud and cloud shadow are shown by black. In (b, d), permanent water bodies are represented by blue, while temporal water-covered areas are shown by green.

Then, we validated the accuracy of the land type classification (described in Section 3.4) by checking the results in regions where water, wet soil and snow cover are coexistent. A region of southern Iceland was selected for this purpose. The RGB composite image and land type classification results are shown in Figure 5. The RGB composite was created by taking the minimum reflectance from multiple scenes in order to remove temporal cloud or ice/snow covers. In general, the G3WBM accurately captured water bodies, including those most commonly challenging e.g. lakes on glacier edges (e.g. W19.85, N64.60 and W17.35, N64.15), lakes above wet lava rocks (e.g. areas around W18.8, N64.1) and lakes in the valleys between high mountains (e.g. W21.45, N64.5). The developed algorithm also succeeded in excluding rocks on glaciers (e.g. W17.2 N64.2) which sometimes show similar reflectance characteristics to water. However, boundaries between permanent water bodies and surrounding wet soils were not clearly represented in coastal areas, probably because of a shallow groundwater table (e.g. rivers around (W17.5, N63.8). The classification in wet coastal regions is difficult because both water body and wet soil show a high NDWI and also because the topography is flat. It should be noted here that the proposed method was designed for accurate delineation of permanent water bodies, and therefore, there might be some misclassification between ice, wet soil and temporal water-covered areas (e.g. rocks on a glacier were classified as temporal water-covered areas at W17.35, N64.40).

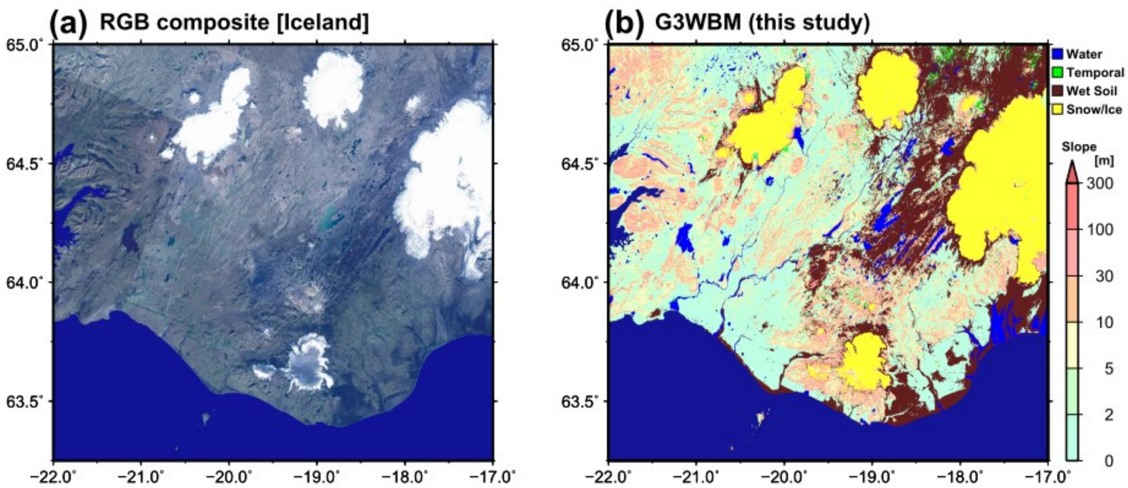

Figure 5: Southern Iceland region with the coexistence of water, wet soil and snow cover. (a) RGB composite image, (b) and land type classification. Permanent water bodies, temporal water-covered areas, wet soil, and ice/snow are represented by blue, green, brown and yellow, respectively.

Figure 6 illustrates a middle reach of the Ganges River, as an example of regions where channel connectivity correction using the SWBD (Section 3.5) was required. Temporal change of channel locations from 1992 to 2009 was calculated from eight GLS images (i.e. GLS1990, GLS2000, GLS2005 and GLS2010 images for path 140 rows 042-043). Because the channel location is not stable, most river segments were classified as temporal water-covered areas (green pixels in Figure 6b). Disconnected channels are not ideal for hydrological research, therefore a flow connectivity correction was performed using the SWBD water mask. Here we can see that gaps in permanent water bodies (see discontinuous channel colored with blue in Figure 6b) were filled by the SWBD water masks (red in Figure 6b). Please note that the SWBD represents the water extent in February 2000, therefore, some water bodies of the SWBD might correspond to temporal water-covered areas. This might result in a small overestimation of permanent water bodies in the modified water body map, but we decided that the benefit of ensuring channel continuity is greater than the disadvantage of the overestimation. Channel connectivity correction was only performed below 60N coinciding with the SWBD coverage. The modification was not required above 60N as almost no connectivity problems were found in boreal regions. Of course, for geomorphological studies, the changing channel location information observed over the GLCF epochs may be of significant value in its own right.

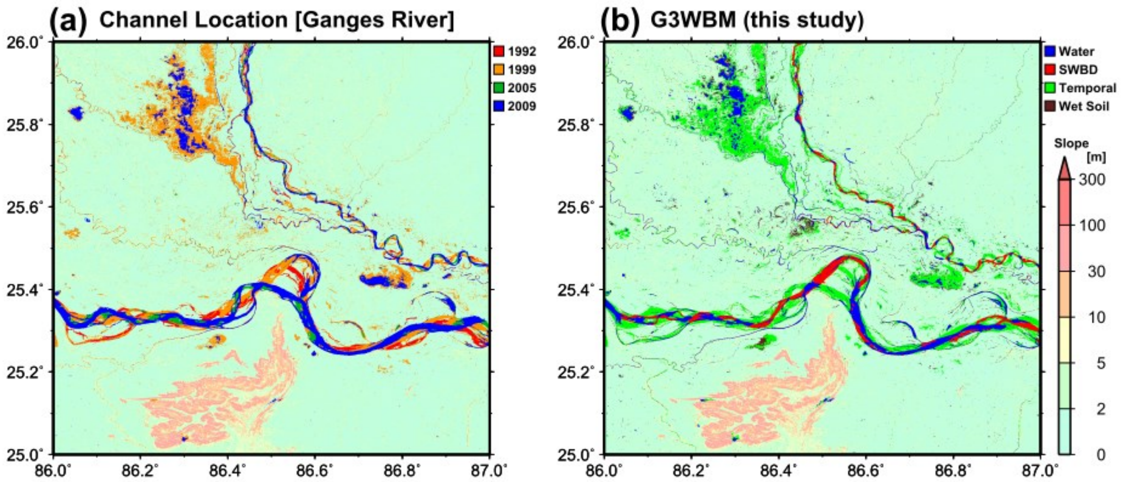

Figure 6: (a) Channel location change in the Ganges River and (b) result of water body classification. Channel locations in 1992, 1999, 2005 and 2009 are shown by red, orange, dark green and blue, respectively in (a). Permanent water bodies and temporal water-covered areas are shown by blue and green in (b), while overlapping SWBD water mask is shown in red.

A global distribution of permanent water bodies is illustrated in Figure 7. The percentage of permanent water bodies within 0.01 degree grid boxes is shown in Figure 7a, whilst the zonal/meridional total water body areas per 1 degree latitude/longitude bin are shown in Figures 6b and 6c. In total, 3.25 million km^2 were classified as permanent water bodies, which is about 2.4% of the total inland area (including the Caspian Sea). The aggregation of smaller water bodies occasionally occupy more than 20% of the area of 0.01 degree grid boxes (grids colored with blue or dark blue in Figure 7a). It can be seen that relatively small water bodies, i.e. lake size <100 km^2 (purple lines in Figures 6b and 6c) are concentrated mainly in boreal regions (i.e. Canada, Scandinavia, Finland, West and North Siberian plain, Kolyma and Indigirka River basins). The second peak is in the Tibetan Plateau. This distribution pattern is consistent with previous studies (e.g. Lehner and Döll, 2004; Fluet-Chouinard et al., 2015). With the exception of boreal regions and the Tibetan Plateau, the zonal total water body area is dominated by very large water bodies (black lines in Figures 6b and 6c), such as the North American Great Lakes, the Caspian Sea, lakes in the African Rift Valley, and the Amazon and Congo Rivers.

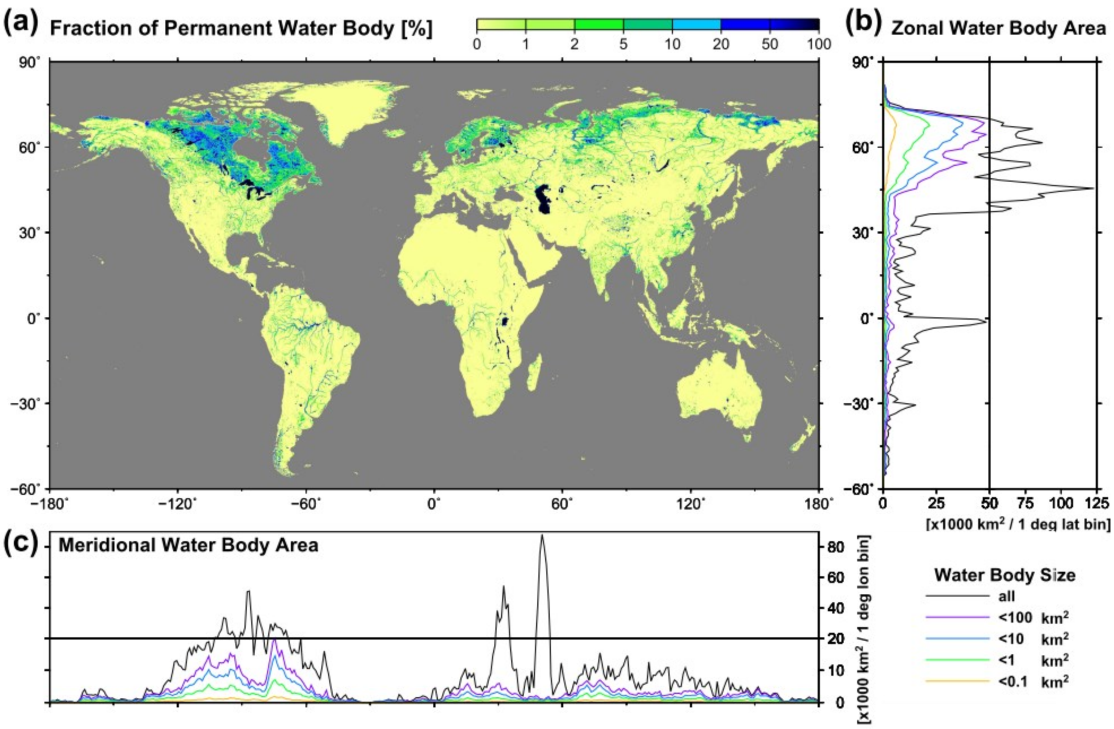

Figure 7: Global distribution of permanent water bodies. (a) Fraction of permanent water body within 0.1 degree grid boxes. (b) Zonal and, (c) meridional total water body area per 1 degree latitude/longitude bin.

5. Discussion

5.1 Importance of multi-temporal analysis

The global distribution of temporal water-covered areas is illustrated in Figure 8. Temporal water-covered areas are concentrated in large river floodplains (e.g. the Ob, Lena, Amazon and Ganges Rivers), boreal climate regions (e.g. northern Canada and northern Siberia) and arid regions (e.g. inland Australia and Central Asia). Temporal water-cover in boreal regions is likely to be dominated by snow melt because the acquisition timing of the GLS images in the boreal region were June to August (Gutman et al, 2013), which overlaps with the snow melt season in high latitude zones (Armstrong et al., 2005). The area around the Aral Sea is classified as a temporal water-covered area because it had been covered by open water in earlier GLS images but was dried up in later images. The global summation of temporal water-covered areas was 0.49 million km^2, about 15% of the global permanent water body area (see Table 3). These areas could be misclassified as permanent water bodies if only flooded images were used in the water body mask development.

Global total area for each land classification type was calculated for the G3WBM and the GLCF GIW (Feng et al., 2015) and summarized in Table 3. Among the 3.81 million km^2 of inland water area in the GLCF GIW, 2.92 million km^2 is classified as permanent water in the G3WBM but other areas are classified as temporal water-covered area (0.17 million km^2) or as non-water surface (0.71 million km^2). The classification discrepancy between water and land (i.e. water in one database but land in the other) is probably due to the difference in spatial resolutions. However, the 0.17 million km^2 areas treated as water in the GLCF GIW but temporal water-covered area in the G3WBM was detected because multi-temporal scenes were used in the G3WBM. This indicates the importance of the water frequency analysis in creating global water body maps. Furthermore, the GLCF GIW includes the 3.61 million km^2 of no-data areas which are mainly due to cloud or cloud shadow. Though most of the no-data areas are considered to be land, the GLCF GIW missed some true water bodies (as shown in Figure 4a).

This also illustrates the importance of using multi-temporal images for eliminating gaps in water body mask.

Please note that the water frequency analysis in this study was mainly performed to delineate an accurate permanent water body mask by excluding temporal water-covered areas. The developed method did not intend to accurately delineate all temporal water-covered areas on the earth. Whether temporal water-covered areas could be detected or not, is decided by images used in the analysis. If flood images are not included in the GLS database, it is, of course, not represented as a temporal water-covered area in the developed database.

Table 3. Confusion matrix of global inland area classification between G3WBM and GLCF GIW.

Global total area [x1000 km^2]		GLCF GIW			
		Water	Land	No Data	**Total**
G3WBM (this study)	Permanent Water Body	2,924	245	72	**3,240**
	Temporal Flood Area	176	303	14	**493**
	Other Land Types	709	126,998	3,520	**131,228**
	Total	**3,809**	**127,545**	**3,606**	

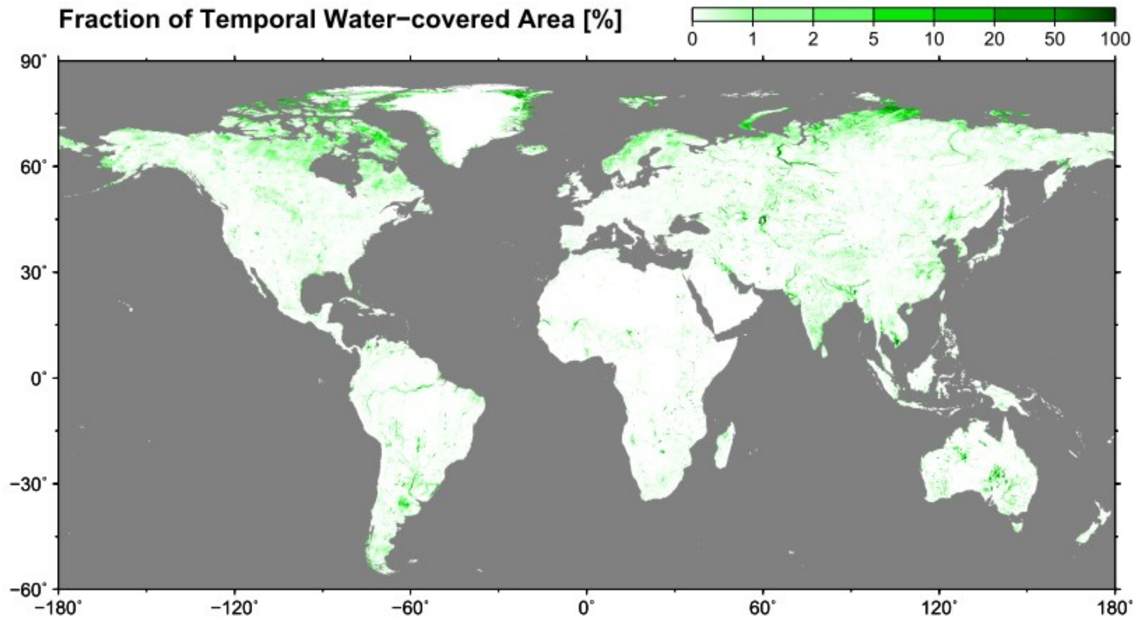

Figure 8: Global distribution of temporal water-covered areas. Percentage of temporal water-covered areas in each 0.1 degree grid box is shown.

5.2 Accuracy of water body detection

We performed extensive validation on Hokkaido Island, northern Japan (Figure 9a) to estimate the accuracy of the water body detection. The delineated water body area in the validation domain was 472 km^2, about 0.6% of the area of Hokkaido Island (77,984 km^2). The accuracy of excluding non-water land types from actual water body area was first examined. We checked whether each water body in G3WBM corresponded to actual rivers/lakes or not, by plotting all permanent water bodies larger than 0.05 km^2 onto topographical maps and space/airborne photos using Google Maps. There were 280 water bodies larger than 0.05 km^2 in the G3WBM (circle and square plots in Figure 9). We found only one exception which did not correspond to an actual water surface (the dark green square in Figure 9a). The commission error was located at the caldera of Taisetsu-zan Mountain (E142.88, N43.68), where snow cover and wet lava soil exist together. All water bodies, except for this commission error, corresponded to rivers and lakes on the topographic map or space/airborne photos. This result suggested that the proposed method can accurately distinguish water-like land type (e.g. snow, wet soil, shadow) from actual water bodies. Therefore, the overestimation of water bodies in the G3WBM was anticipated to be very small.

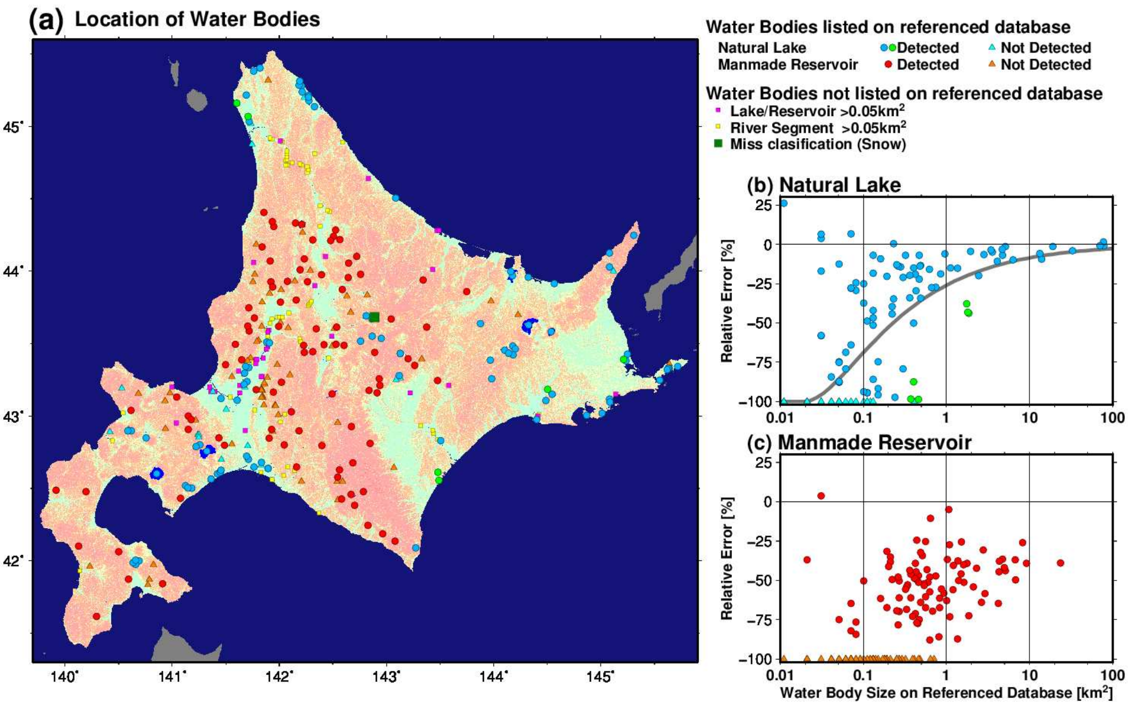

Figure 9: (a) Location of water bodies on Hokkaido Island, Japan. (b) and (c) Relative error of detected natural/manmade water bodies to lake/reservoir database. Grey lines in (b) represent an expected error due to coastline misclassification for circle-shape lakes.

We then compared the surface area of individual natural lakes found using our newly developed map against an existing database. We used the GIS database of natural lakes developed by Hokkaido Research Organization (available at http://envgis.ies.hro.or.jp/). All lakes registered in the GIS database (117 lakes excluding lagoons which were treated as ocean in the G3WBM) were used in the comparison. The size of referenced lakes varied from 77.76 km^2 to 0.01 km^2. Among the 117 referenced lakes, 94 lakes were detected in the G3WBM (blue and green circles in Figure 9) but 23 lakes were missed (light blue triangles in Figure 9). The relative error of each lake area is plotted on Figure 9b. It was found that surface areas were underestimated in most lakes, except for very small ones which consisted of two or three pixels. Lakes with relatively large errors (green circles in Figure 9) were found to be shallow marsh with a large surface area variation (Penketo Marsh; E141.71, N45.07, Shirarutoro Marsh; E144.50, N43.18, and Oikamanai Lake; E143.45, N42.56) and lakes covered by dense aquatic vegetation (Kimonto Marsh; E143.48, N42.61, Kanekinto Marsh; E145.21, N43.39, and Junsai Marsh; E141.60, N45.16). The discrepancy in shallow marsh with large surface area variation can be explained by the timing of observations. Underestimation in vegetated lakes was caused by the mixing of water and vegetation. As pixels covered by aquatic vegetation (such as water lily) have lower NDWI and higher NDVI than open water, most of them were judged to be wet soil/vegetation. Underestimation in other lakes can be mainly explained by the treatment of shoreline pixels. Given that water and land are mixed in shoreline pixels, they tend to be judged as wet soil because of their moderate NDWI. The ratio of shoreline pixels to all lake pixels becomes larger for smaller lakes, thus a larger underestimation was observed in smaller lakes in Figure 9b. The underestimation expected from shoreline misclassification was calculated by assuming all coastline pixels of a circle-shaped lake had been judged to be non-water. The expected error (the gray line in Figure 9b) well explain the actual underestimation ratio for each size class. Similar to small lakes, rivers narrower than one pixel size (about 90 m at the equator) were not well represented in the G3WBM.

We also compared the lake surface area of manmade reservoirs from the G3WBM against an existing dam database. All manmade reservoirs constructed before 2000 as listed in the handbook of Japanese dams (Japan Dam Association, 2014) were used in the comparison. In the delineated water body map, 93 out of 152 reservoirs were represented in the G3WBM (red circles in Figure 9) but 59 reservoirs were missing (orange triangles in Figure 9). The relative error of lake surface area is plotted on Figure 9c. Similar to natural lakes, smaller reservoirs generally showed a larger underestimation of area. However, the underestimation ratio was larger in manmade reservoirs than natural lakes. This is probably because reported lake areas in the dam handbook denote surface areas at maximum storage capacity. Actual dam storage is usually smaller than the maximum capacity, therefore the reservoir surface areas are likely to be underestimated using observations, and this makes classification accuracy of manmade reservoir areas in Hokkaido particularly challenging. Another reason for underestimation may be due to the locations of manmade reservoirs. While natural lakes tend to be located in flat regions, manmade reservoirs are generally created in mountainous areas. As we used elevation gradient of the DEM for separating shadow from water bodies, some reservoirs were mistakenly classified as shadows. In order to improve the classification accuracy, higher resolution topographic data is needed, especially in mountainous area.

The accuracy of water body delineation was also validated using high quality geospatial data of the contiguous United States. The U.S. Geological Survey National Hydrography Dataset (NHD) (Simley and Carswell, Jr., 2009) (Figure 10a) was selected for this purpose. Given that the NHD is high-resolution vector data of water bodies based on the U.S. topography maps, its accuracy is considered to be adequate for validation of satellite-derived water maps. The NHD "waterbody" and "river stream area" polygons were converted to a 1 arc-second raster, and then water body areas were compared between the NHD and the G3WBM (Figure 10b). The NHD polygons were converted to a higher resolution raster (1 arc-sec) than the G3WBM because they were treated as "truth" data for validation purpose. It was found that large water bodies (red in Figure 10a) were well represented in the G3WBM (yellow color in Figure 10b), while most small water bodies (<0.1 km^2, pale violet colors in Figure 10a) were not captured (white in Figure 10b). Overestimation of water body area was limited to some flood prone regions (red colored area in Figure 10b), so that commission error is expected to be very small in the G3WBM.

Figure 10c shows the relative water area error of 80,312 water bodies in the NHD whose size is larger than 0.1 km^2. Blue dots represent "waterbody" features in the NHD database (i.e. lakes, ponds, and reservoirs), while green dots represent "river stream area". It was found that about 70% of water bodies >1 km^2 show relative water area error smaller than 25%. Similar to the case of Hokkaido Island (Figure 9b), underestimation error was larger for smaller water bodies.

In general, lakes, ponds, and reservoirs are better delineated compared to river stream areas. In order to analyze why water body area was underestimate, the ratio of shoreline pixels to water body pixels within each water body was calculated and plotted against the relative water area error (Figure 10d). A strong relationship was observed between the shoreline pixel ratio and the relative water area error (the red line in Figure 10d). This suggests that the underestimation of water body area is mainly due to the difficulty of classifying shoreline pixels on water-land boundaries. Thus, narrow river segments or lakes in mountainous valley regions are not well represented in the G3WBM because they have relatively long shorelines compared to their water body size. We visually checked the location of 1,466 water bodies >1 km^2 whose relative error is more than 20% larger than the shoreline pixel ratio (i.e. below the orange line in Figure 10d). It was found that these large errors mainly correspond to water bodies with large surface area fluctuation (e.g. floodplains, salt marshes, and reservoirs with frequent water level change). Given that the proposed algorithm was designed to detect only permanent water bodies, this underestimation was expected because temporal change of water body area was not included in the NHD. However, we also found that some omission errors were caused by vegetation coverage over permanent water bodies (e.g. swamps, algae blooms). In order to further improve water body detection accuracy, classification of vegetated water bodies should be considered, in addition to a better shoreline classification.

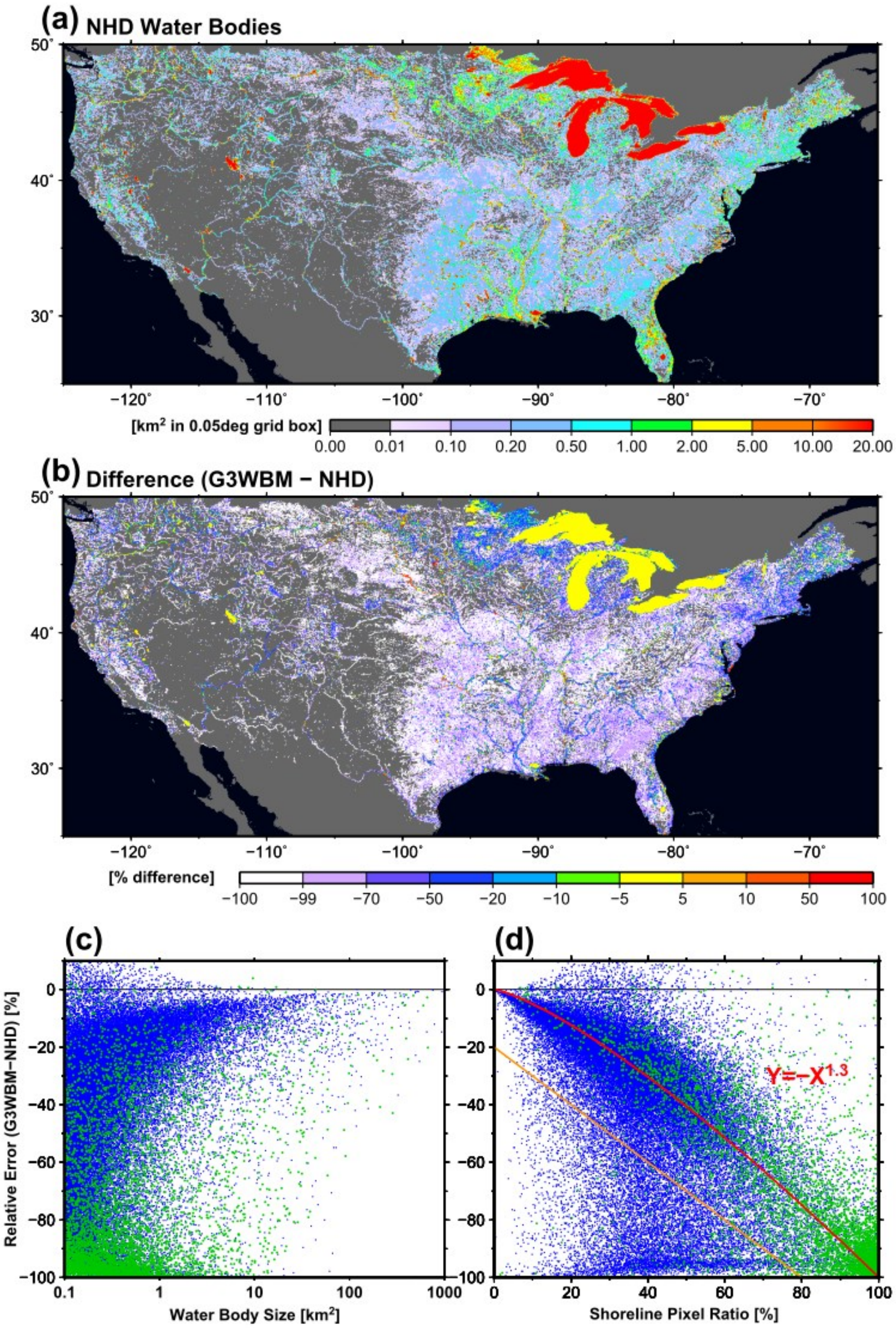

Figure 10: Distribution of water bodies in the NHD (a) and the difference between the G3WBM and the NHD (b). For visualization purpose, the resolution of water body map was converted to 0.05 degree. Scatter plots of relative water area error versus water body size (c) and versus shoreline pixel ratio (d). Blue and green dots represent "waterbody" and "stream area" features in the NHD database.

5.3 Comparison of Global Water Body Area

The total inland water body area (both river and lakes) in the constructed water mask database was 3.25 million km^2 (including the Caspian Sea at 0.36 million km^2). The total water surface area of large water bodies (>100 km^2) accounts for 63% of the total inland water body area, and the remaining water surface is accounted for by smaller water bodies. Lakes smaller than 0.1 km^2, 1 km^2 and 10.0 km^2 account for 4%, 13 % and 26% of the total inland water body area, respectively. Given that most of the small lakes (<0.1 km^2) are not well represented in the constructed database, the global total water extent is expected to be underestimated. Note that the size of a pixel is 0.0085 km^2 at the Equator and 0.0043 km^2 at 60 degree north/south.

Global total water body areas are compared between the 6 water body datasets in Table 4. In the case of the products based on optical sensors (i.e. MODIS and Landsat), the global water body area generally increases with the resolution of the product because smaller water bodies are detected at higher resolution. However, the global total water body area is also affected by the coastline definition and treatment of temporal water-covered areas (e.g. floodplains, salt marsh). For example, the G3WBM shows a relatively smaller area compared to other products because floodplains and salt marsh are excluded from permanent water bodies. The global water body area of the GLOWABO (5.37 million km^2) (Verpooter et al., 2014) is significantly larger than other products based on Landsat or MODIS (between 3.25 and 3.65 million km^2). This is probably because small water bodies are represented at the 0.5 arc-second resolution, but without direct comparison of the products, the exact reason for this large difference is unclear. The global water body area of the GIEMS-D15 (Fluet-Chouinard et al. 2015) is the largest among the all databases, probably due to the downscaling procedure. The GIEMS-D15 was generated by downscaling 25-km resolution water extent data (Papa et al., 2010) onto a 15 arc-second topography, which may cause over-representation of rivers and lakes smaller than the 15 arc-second pixel size (about 500 m at the equator). The statistical estimate by Downing et al. (2006) was analyzed to be an overestimation (McDonald et al., 2012), so that the global water estimations by the three databases (GLCF MODIS, GLCF GIW, and G3WBM) are considered to be consistent.

Table 4. Comparison of global total water body area between 5 databases.

Water body database	Source	Resolution	Global water area	Cutoff Threshold
GLCF MODIS[a]	MODIS	7.5 sec	3.29 million km^2	N/A
G3WBM (permanent water)	Landsat	3 sec	3.25 million km^2	1 pixel (~0.008 km^2)
GLCF GIW	Landsat	30 m	3.65 million km^2	5 pixels (~0.005 km^2)
GLOWABO[b]	Landsat	0.5 sec	5.37 million km^2	0.002 km^2
GIEMS-D15 (annual min)	Multi-satellite	15 sec	6.5 million km^2	N/A
Downing et al. 2006[c]	Statistical	-	4.2 million km^2	0.001 km^2

[a] The water body area of GLCF MODIS was calculated by the authors because it's not available in the description paper (Carroll et al., 2009).

[b] The water body area of the Caspian Sea was added to the GLOWABO for comparison.

[c] Downing et al. excluded river water surface from global water body area.

5.4 Possibility of further improvement

We restricted the resolution of G3WBM to 3 arc-second due to limitations in human and computational resources. However, given that the original resolution of Landsat images is about 1 arc-second, developing a global 1 arc-second water body map with water frequency information is certainly possible. Given that the underestimation of lake area in G3WBM is likely due to omission of shoreline water pixels, the accuracy of water classification is anticipated to increase in a higher-resolution water body map.

Even with the large number of scenes used, not all seasonal or extreme flood events will be captured in the water map developed here. Part of the reason for this will be due to the fact that cloud free scenes from leaf-on growing seasons were selected in the GLS collection, meaning that all the GLS images utilized come from the same season and therefore may "miss" flooding in other seasons. Given that temporary water bodies are hotspots for biodiversity and biogeochemical processes, accurate estimate of global temporal water extent is essential in this regard. Including future images as they become available, as well as broadening the number of images processed to include non-GLCF scenes may help reduce these issues, but will bring a higher computation cost and may increase observational uncertainty. Furthermore, while most temporal water-covered areas represent seasonal flooding, some represent long-term trends (such as shrinking of the Aral Sea, construction of dams, disappearing water bodies in Alaska and Siberia). Separation of seasonal and long-term water body change may be important for estimating global dynamics of surface waters. In addition, it's better to use topography data consistent with the time of Landsat image acquisition, because the elevation gradient from the DEM is used for water body classification.

The classification algorithm could be further improved to capture water bodies more accurately. We used the classification criteria with globally-constant thresholds (as in Table 3), but the threshold could be different in different regions. For example, sediment-rich and/or turbid water tends to show lower NDWI and higher NDVI than sediment-free water, so that some sediment-rich rivers are misclassified as wet soil in G3WBM (e.g. small tributaries of the Indus River). Vegetated water surface (e.g. lakes with algal blooms, floating plants) has similar characteristics to sediment-rich water, thus it is difficult to be detected by classification criteria. Using variable thresholds (e.g. Feng et al., 2015) based on local reference water body data may be a good solution for improving classification accuracy. Shoreline pixels with land and water mixing are likely to be omitted as wet soil because they have a lower NDWI than pure water pixels. Applying an additional classification step for shoreline pixels, after determining water body pixels, may improve the overall accuracy of water body mapping because mixed shoreline pixels are considered to be a major source of water area underestimation.

We did not applied atmospheric correction in this study in order to reduce computational requirements. The previous study by Verpooter et al. (2014) argued that atmospheric correction is not necessary for global water body mapping. Given that the GLCF GLS database consists of mostly-cloudless Landsat images from leaf-on growing seasons, atmospheric conditions are expected to be similar between different images and this probably decreases the need in atmospheric correction. However it should be better to remove the inconsistency due to atmospheric conditions, especially when multi-temporal images are used (Song et al., .2001). Application of atmospheric correction is therefore a possible strategy for improving the accuracy of water body detection in the G3WBM.

6. Conclusion

We developed the Global 3 arc-second Water Body Map (G3WBM) using 33,890 multi-temporal Landsat GLS images. In addition to the conventional land/water classification used in a previous global water body database, we separated permanent water bodies from temporal water-covered areas by calculating the frequency of water body existence from multi-temporal images. The G3WBM identified 3.25 million km^2 of permanent water bodies in the global inland areas, while the global total of temporal water-covered areas was 0.5 million km^2 (~15% of the global permanent water body area). The abundance of temporal water-covered areas suggests the importance of water frequency analysis using multi-temporal images. From the Comparison to a 30-m resolution water body map (the GLCF GIW), we concluded that the use of multi-temporal images is as important as analysis at a higher resolution for depicting global-scale dynamics of surface water bodies.

The accuracy of water body delineation was validated using space/airborne photos and the existing database of waterbodies in Hokkaido (Japan) and in the contiguous United States. There was almost no commission error of water bodies in the G3WBM, which suggests that the proposed classification algorithm has a very high accuracy. The areas of small lakes in the G3WBM tend to be underestimated, mainly due to the mixing of land and water in shoreline pixels, however the accuracy will be improved if a water body map is generated at higher resolution. Given that the proposed method is automated, it is not impossible to generate a global water body map at 1 arc-second (~30 m) or higher resolutions.

The G3WBM is distributed free of charge for research and educational purposes. Please visit the product webpage (http://hydro.iis.u-tokyo.ac.jp/~yamadai/G3WBM) to get access to the database.

Acknowledgements

This research is funded by “JSPS Grant-in-Aid for Scientific Research #26889077”. A part of the computational resources is also supported by “JSPS Grant-in-Aid for Scientific Research #23226012”. Mark Trigg’s contributions were completed under funding provided by the Willis Research Network.

References

Allen, G.H. & Pavelsky T.M. (2015). Patterns of river width and surface area revealed by the satellite-derived North American River Width data set, *Geophysical Research Letters.*, 42, 395–402, doi:10.1002/2014GL062764.

Armstrong, R., Brodzik M., Knowles K., & Savoie M. (2005). Global Monthly EASE-Grid Snow Water Equivalent Climatology. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center.

Bridgham, S.D., Cadillo-Quiroz H., Keller J.K., & Zhuang Q., (2013). Methane emissions from wetlands: Biogeochemical, microbial, and modeling perspectives from local to global scales. *Global Change Biology*, 19(5), 1325–1346.

Carroll, M.L., Townshend J.R., Di Miceli C.M., Noojipady P., & Sohlberg R.A. (2009), A new global raster water mask at 250 m resolution, *International Journal of Digital Earth*, 2(4), 291-308, doi:10.1080/17538940902951401.

Chander, G. & Markham B.L. (2003). Revised Landsat-5 TM radiometric calibration procedures, and post-calibration dynamic ranges. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 2674−2677, doi: 10.1109/TGRS.2003.818464

Chander, G., Markham B.L., & Helder D.L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113, 893−903, doi:10.1016/j.rse.2009.08.011.

Cole, J. J., Prairie Y.T., Caraco N.F., McDowell W.H., Tranvik L.J., Striegl R.G., Duarte C.M., Kortelainen P., Downing J.A., Middelburg J.J., & Melack J. (2007). Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget, *Ecosystems*, 10 (1), 172–184, doi: 10.1007/s10021-006-9013-8.

Downing, J.A., Prairie Y.T., Cole J.J., Duarte C.M., Tranvik L.J., Striegl R.G., McDowell W.H., Kortelainen P., Caraco N.F., Melack J.M., & Middelburg J.J. (2006). The Global Abundance and Size Distribution of Lakes, Ponds, and Impoundments, *Limnology and Oceanography*, 51 (5), 2388-2397.

Downing, J. A. (2010). Emerging global role of small lakes and ponds: little things mean a lot, *Limnetica*, 29 (1), 9-24.

Downing, J.A., Cole J.J., Duarte C.A., Middelburg J.J., Melack J.M., Prairie Y.T., Kortelainen P., Striegl R.G., McDowell W.H., & Tranvik L.J. (2012). Global abundance and size distribution of streams and rivers, *Inland Waters*, 2 (4), 229-236

Farr, T. G., et al. (2007). The Shuttle Radar Topography Mission. *Review of Geophysics*. 45, RG2004, doi:10.1029/2005RG000183.

Feng, M., Sexton J.O., Channan S., & Townshend J.R. (2015). A global, high-resolution (30 m) inland water body dataset for 2000: first results of a topographic-spectral classification algorithm, *International Journal of Digital Earth*, *published online*, DOI:10.1080/17538947.2015.1026420

Fluet-Chouinard, E., Lehner B., Rebelo L.M., Papa F., & Hamilton S.K. (2015). Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data, *Remote Sensing of Environment*, 158, 348-361, doi:10.1016/j.rse.2014.10.015

Gutman, G., Huang C., Chander G., Noojipady P., & Masek J.G. (2013). Assessment of the NASA-USGS Global Land Survey (GLS) datasets, *Remote Sensing of Environment*, 134, 249–265, doi: 10.1016/j.rse.2013.02.026.

Irish, R. (2000). Landsat 7 automatic cloud cover assessment: Algorithms for multispectral, hyperspectral, and ultraspectral imagery, *Proceedings of SPIE*, 4049, 348–355.

Japan Dam Association (2014). Dam Nenkan 2014 (in Japanese), Tokyo.

Ji, L., Zhang L., & Wylie B. (2009). Analysis of dynamic thresholds for the normalized difference water index, *Photogrammetric Engineering and Remote Sensing*, 75 (2009), 1307–1317.

Lehner, B. & Döll P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands, *Journal of Hydrology*, 296(1–4), 1–22.

Maxwell, S.K., Schmidt G.L., & Storey J.C. (2007). A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images, *International Journal of Remote Sensing*, 28(23), 5339–5356. doi:10.1080/014311601034902.

McDonald, C.P., Rover J.A., Stets E.G., & Striegl R.G. (2012). The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of global lake extent, *Limnology and Oceanography*, 57 (2),597-606, doi: 10.4319/lo.2012.57.2.0597.

McFeeters, S.K. (1996). The use of Normalized Difference Water Index (NDWI) in the delineation of open water features, International Journal of Remote Sensing, 17(7),1425–1432.

NASA/NGA (2003). SRTM Water Body Data Product Specific Guidance, Version 2.0, available online: http://dds.cr.usgs.gov/srtm/version2_1/SWBD/SWBD_Documentation/

Oki, T. & Kanae S. (2006). Global hydrological cycles and world water resources, *Science*, 313, 1068 – 1072, doi:10.1126/science.1128845.

O'Loughlin, F., Trigg M.A., Schumann G.J.-P., & Bates P.D. (2013). Hydraulic characterization of the middle reach of the Congo River, *Water Resources Research*, 49, 5059–5070, doi:10.1002/wrcr.20398.

Palmer, S.C.J., Kutser T., & Hunter P.D. (2015). Remote sensing of inland waters: Challenges, progress and future directions, *Remote Sensing of Environment*, 157, 1-8, doi: 10.1016/j.rse.2014.09.021.

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*, 115, D12111, doi:10.1029/2009JD012674.

Pappenberger, F., Dutra E., Wetterhall F., & Cloke H.L. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade, *Hydrology and Earth System Science*, 16, 4143-4156, doi:10.5194/hess-16-4143-2012.

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*, 112, D12107, doi:10.1029/2006JD007847.

Sampson, C.C, Smith A.M., Bates P.D., Neal J.C., Alfieri L., & Freer J.E. (2015). A high-resolution global flood hazard model, *Water Resources Research*, published online, doi:10.1002/2015WR016954

Sjögersten, S., Black C.R., Evers S., Hoyos-Santillan J., Wright E.L., & Turner B.L. (2014). Tropical wetlands: A missing link in the global carbon cycle?, *Global Biogeochemical Cycles*, 28, 1371–1386, doi:10.1002/2014GB004844.

Song, C., Woodcock C.E., Seto K.C., Lenney M.P., & Macomber S.A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, 75, 230-244, doi:10.1016/S0034-4257(00)00169-3.

Verpooter, C., Kutser T., Seekell D.A., & Tranvik L.J. (2014). A global inventory of lakes based on high-resolution satellite imagery, *Geophysical Research Letters*, 41, 6396 – 6402, doi:10.1002/2014GL060641.

Warmerdam, F. (2008). The Geospatial Data Abstraction Library, Open Source Approaches in Spatial Data Handling, Springer, 87-104.

Xu H (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, *International Journal of Remote Sensing*, 27 (14), 3025-3033, doi:10.1080/01431160600589179.

Yamazaki, D., O ' Loughlin F., Trigg M.A., Miller Z.F., Pavelsky T.M., & Bates P.D. (2014a). Development of the global width database for large rivers, *Water Resources Research*, 50, doi:10.1002/2013WR014664.

Yamazaki, D., Sato T., Kanae S., Hirabayashi Y., & Bates P.D. (2014b). Regional flood dynamics in a bifurcating mega delta simulated in a global river model, *Geophysical Research Letters*, 41, doi:10.1002/2014GL059744.

Appendix

A.1 SLC gap filling

Landsat7 images after 31st May 2003 have striped gaps due to the failure of the Scan Line Corrector (SLC). This could result in striping patterns in water body classification, thus we removed the SLC gaps by the following interpolation method. After calculating water frequency, Fw, using multiple Landsat images (Section 3.2), pixels with Fw>0.1 were marked as potential water bodies. Remaining pixels were marked as potential land areas. Then, SLC gap filling was applied for each Landsat scene, using this extra information from non-SLC-gap scenes. If a pixel within an SLC gap was a potential water body in non-SLC-gap scenes, reflectance values were copied from its nearest water body pixel (outside the gap), and for potential land areas the nearest potential land pixel reflectance was copied. This interpolation is based on the assumption that reflectance values must be similar within adjacent water body pixels or within adjacent land pixels. Then, water frequency and multi-scene mean indexes were recalculated using the gap-filled Landsat images. Water mask classification was carried out with these recalculated indexes.

A.2 Correction factor for observation confidence

The correction factor f_{NDLI} in equation (5) was introduced to distinguish highly reflective vegetation/rock from cloud or ice/snow. The Normalized Difference Land Index (NDLI) was defined as follows:

$$NDLI = \frac{\min[\rho_G, \rho_R] - \max[\rho_{NIR}, \rho_{SWIR}]}{\min[\rho_G, \rho_R] + \max[\rho_{NIR}, \rho_{SWIR}]} \qquad \text{(A1).}$$

Given that vegetation and rock have relatively low reflectance in visible bands compared to infra-red bands, land shows higher NDLI than cloud and ice/snow. The correction factor f_{NDLI} was calculated by equation (A2):

$$f_{NDLI} = \begin{cases} 1 & (NDLI < 0) \\ 1 - NDLI \times 2 & (0 \le NDLI < 0.5) \\ 0 & (0.5 \le NDLI) \end{cases} \quad \text{(A2).}$$

No correction is made on the probability index Pci if f_{NDLI} equals 1, while probability of cloud/ice existence becomes lower when f_{NDLI} is smaller.

The correction factor using brightness temperature f_{Tb} in equation (5) was applied to improve the accuracy of cloud/ice detection. Given that cloud and ice/snow are relatively cold, brightness temperature of pixels covered by cloud or ice/snow is expected to be low. The correction function f_{Tb} was defined separately for ice/snow and cloud. Given that ice has very low reflectivity in the short wave infra-red band, ice shows higher NDWI than cloud. We assumed that NDWI smaller than 0.3 represents cloud. The correction function f_{Tb} was given as follows:

$$\begin{aligned} & if\,(NDWI < 0.3); \\ & \quad f_{Tb} = \begin{cases} 1 & (Tb < 25) \\ (30 - Tb) \times 0.2 & (25 \le Tb < 30) \\ 0 & (30 \le Tb) \end{cases} \\ & if\,(NDWI \ge 0.3); \\ & \quad f_{Tb} = \begin{cases} 1 & (Tb < 0) \\ (5 - Tb) \times 0.2 & (0 \le Tb < 5) \\ 0 & (5 \le Tb) \end{cases} \end{aligned} \quad \text{(A3).}$$

We assumed that pixels are not likely to represent cloud and ice/snow when the brightness temperature was higher than 25 degrees centigrade and 0 degrees centigrade, respectively. No correction was made on the probability index Pci if f_{Tb} equals 1, while probability of cloud/ice existence becomes lower when f_{Tb} is smaller.

A.3 Correction factor for water probability

Because shadows sometimes show a high NDWI similar to water, the correlation function using NDVI f_{NDVI} was used to modify water probability in equation (6). The reflectivity of water is very low in both near infra-red and short wave infra-red bands, while the reflectivity of shadow is not as low as water in near infra-red band. Therefore, pixels with high NDWI and high NDVI are potentially affected by shadows. The correction function f_{NDVI} was given by Equation (A4):

$$f_{NDVI} = \begin{cases} 1 & (NDVI < 0.1) \\ (0.2 - NDVI) \times 10 & (0.1 \le NDVI \le 0.2) \\ 0 & (0.2 < NDVI) \end{cases} \quad \text{(A4).}$$

No correction is made on P_{NDWI} when f_{NDVI} is 1, P_{NDWI} is reduced when f_{NDVI} is smaller than 1. As non-vegetated areas have a low NDVI, the correction function f_{NDVI} is expected to identify shadow well in vegetated areas, but may be less useful in detecting shadow in non-vegetated areas.