Geographical Information Management and Applications

MSC THESIS

EVALUATING THE CONSISTENCY AND ACCURACY OF WETLAND MAPPING IN GLOBAL AQUATIC LAND COVER PRODUCTS

February 26th, 2021

Name Virginia De Melo Ferreira De Sousa

E-mail v.demeloferreiradesousa@students.uu.nl

Supervisors Nandika Tsenbazar & Panpan Xu (WUR)

Professor Ron van Lammeren (WUR)

Universiteit Utrecht
TU Delft
University of Twente.
Wageningen UR

Acknowledgements

This thesis was written as part of the MSc. program Geographical Information Management and Applications. The research started in September 2020 and ended in February 2021 under the guidance of my supervisors Nandika Tsendbazar and Panpan Xu, whom I thank for their immense patience and valuable feedback.

I would like to thank Ruben for his daily support and for helping me when I struggled with R coding. Special thanks to my fellow GIMA student Chia-Yun for working alongside me on Microsoft Teams every weekday with my cat Grethe keeping us company on my desk. Lastly, I thank my family and friends for their support.

Abstract

Wetlands are valuable ecosystems that are under threat due to the impact of human activities and climate change. Assessing the distribution of wetlands and their extent is relevant to manage these environments. However, the lack of a universal definition and the use of different input data and methodologies to map wetlands can complicate the assessment. This study aims to evaluate the consistency and accuracy of three selected global products: GLWD-3, LC-CCI, and CW-TCI. Specifically, it defines the consistency on a thematic, spatial, and temporal level, while the accuracy assessment was completed with an independent integrated reference dataset. A comparison among the products was completed to determine their consistency to a global extent, and a regional analysis was included to assess the consistency of the global products in wetlands hotspots such as the Amazon and Congo Basins with regional datasets. The accuracy assessment showed an overrepresentation of wetlands for all maps at the global extent. Furthermore, the results reveal some definition inconsistencies that created disagreement for LC-CCI, while CW-TCI was considered unreliable for mapping wetlands due to the high disagreement and low classification accuracies. GLWD-3 had better definition consistency but missed mapping wetland areas in the Amazon and Congo Basins while having a high disagreement and overestimation for the class wetlands. As a result of the assessment, despite the overrepresentation of wetlands by all products, LC-CCI is more suitable for contemporary studies and GLWD-3 is an old dataset that did not perform better than the other products, while CW-TCI is not a reliable dataset. Based on these results, it is necessary for users to have a prior understanding of the wetland definition before choosing an appropriate product that meets their needs. The LCCS-based aquatic land cover characterization framework presented recently is recommended as a possible solution to improve the differences between aquatic land cover and generic land cover mapping. Lastly, researchers need to continue their assessments on products to highlight disagreements and determine their consistency over time.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Context

Wetlands are considered valuable ecosystems, providing ecological and socioeconomic benefits such as regulating regional climate, providing a habitat to a great variety of species, and being a source of recreational opportunities (Gallant, 2015; Wu, 2017). However, they are also one of the most threatened environments due to the increase in human demand to expand agricultural and urban areas (Wu, 2017). Other than the impact of human activities, climate change is affecting the hydrologic cycle of wetlands due to the rising sea level and carbon sequestration (Aires et al., 2018). It was estimated a drastic reduction of 64% of global wetlands since 1900, with a rate that varies from country to country (Wu, 2017). The Amazon Basin is an example of an area with high-risk wetlands composed of a complex system with numerous types of wetlands (Wittmann & Junk, 2016). Second to the Amazon in size is the Congo River Basin, which has a wide variety of freshwater species, and it has seen an increase in unsustainable management of its resources (Harrison, Brummett, & Stiassny, 2016). Other relevant areas are represented by the river valleys of the Mississippi, Ganges-Brahmaputra, and Yangtze, but extensive wetlands can also be found in arctic and boreal regions (Tiner, 2009).

To manage wetland environments, it is vital to assess the distribution and the extent of these ecosystems. However, the assessment is complicated by the lack of a universal definition of wetlands, which is due to the presence of different classification schemes that aim to reflect the variety of wetlands (Wu, 2017). Indeed, wetland areas can refer to aquatic habitats, transitional areas between wet and dry environments, or even open-water ones in some definitions (Wu, 2017).

Wetland maps and inventories are required for wetland preservation and management (Wu, 2017). Field monitoring provides valuable information on environmental factors and the consequences of changes that affect the functioning of the ecosystems (Gallant, 2015). However, it is an expensive method that is limited by access to remote locations and it only allows to monitor areas to a smaller extent (Gallant, 2015). On the other hand, Geographical Information System (GIS) and remote sensing are effective solutions for monitoring wetlands to a global extent and with repeated observations while being cost-effective (Hu, Niu, & Chen, 2017; Wu, 2017). Remote sensing-based datasets are one type of product that can be selected to map wetlands. One product is the global land cover dataset such as the Land Cover Climate Change Initiative with 300 m spatial resolution (LC-CCI, Defourny, Bontemps, Lamarche, Brockmann, & Boettcher, 2017). Compilation datasets are a second option that compiles historical data regarding water, vegetation, and soil, with wetland-related datasets such as wetland maps and land cover products (Hu et al., 2017). An example of a commonly used product is Global Lakes and Wetland Database Level-3 (GLWD-3), which combines water bodies and various wetland types in a global raster map at around 1 km resolution (Lehner & Döll, 2004). As a supplement of the former two types of wetland-related datasets, simulation datasets aim to replicate the extent and distribution of wetlands by modeling potential wetland areas using hydrological models or Land Surface Models (LSM) (Hu et al., 2017). An example of such dataset is the groundwater-driven wetland (GDW) map of the Composite Wetland (CW) map which is created by three maps of topography-climate wetness index (TCI) and water table depth estimates (Tootchi, Jost, & Ducharne, 2019).

The distribution and extent of wetland areas are relevant to different fields that involve the monitoring of this land cover type, such as climate modeling, hydrological modeling, and agricultural management (Xu, Herold, Tsendbazar, & Clevers, 2020). Moreover, wetland datasets are critical for applications such as simulation of species distribution, exploration of migration patterns, and establishment of natural reserves (Hu et al., 2017). These examples represent a part of a large variety of users that require wetland maps. However, choosing the right product according to specific requirements, and the application it is needed for, is not a straight-forward decision.

Furthermore, the type of classification used is strongly related to the purpose of the dataset. Xu et al. (2020) divided 33 global aquatic land cover (GALC) datasets into four groups; inundation/extent datasets are dynamic products that provide a delineation of flooded areas, with a coarse spatial resolution; Global Land Cover (GLC) datasets have limited classes related to water, which underrepresents the complexity of aquatic land cover; single-type GALC datasets have a finer spatial resolution and are thematically specific since they present one type of aquatic land cover, which is not suitable for various user needs; lastly, multi-type GALC datasets have several aquatic land cover classes, but they have a coarse spatial resolution and are outdated.

1.2 Problem statement

Although remote sensing is an effective technique to map wetlands, it presents difficulties in terms of consistency that allows the monitoring of change (Gallant, 2015). Indeed, knowledge on the status of wetlands is limited and global datasets present disagreement on their distribution (Hu et al., 2017). This occurs due to the numerous definitions of wetlands that create inconsistencies among wetland classes, and because different input data and methodologies are applied to map wetlands, which produce different results (Hu et al., 2017). Therefore, it is necessary to explore inconsistencies among these datasets on a thematic, spatial, and temporal level.

Another issue is the lack of an extensive assessment for determining the accuracy of inundation/extent datasets and multi-type GALC datasets that creates uncertainties about their quality (Xu et al., 2020). Indeed, users may need more information to determine the right product for a specific research application (Xu et al., 2020). Datasets such as LC-CCI, CW-TCI, and GLWD-3 are commonly used by users, even though they provide little information on the accuracy based on independent validation data. Furthermore, identifying wetlands is a challenging land cover mapping task and it is often performed at a regional scale (Hu et al., 2017). On the other hand, accurate global datasets are rare and scattered among land cover and water products, peatland, and mangrove datasets (Hu et al., 2017).

Therefore, a comparison among these products is necessary to assess their consistency and accuracy, which would guide users in choosing the most suitable one according to the purpose of use and requirements. Other than the comparison on a global extent, two regions will be looked into since they have a large ecosystem of wetlands: Amazon Basin and Congo Basin. The analysis would provide further information on the consistency of the global products on a regional extent.

1.3. Objective and research questions

The main objective of this research is to evaluate the consistency and accuracy of three commonly used global wetland products. The following products were chosen to represent Inundation/Extent datasets, GLC datasets, and multi-type GALC datasets respectively: CW-TCI, LC-CCI, and GLWD-3. The CW-TCI dataset was chosen because it records the largest extent of wetlands, while GLWD-3 has been widely used by multiple user groups even though it could be considered outdated. LC-CCI was selected to represent a GLC dataset that does not have an extensive number of classes related to wetlands and water.

The main objective can be achieved by answering the question:

"To what extent are the three selected global wetland products consistent and accurate:

CW-TCI, LC-CCI, and GLWD-3?"

Furthermore, the following questions will be answered to satisfy the main objective:

1. To what extent are they spatially, thematically, and temporally consistent among each other on a global extent?
2. How accurate are they when compared with an integrated validation dataset on a global extent?
3. What insights can the Amazon Basin and Congo Basin regions provide to the global extent consistency and accuracy analysis?

1.4 Reading guide

The following chapters provide an outline of the research, starting with Chapter 2 that includes the material and methods selected to perform the analysis. Indeed, the aforementioned chapter presents products, reference data, and a detailed methodology for each research question. Chapter 3 consists of the results of the analysis, which are then discussed in Chapter 4. Lastly, recommendations and conclusions are drawn in Chapter 5.

2. Materials and Methods

2.1 Research workflow

The objective of this research is to determine how consistent and accurate the three selected products are. Figure 1 illustrates the workflow of the study. To answer the first research question defined in Chapter 1.3, a comparative analysis first assessed the consistency of products in three aspects: thematic, spatial, and temporal. As different products rely on various definitions of wetlands that result in the implementation of different classification schemes, there is an inconsistency of land cover types and their spatial distribution (Nakaegawa, 2012; Xu et al., 2020). On the other hand, products refer to different time frames, and that has to be accounted for when a comparison is conducted. To compare the products on a thematic and spatial level, a legend harmonization was first required. The UN Land Cover Classification System (LCCS, Gregorio, 2016) is the protocol chosen to convert different legends into a uniform one. The harmonization allowed comparing the products since the differences and inconsistencies of the legends were highlighted (Herold, Mayaux, Woodcock, Baccini, & Schmullius, 2008). Therefore, a disagreement map was developed to reveal how spatially consistent the maps are.

The second part of the research focuses on the accuracy of the products since not all of them were validated rigorously before (Xu et al., 2020). Indeed, GLWD-3 and CW-TCI maps have a Stage 0-1 according to the validation hierarchy of Committee on Earth Observation Satellites (CEOS, Land Product Validation Subgroup, 2003), which means that the product was either not validated or only assessed from a small set of samples. On the other hand, LC-CCI has Stage 3 since the product was assessed with a larger set of samples and periods on a global scale (Xu et al., 2020). Therefore, an accuracy assessment was performed with a reference dataset created by combining three different ones, after a legend harmonization. The final dataset was created to assess the products by creating a confusion matrix that allowed the calculation of their overall accuracy, producer's accuracy, and user's accuracy. The calculations were performed in R.

The third objective is to determine the consistency of the products on a regional extent. Amazon Basin and Congo Basin were selected as areas of interest for a more in-depth analysis. Indeed, the two regions were evaluated in more detail, along with regional datasets to determine the spatial consistency of the global products when applied to regional areas.

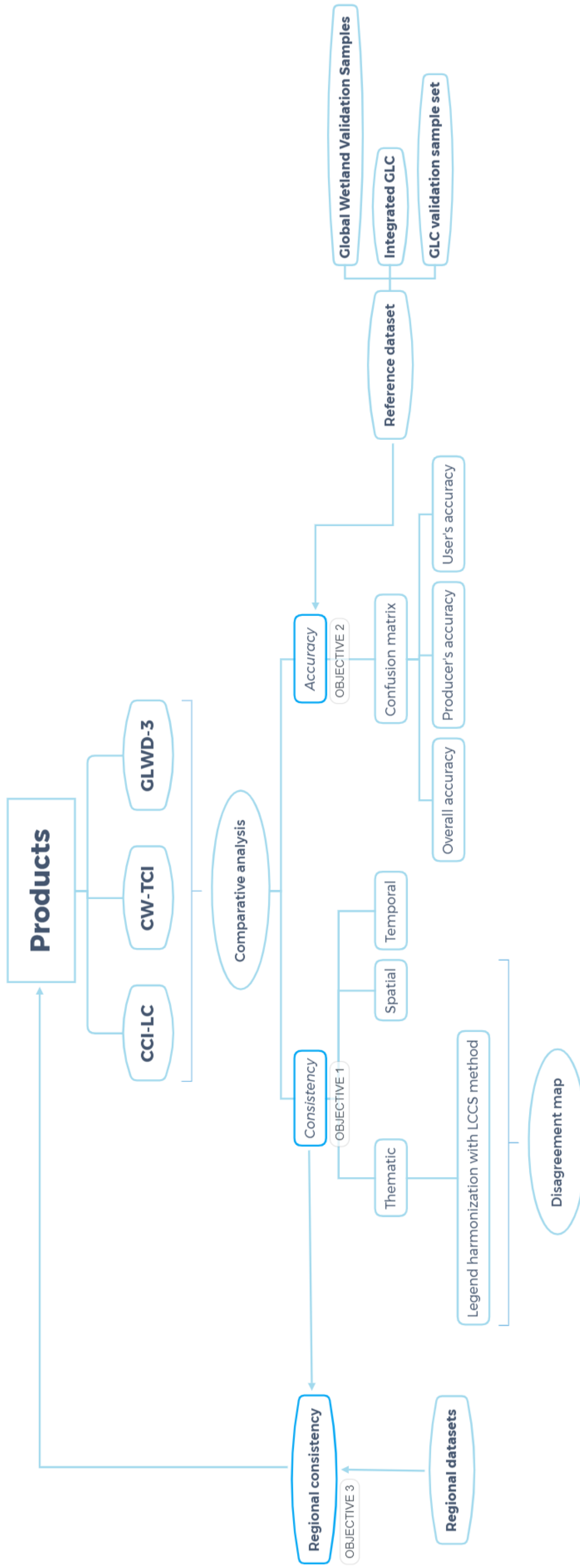


Figure 1: Research workflow

2.2 Data

2.2.1 Global datasets

Three commonly used global wetland-related products of different types were selected to be compared and assessed (Table 1).

Land Cover Climate Change Initiative (LC-CCI)

The product consists of global land cover maps with a spatial resolution of 300 m and an annual temporal resolution that refers to the time frame from 1992 to 2015 (Defourny et al., 2017). However, only the map referring to 2015 was considered for the research as it represents the latest extent of wetlands. The legend was developed using the Land Cover Classification System (LCCS, Gregorio, 2016) to make the dataset as compatible as possible with other products. Wetland extent corresponds to 4.7% of the global land surface area (Tootchi et al., 2019).

Global Lakes and Wetlands Database (GLWD-3)

GLWD-3 aims to provide the maximum extent of wetland areas, as well as lakes, reservoirs, and rivers (Lehner & Döll, 2004). The map has around 1 km resolution and it refers to the 1980s. Wetlands cover 6.2-7.6% of the global land surface area. However, Tootchi et al. (2019) concluded that the dataset may underestimate the extent of wetlands.

Composite Wetland Maps (CW-TCI)

This product was developed based on a wetland definition that supports hydrological and land surface modeling applications (Tootchi et al., 2019). Indeed, wetland areas are considered persistently saturated or near saturated. The maps were created by combining (1) regularly flooded wetlands (RFWs) and (2) ground-driven wetlands (GDWs) and have around 500 m resolution. The map CW-TCI15% was chosen given the 21.6% of wetland extent representation, which is the highest percentage among the other products selected since it includes small and scattered wetlands that are not easily detected with remote sensing techniques (Tootchi et al., 2019).

Table 1: Global datasets used in this study

Dataset	Type	Data source	Temporal frame	Wetland classes	Spatial resolution	Reference	Data access
Land Cover Climate Change Initiative (LC-CCI)	GLC dataset	MERIS, SPOT VGT, PROBA-V	1992-2015	Tree cover, flooded, fresh or brackish water; Tree cover, flooded, saline water;	300 m	Defourny et al. 2017	https://lcviewer.vito.be/download

Dataset	Type	Data source	Temporal frame	Wetland classes	Spatial resolution	Reference	Data access
				Shrub or herbaceous cover, flooded, fresh/saline/ brackish water; water bodies			
Composite Wetland Map (CW-TCI15%)	Inundation extent dataset	Land Cover CCI, GIEMS-D15, water table depth estimates (Fan, Li, & Miguez-Macho, 2013), etc.	1984-2015	Regularly flooded wetlands; groundwater-driven wetlands; intersection of RFWs and GDW-TCI15%; lakes	15 arc-second (~500 m)	Tootchi et al. 2019	https://doi.pangaea.de/10.1594/PANGAEA.892657?format=html#download
Global Lakes and Wetlands Database (GLWD-3)	Multi-type GALC datasets	MGLD, LRs, WRD, DCW, ArcWorld, WCMC wetlands map, GLCC	1980s	Freshwater marsh, floodplain; swamp forest, flooded forest; coastal wetland (mangrove, estuary, delta, lagoon); pan, brackish/saline wetland; bog, fen, mire; intermittent wetland/lake; 50–100% wetland; 25–50% wetland; wetland complex (0–25% wetland); lake; reservoir; river	30 arc-second (~1 km)	Lehner and Döll 2004	https://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3

2.2.2 Reference datasets

Three reference datasets (Table 2) were selected to be integrated into one final validation dataset. A legend harmonization was performed (see Chapter 2.3.2) to combine the different classes that each dataset has.

The Global Wetland Validation Samples (GWVS) dataset was created by researching documents through Web Of Science on wetland-related keyword searches with a result of 803 samples (Zheng, Niu, Gong, & Wang, 2017). However, the class "spring" was not included in the final reference dataset since the size of the samples was much smaller than the pixel size of the maps, bringing the final number of samples to 757. A different approach was applied to the integrated GLC map, which was produced through the integration of four GLC maps and six GLC reference datasets (Tsendbazar, de Bruin, & Herold, 2017). Lastly, the GLC validation sample set was developed using Landsat TM/ETM+ images, MODIS EVI times series data, and other high-resolution images (Zhao et al., 2014).

Table 2: Reference datasets

Dataset	Number of aquatic samples	Number of non-aquatic samples	Classification scheme	Wetland class	Reference
Global Wetland Validation Samples (GWVS)	757	NA	Ramsar Convention-based 15 classes	Coastal swamp, coastal marsh, coastal mudflats, estuary, lagoon, shallow water, swamp, marsh, peatland, river, lake, spring, mudflats, water storage area, salt field and fish farm	Zheng et al. 2015
Integrated GLC map	1325	20893	LCCS-based 9 classes	Wetland vegetation, open water	Tsendbazar et al. 2016
GLC validation sample set	1280	37384	LCCS-based 11 classes	Wetland, water bodies	Zhao et al. 2014

2.2.3 Regional datasets

Amazon Basin

Amazon Basin is presented by a dataset of around 90 m resolution that includes wetland extent, vegetation type, and dual-season flooding, of which only the high-flood season (May-June 1996) was considered since it represents the largest extent of wetlands (Hess, Melack, Novo, Barbosa, & Gastil, 2015).

Congo Basin

The final map selected for this region is an integration of various datasets to include both wetlands and water bodies of the Basin area. Indeed, the map representing wetlands did not include the water body class. Therefore, other maps representing water bodies of the countries that are part of the Congo Basin were used to be combined into a final integrated dataset.

Table 3: Regional datasets

	Dataset	**Temporal frame**	**Wetland classes**	**Spatial resolution**	**Reference**	**Data access**
Amazon Basin	LBA-ECO LC-07	Oct.-Nov. 1995 and May-June 1996	Flooded herbaceous; flooded shrub; flooded woodland; flooded forest	3 arc-second (~90 m)	Hess et al., 2015	https://daac.ornl.gov/LBA/guides/LC07_Amazon_Wetlands.html
Congo Basin	Congo_Basin_HCV	Na	Swamp forest; Marantaceae; mangrove	100 m	Grantham et al., 2020	https://www.globildata.org/arcgis/rest/services/CongoBasin/Congo_Basin_HCV/MapServer
	Cameroon DCW Water Bodies	Na	Water bodies	-	DIVA-GIS	https://hub.arcgis.com/datasets/nga::cameroon-dcw-water-bodies-110000000?geometry=5.621%2C2.669%2C20.167%2C6.501
	Gabon Water Areas	Na	Water areas	-	DIVA-GIS	https://hub.arcgis.com/datasets/nga::gabon-water-areas
	Central African Republic DCW Water Bodies	Na	Water bodies	-	DIVA-GIS	https://hub.arcgis.com/datasets/nga::central-african-republic-dcw-water-bodies-110000000?geometry=-7.440%2C-1.039%2C50.743%2C14.194

Dataset	Temporal frame	Wetland classes	Spatial resolution	Reference	Data access
Equatorial Guinea DCW Water Bodies	Na	Water bodies	-	DIVA-GIS	https://hub.arcgis.com/datasets/nga::equatorial-guinea-dcw-water-bodies-110000000
Water bodies (Democratic Republic of Congo)	Na	Water bodies	-	Global Forest Watch	https://data.globalforestwatch.org/datasets/6ed8d2bfd54a42a099f9ffdf660fb8d2_27?geometry=-7.002%2C-8.893%2C51.094%2C6.438
Republic of the Congo DCW Water Bodies	Na	Water bodies	-	DIVA-GIS	https://hub.arcgis.com/datasets/nga::republic-of-the-congo-dcw-water-bodies-110000000

2.3 Research approach

The following sections provide the methodology to determine the consistency and accuracy of the chosen maps and a closer evaluation of the Amazon Basin and Congo Basin regions.

2.3.1 Consistency evaluation of existing wetland datasets on a global extent

Thematic consistency

Thematic harmonization is needed to provide consistency among legends of different maps, with a simplified legend used to aggregate classes and that satisfies all the legends of the maps. However, aggregation implies a generalization which leads to a loss of information (Fritz & See, 2008). Land Cover Classification System (LCCS, Gregorio, 2016) is a translation protocol that allows harmonizing legends, while also highlighting inconsistencies among them (Herold et al., 2008).

It was decided to focus the comparison of the maps on two classes, "wetlands" and "water bodies", while the remaining classes were considered "other". Therefore, the new legend was established into which the original legends were translated. LCCS was used to create the aggregated classes, by examining the description of the classes and translating them into the LCCS-code (Appendix, Table A1).

Table 4 presents the final result of the harmonized legend, which was then used to both compare the consistency of the maps as well as to determine their accuracy in Chapter 2.3.2.

In order to harmonize the legends of each dataset, a few steps were taken. A pre-processing phase included the reprojection of the GLWD-3 map to the WGS84 geographic coordinate system and datum. Then, maps were resampled to around 500 m resolution, which is 0.004167 by 0.004167 decimal degrees, by the nearest-neighbor method. An additional step was needed for the LC-CCI map because the class "water" included both oceans and inland water bodies. Therefore, the tool Reclassification and Raster calculator, along with the dataset CCI WB v4.0 (Lamarche et al., 2017) that has only two classes representing ocean and inland, were employed to separate the class into "ocean" and "water bodies". Then, the legends were reclassified using the Reclassification tool to create harmonized legends according to Table 4. Lastly, since GLWD-3 only includes wetlands and water bodies classes, NoData was used for the harmonized class "other" to allow the comparison for this class as well. The same step was also applied for CW-TCI to include oceans in the final global comparison, thus, providing an equal starting point among the maps.

Table 4: Harmonized legends of global datasets

Harmonized class value		**Class value**	**LC-CCI classes**	**Class value**	**CW-TCI classes**	**Class value**	**GLWD-3 classes**
1	*Wetlands*	160	Tree cover, flooded, fresh, or brackish water	1	Groundwater-driven wetlands	4	Freshwater marsh, floodplain
		170	Tree cover, flooded, saline water	2	Regularly flooded wetlands	5	Swamp forest, flooded forest
		180	Shrub or herbaceous cover, flooded, fresh/saline/brackish water	3	Intersection of RFWs and GDW-TCI15%	6	Coastal wetland (mangrove, estuary, delta, lagoon)
						7	Pan, brackish/saline wetland
						8	Bog, fen, mire
						10	50-100% wetland
						11	25-50% wetland
						12	Wetland complex (0-25% wetland)
2	*Water bodies*	222	Water bodies	4	Lakes	1	Lake
						2	Reservoir
						3	River
						9	Intermittent wetland/lake
3	*Other*	0:150, 190:202, 221	No Data; cropland, rainfed, cropland, irrigated or post-flooding; mosaic	0	Non-wetlands		No Data

Harmonized class value	Class value	LC-CCI classes	Class value	CW-TCI classes	Class value	GLWD-3 classes
		cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%); tree-cover, broadleaved, evergreen, closed to open (>15%); tree cover, needleleaved, evergreen, closed to open (>15%); tree cover, needleleaved, deciduous, closed to open (>15%); tree cover, mixed leaf type (broadleaved and needleleaved); mosaic tree and shrub (>50%) / herbaceous cover (<50%); mosaic herbaceous cover (>50%) / tree and shrub (<50%); shrubland; grassland; lichens and mosses; sparse vegetation (tree, shrub, herbaceous cover) (<15%); urban areas; bare areas; permanent snow and ice; ocean				
				NoData		

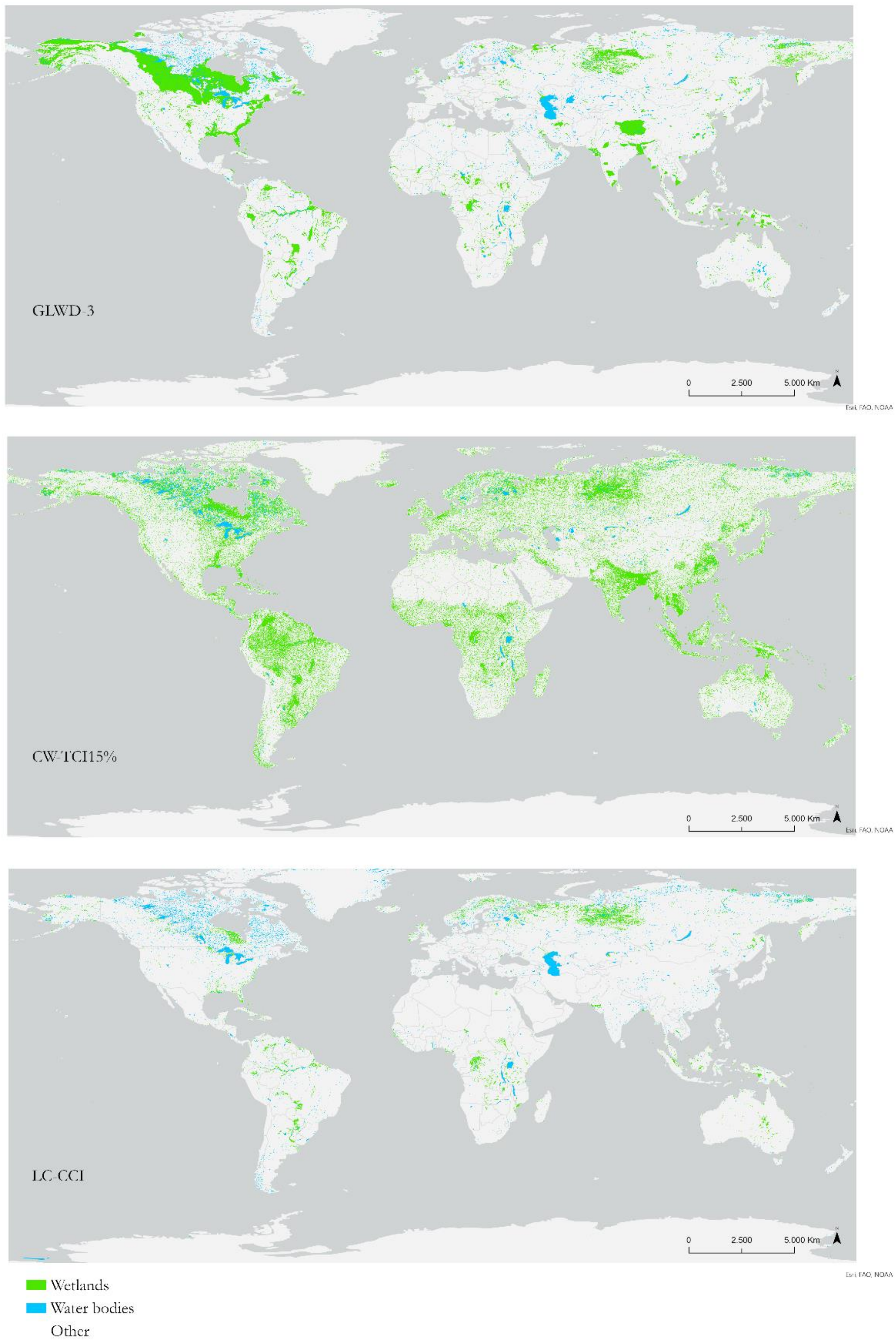


Figure 2: Global maps with a harmonized legend

Note: the class "Other", which corresponds to every pixel that is not "Wetland" or "Water bodies", was assigned no color to improve the visualization of the map.

Spatial consistency

Disagreement among LC products exists due to reasons related to the data source, classification system, spatial resolution, or differences in class definition (Gao & Jia, 2013). A disagreement map can be produced by comparing the maps pixel-by-pixel, which provides an output that can offer a visualization of the disagreement.

Figure 3 illustrates the workflow to create a disagreement map. Raster Calculator tool in ArcGIS Pro was used to compute the pixel-by-pixel comparison analysis between the three maps. The result is a disagreement map showing pixels that agree for each class among all three maps with the following expression:

*"GLWD-3"*100 + "LC-CCI"*10 + "CW-TCI"*

After the reclassification, the result is a final disagreement map showing pixels that agree for each class either for all three maps or only two of them. Moreover, disagreement is shown where no map agrees on any class.

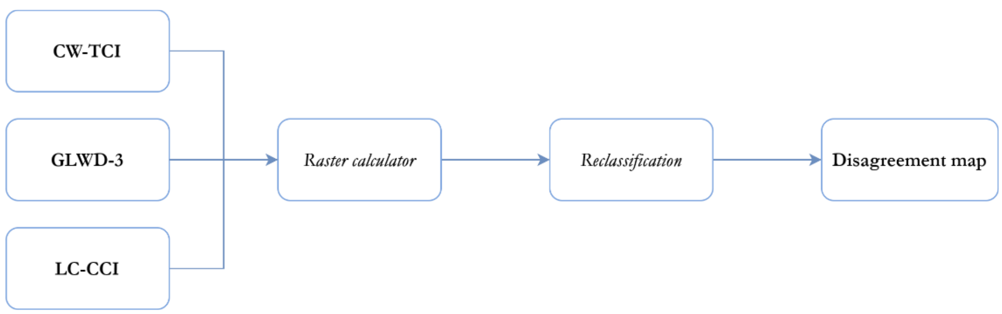


Figure 3: Workflow for disagreement map

The final disagreement map was used to investigate the original classes that cause disagreement or partial agreement. Raster Calculator tool was used to assign the original classes of the global datasets (*Global Map*) to the disagreement areas or to the areas where two maps agree on a class (*Class Disagreement Map*). The disagreement or partial agreement maps were first reclassified so that the selected class was assigned to value 1, while value 0 was applied to the rest of the map. The step was repeated for each class of the disagreement map representing full disagreement and partial agreement. The following generalized expression allowed to assign the original classes to the disagreement or partial agreement areas:

*Class Disagreement Map * Global Map*

Lastly, the percentages were calculated to determine which classes caused the most disagreement or partial agreement. The number of pixels for each class was used to calculate the percentages.

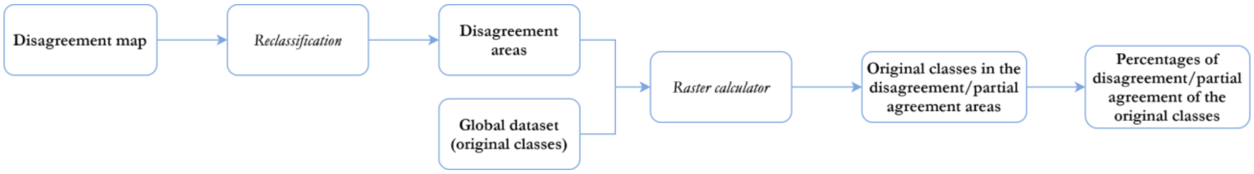


Figure 4: Workflow for percentages of disagreement of original classes for each global dataset

Temporal consistency

The data sources used to create the maps present different temporal ranges. As indicated in Table 5, the LC-CCI map chosen for this research refers to 2015. On the other hand, the CW-TCI map was created with data collected from 1984 to 2015. Lastly, the GLWD-3 map was generated with data from the 1980s, and even though it could be considered outdated, it is still widely used by various user groups (Xu et al., 2020).

The maps were validated with the reference datasets previously mentioned (Table 2) which refer to the year between 2014 and 2016. Therefore, the validation could provide an understanding of whether these datasets are suitable for contemporary studies.

Table 5: Temporal characteristics of the global products

Dataset	Temporal frequency	Reference period
LC-CCI	Yearly	2015
CW-TCI	Static	1984-2015
GLWD-3	Static	1980s

2.3.2 Accuracy assessment of existing wetland datasets on a global extent

Accuracy assessment requires high-quality data, also called reference data, with its legend harmonized with the map. This step is necessary to properly compare the classifications and determine the quality of the map (Herold et al., 2008; Strahler et al., 2006). A confusion or error matrix is a commonly chosen approach to calculate accuracy. The classes of the map and the reference data, which were previously harmonized, are cross-tabulated and the accuracy is quantified.

Three different accuracies can be calculated with the confusion matrix: overall accuracy, producer's accuracy, and user's accuracy (Olofsson et al., 2014). The overall accuracy indicates the quality of the map with the percentage of pixels correctly classified, calculated with the division of the total number of correctly classified points by the total number of reference data samples. The producer's accuracy represents the percentage of correctly classified reference points, calculated by dividing the number of correctly classified samples of each class by the number of reference samples of the same class (column total). Lastly, the user's accuracy is the percentage of the probability of correctly classified pixels in the map, obtained by dividing the number of correctly classified points in each class by the reference samples classified of the same class (row total).

Reference dataset

After the selection of the reference datasets, their legends were harmonized (Table 6). The datasets were added to ArcGIS Pro as XY Table to Point and then merged into a final dataset.

Table 6: Classes of reference datasets

Harmonized class value	Harmonized class		Global Wetland Validation Samples (GWVS)		Integrated GLC map		GLC validation sample set
1	*Wetland*	0	Coastal swamp	5	Wetland vegetation	51	Wetlands/Marshland
		6	Swamp				
		7	Marsh				
		8	Peatland				
2	*Water bodies*	3	Estuary	9	Water	61	Water bodies/Lake
		4	Lagoon			62	Water bodies/Reservoir-pond
		5	Shallow water			63	Water bodies/River
		9	River				
		10	Lake				
		12	Water storage area				
		13	Salt field and fish farm				
3	*Other*	1	Coastal marsh	1	Forest	10	Croplands (rice fields, greenhouse farming, other croplands)
		2	Coastal mudflats	2	Shrubland	20	Forests (broadleaf forests, needleleaf forests, mixed forests, orchard)
		11	Mudflats	3	Grassland	30	Grasslands (pastures, other grassland)

Harmonized class value	Harmonized class	Global Wetland Validation Samples (GWVS)	Integrated GLC map		GLC validation sample set	
			4	Cropland (incl. mixtures)	40	Shrubland
			6	Urban/built-up	52	Wetlands/Mudflats
			7	Snow and ice	64	Ocean
			8	Bare/sparse vegetation	70	Tundra (shrub and brush tundra, herbaceous tundra)
					80	Impervious surfaces (impervious-high albedo, impervious-low albedo)
					90	Barren lands (dry salt flats, sandy areas, bare exposed rock, bare herbaceous croplands, dry lake/river bottoms, other barren lands)
					100	Snow and ice (snow, ice)
					999	Cloud

In order to calculate the accuracy, the maps were firstly harmonized, and their pixel values on the coordinates of the reference samples were extracted. After this step, the confusion matrix was computed in R (Appendix, - Figure A1), along with the overall accuracy, and class-specific accuracies.

2.3.3 Consistency evaluation of existing wetland dataset on a regional extent

Spatial consistency

Figure 5 illustrates a similar workflow on creating the disagreement map as previously presented in Chapter 2.3.1. After harmonizing the legend of the regional dataset (Figure 6, Table A3), the map was resampled to match the resolution (0.004167 by 0.004167 decimal degrees) of the global products. Afterward, each global map was clipped to match the extent of the regional map to allow the comparison. Then, Raster Calculator was used to compute the pixel-by-pixel comparison analysis between the regional map and one global map at a time. The result was reclassified to a final disagreement map that shows pixels that agree for each class, while disagreement is shown where maps do not agree on a class.

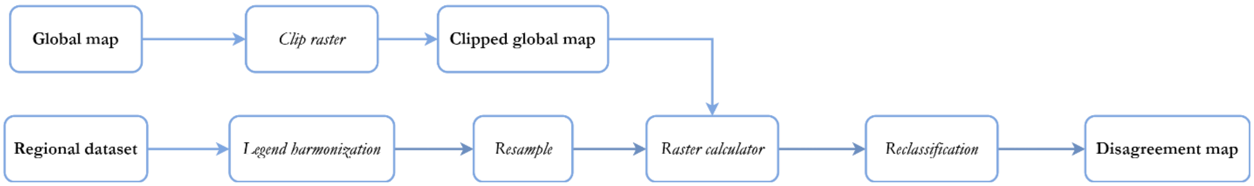


Figure 5: Workflow for regional disagreement map

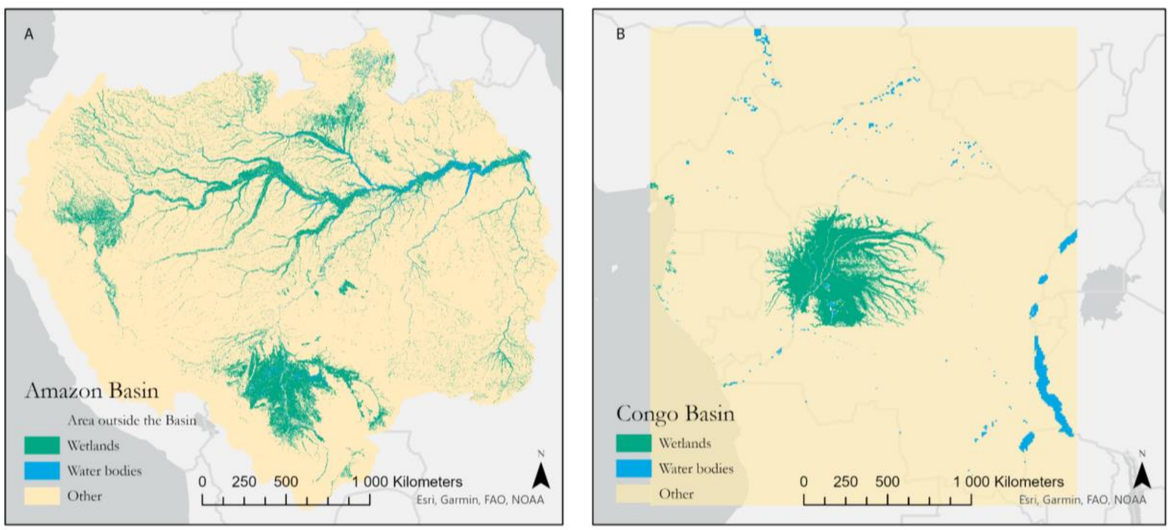


Figure 6: Harmonized regional datasets. No color was assigned to "area outside the Basin" in the Amazon Basin product

After clipping the global maps with the original legends, Raster Calculator was implemented to assign the original classes to the disagreement areas. For example, the disagreement map between GLWD-3 and the Congo regional product was used to extract areas of disagreement. Then, the output was used to assign the original classes of the global dataset, and later of the regional dataset, to the disagreement areas to determine which classes caused the most disagreement. The steps were repeated for each global dataset in both areas of interest. Lastly, the number of pixels for each class was used to calculate the percentages of the classes that caused the most disagreement.

3. Results

This chapter describes the results of the disagreement maps and accuracy assessment calculated for each dataset. The global comparison among the datasets is outlined in Chapter 3.1, while Chapter 3.3 assesses the global maps at a regional scale along with regional datasets. Chapter 3.2 illustrates the result of the accuracy assessment.

3.1 Consistency evaluation of wetland datasets on a global extent

Figure 7 illustrates the disagreement map between the three selected global products. Full and partial agreement on wetlands is visible particularly in northern Canada and northern Siberia, which is also where most disagreements occur, as well as in South America and India. Furthermore, there are noticeable disagreements in some areas in southern Africa and Australia.

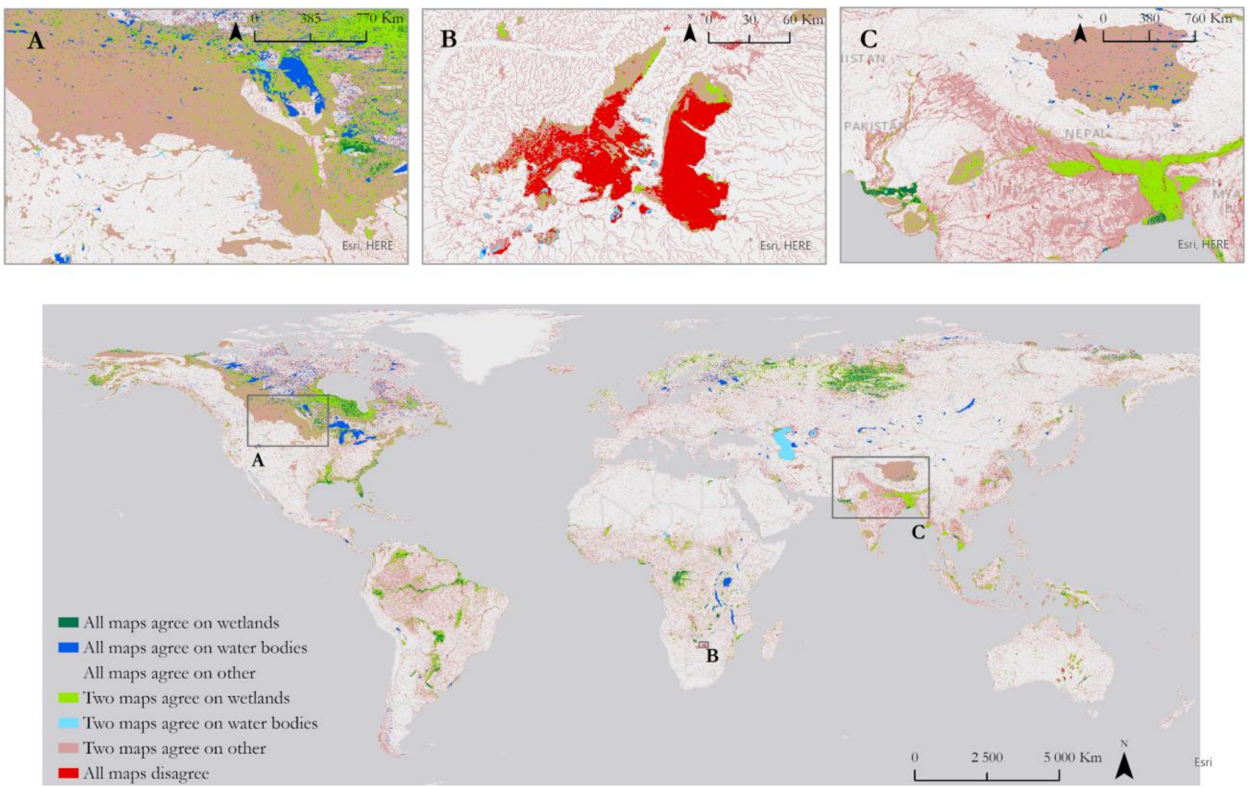


Figure 7: Disagreement map of global products (A: Prairie Pothole Region, B: Makgadikgadi wetlands in Botswana, C: Tibetan Plateau, Bangladesh, and Northern India)

Map A of Figure 7 shows a close-up of the Prairie Pothole Region in North America. This area presents partial agreement of two maps that consider the area with classes classified as "other", while a third map classified is as wetlands. A similar account occurs in Map C, where the Tibetan Plateau is a large area where two maps agree on the class "other", while a third map classifies the area as wetlands. The map also includes the northern part of India where partial agreement is shown for class "other" and “wetlands”. Lastly, Map B provides a look into the Makgadikgadi wetlands in Botswana where maps disagree on the classification.

Regarding the class water bodies, maps present a full agreement and partial agreement mostly. A visible example of partial agreement is the Caspian Sea that belongs to the class “other” in the CW-TCI map after the reclassification because it is considered a sea, in contrast to the other two maps that classify it as water bodies. However, partial agreement occurs in various lakes in each continent. An example is shown in Appendix - Figure A3 with lake Chad, which has an extent that varies greatly among the maps.

A further step was taken to determine which original classes contributed the most and the least to the disagreement among the maps. The bar charts in Figure 8 summarize the percentages of the harmonized classes of each global product when compared to the areas of disagreement, and the detailed percentages of the original classes for each global dataset can be found in the Appendix in - Table A4.

The chart reveals a high percentage of disagreement of 87.7% in the CW-TCI dataset for classes that were reclassified into wetlands, which indicates an overrepresentation of the class. The original classes that create the most disagreement are “regularly flooded wetlands” with 51.3%, while the class “intersection of RFWs and GDW-TCI15%” and class “groundwater-driven wetlands” reach 18.9% and 17.4% of the disagreement, suggesting unreliability of the product to map wetlands. GLWD-3 and LC-CCI datasets present a low disagreement of wetlands with 8.7% and 3.6%, respectively, indicating the products have a less overestimation of wetlands compared to CW-TCI. The original wetland classes for the GLWD-3 dataset causing disagreement are “freshwater marsh, floodplain”, “swamp forest, flooded forest”, “coastal wetlands”, while the ones of the LC-CCI datasets are “tree cover, flooded, fresh or brackish water”, “tree cover, flooded, saline water”, and “shrub or herbaceous cover, flooded, fresh-saline or brackish water”. Thus, classes that create disagreement are mostly related to wetland types that are prone to seasonal changes and should be validated separately.

The water bodies class has a similar trend to the previous one described above. Indeed, GLWD-3 and LC-CCI have a higher disagreement of 39.3% and 50.4% respectively, compared to the CW-TCI dataset showing a disagreement of 8.5%. Classes of GLWD-3 that cause disagreement are “lake”, “river”, “intermittent wetland/lake”, and “reservoir”. The LC-CCI and CW-TCI datasets only have one class that belongs to the harmonized class of water bodies. The class of the LC-CCI dataset was created by dividing the class water into “ocean” and inland “water bodies”.

Lastly, the GLWD-3 dataset presents disagreement for the class belonging to “other” at 52.0%. The class was created by turning NoData pixels into a class since the original map does not have classes that could be harmonized into the class named “other”. The LC-CCI dataset has a high percentage of disagreement for the same class (46.0%), deriving mostly from the original classes of “cropland”, “tree cover”, and “bare/sparse vegetation”. On the other hand, due to overestimation of wetlands, CW-TCI has disagreement of class “other” at 3.8%. NoData pixels, which represent oceans, were reclassified into the class “other” to provide a cohesive comparison between the maps.

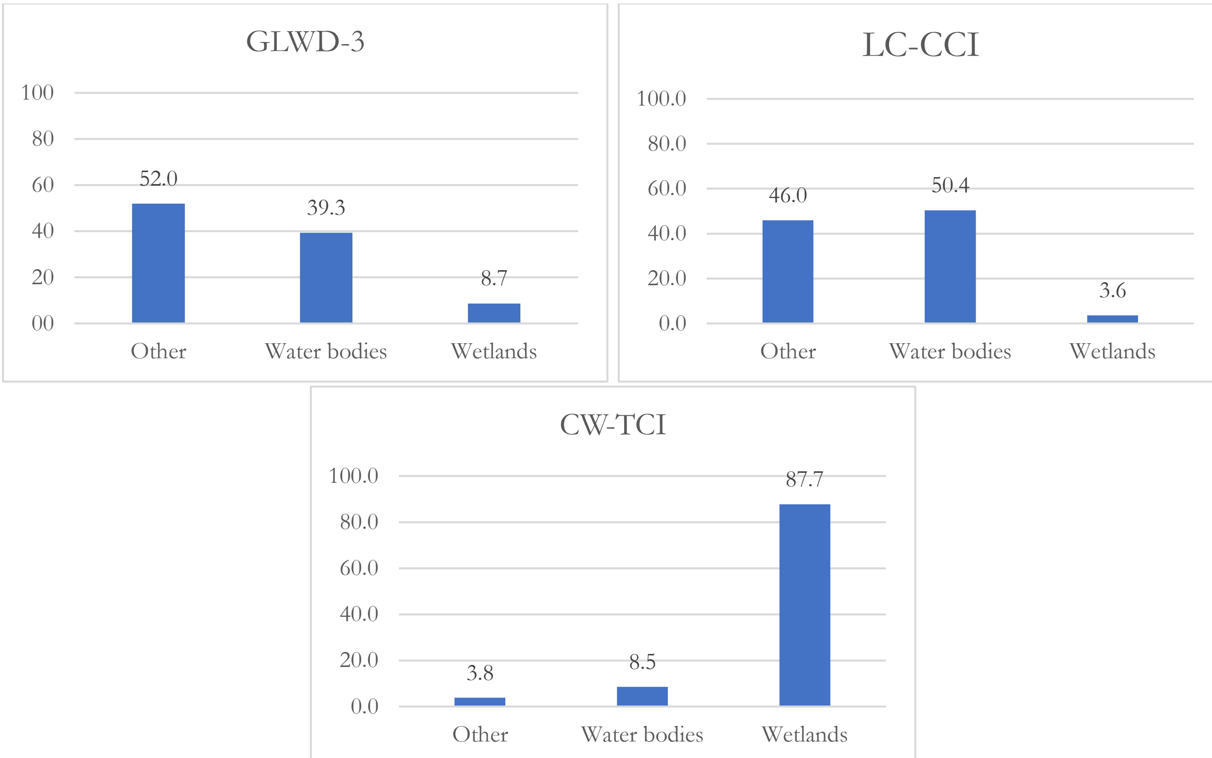


Figure 8: Percentages of class disagreement according to the original classes of the global product

Figure 9 illustrates the percentages of the harmonized classes of each global product when two maps agree on one harmonized class and the third one disagrees. For example, the chart on wetlands presents the percentages of wetland classes for each global dataset when there is partial agreement. On the other hand, percentages of disagreement are represented by the classes water bodies and "other", indicating that the dataset disagrees in some areas and classified them differently than the other two maps. The detailed percentages can be found in the Appendix from - Table A5 to - Table A7.

CW-TCI has the highest agreement with a second map for classes that were harmonized as wetlands at 95.8%, followed by GLWD-3 with 72.8%. The original classes that agree the most are "freshwater marsh, floodplain", "25-50% wetland", and "50-100% wetland" for the GLWD-3 dataset, while all wetland classes of CW-TCI present high percentages overall, with "regularly flooded wetlands" representing the largest portion of the percentage with 44.6%. LC-CCI presents the lowest agreement with another map on wetlands with only 31.5% represented mainly by "shrub or herbaceous cover, flooded, fresh-saline or brackish water". On the other hand, LC-CCI is the dataset with the highest agreement with a second map on water bodies with 91.4%. GLWD-3 and CW-TCI have 58.6% and 49.9% respectively of water bodies classes that partially agree with a second map. Lastly, LC-CCI has 97.8% of classes classified as "other" that partially agree, followed by GLWD-3 with 73.4%. On the other hand, CW-TCI only reaches 28.7% of partial agreement for this class due to 70.5% of wetlands classes being part of these partial agreement areas.

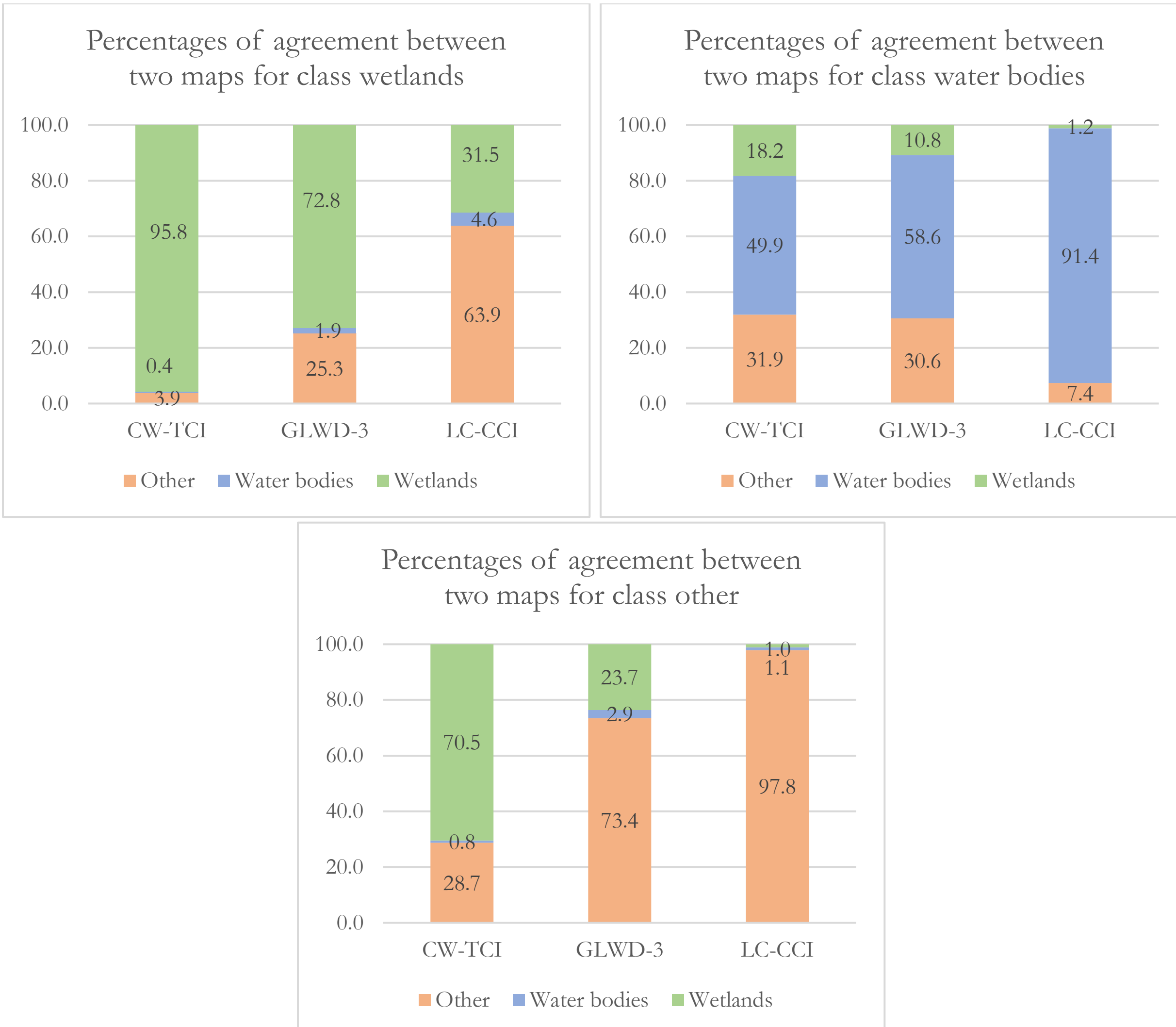


Figure 9: Percentage of the class agreement for each of the harmonized classes when two maps agree on the class and the third one disagrees

3.2 Accuracy assessment of wetland datasets on a global extent

The confusion matrices were calculated in R (Appendix - Figure A1) and reported in Table 7. The overall accuracies for GLWD-3, LC-CCI, and CW-TCI maps differ significantly at 87.2%, 93.9%, and 73.0% respectively. However, class-specific accuracies are needed to provide information about the accuracy of each class, which is not represented by the overall accuracy (Herold et al., 2008). Wetlands are overestimated by all maps, especially by the CW-TCI map, since there is a high producer's accuracy and a low user's accuracy. Furthermore, when comparing LC-CCI and GLWD-3, the first map overestimates wetlands class less than the other one because its user's accuracy is higher while the producer's accuracy is similar to GLWD-3. On the other hand, water bodies are underrepresented by the three maps since the producer's accuracy is lower than the user's accuracy. Lastly, GLWD-3 and CW-TCI maps underestimate classes harmonized as "other", while LC-CCI slightly underrepresents it.

Table 7: Confusion matrix of each global product

GLWD-3							
	Reference classes				Correct	Total	User's Accuracy %
Map classes		Wetlands	Water bodies	Other			
	Wetlands	**374**	249	5172	374	5795	**6.5**
	Water bodies	84	**937**	725	937	1746	**53.7**
	Other	554	1129	**52415**	52415	54098	**96.9**
Correct		374	937	52415	**53726**		
Total		1012	2315	58312		**61639**	
Producer's Accuracy %		**37.0**	**40.5**	**89.9**			**Overall Accuracy 87.2 %**

LC-CCI							
	Reference classes				Correct	Total	User's Accuracy %
Map classes		Wetlands	Water bodies	Other			
	Wetlands	**350**	79	1548	350	1977	**17.7**
	Water bodies	110	**1215**	478	1215	1803	**67.4**
	Other	552	1021	**56286**	56286	57859	**97.3**
Correct		350	1215	56286	**57851**		
Total		1012	2315	58312		**61639**	
Producer's Accuracy %		**34.6**	**52.5**	**96.5**			**Overall Accuracy 93.9 %**

CW-TCI							
	Reference classes				Correct	Total	User's Accuracy %
Map classes		Wetlands	Water bodies	Other			
	Wetlands	**700**	762	14508	700	15970	**4.4**
	Water bodies	52	**776**	286	776	1114	**69.7**
	Other	260	776	**43517**	43517	44553	**97.7**
Correct		700	776	43517	**44993**		
Total		1012	2314	58311		**61637**	
Producer's Accuracy %		**69.2**	**33.5**	**74.6**			**Overall Accuracy 73.0 %**

Figure 10 is a representation of some error points of the confusion matrices in the Tibetan Plateau. In the eastern part of the area, there is a common pattern of wetlands of the reference dataset being classified as "other" in the global datasets. Therefore, all maps show an error of commission since they misclassified wetlands in the classified map. GLWD-3 and CW-TCI maps have a higher number of points mapped as wetlands, while the reference dataset classified the points as "other". CW-TCI and GLWD-3 classify many points as wetlands, while the reference dataset considers them as "other". Lastly, GLWD-3 classifies some points in the northern area as water bodies instead of "other", while CW-TCI classifies the same points mostly as wetlands. On the other hand, LC-CCI correctly classifies the area overall.

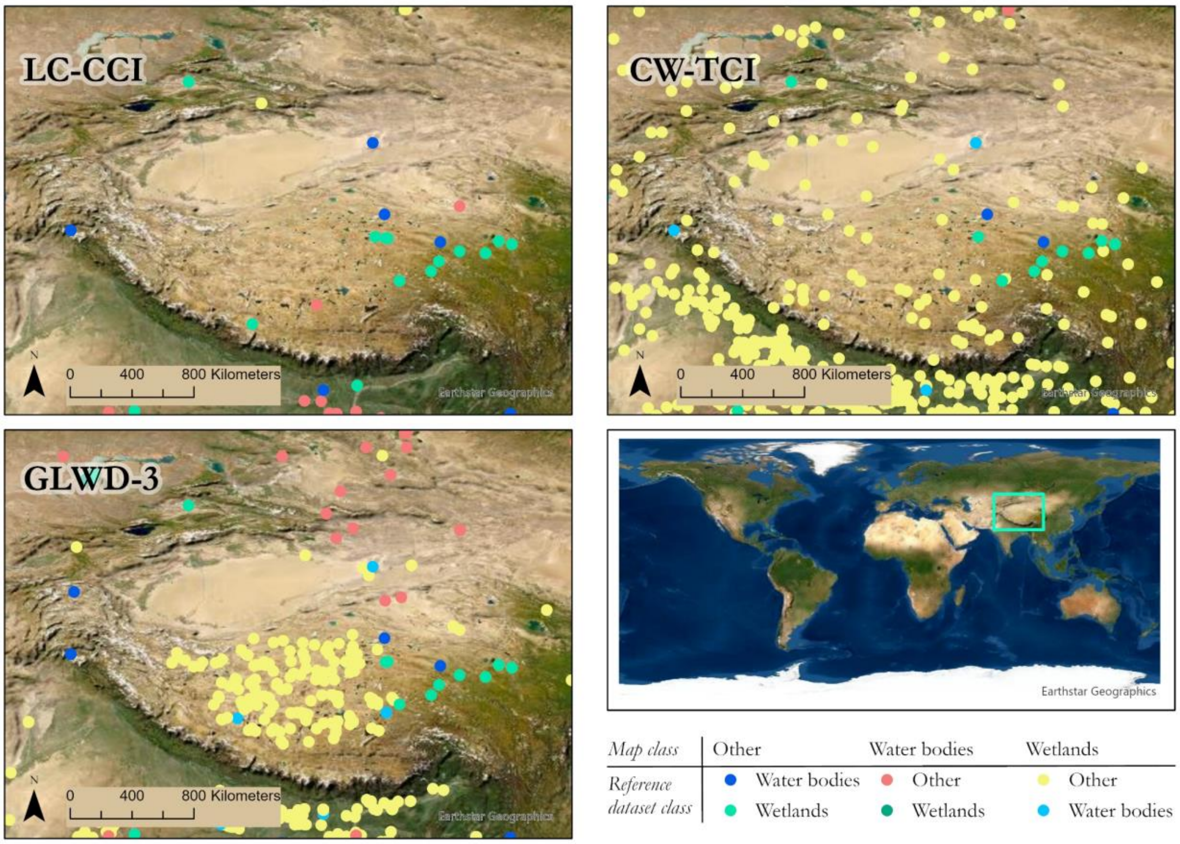


Figure 10: Error points in the Tibetan Plateau. The points represent the map class values that were classified differently by the reference dataset.

3.3 Consistency evaluation of wetland datasets on a regional extent

The global maps were compared to the regional datasets of Amazon and Congo Basins. Similar to the methodology of global consistency evaluation, the percentages of the original classes for each global and regional dataset were included to provide a detailed account of the classes that have a higher disagreement. Global datasets were compared to the regional ones individually to better understand the consistency of the global products at a regional extent.

Amazon Basin

Figure 11 represents the disagreement maps between the Amazon Basin regional dataset LBA-ECO LC-07 and each global dataset. The maps present high disagreement in the south area for both the LC-CCI and GLWD-3 maps. On the other hand, the CW-TCI map shows the highest degree of disagreement throughout the whole area of interest, a similar outcome that occurred in the global comparison as well. However, the CW-TCI map also shows more areas of agreement on wetlands class in contrast to the other two maps.

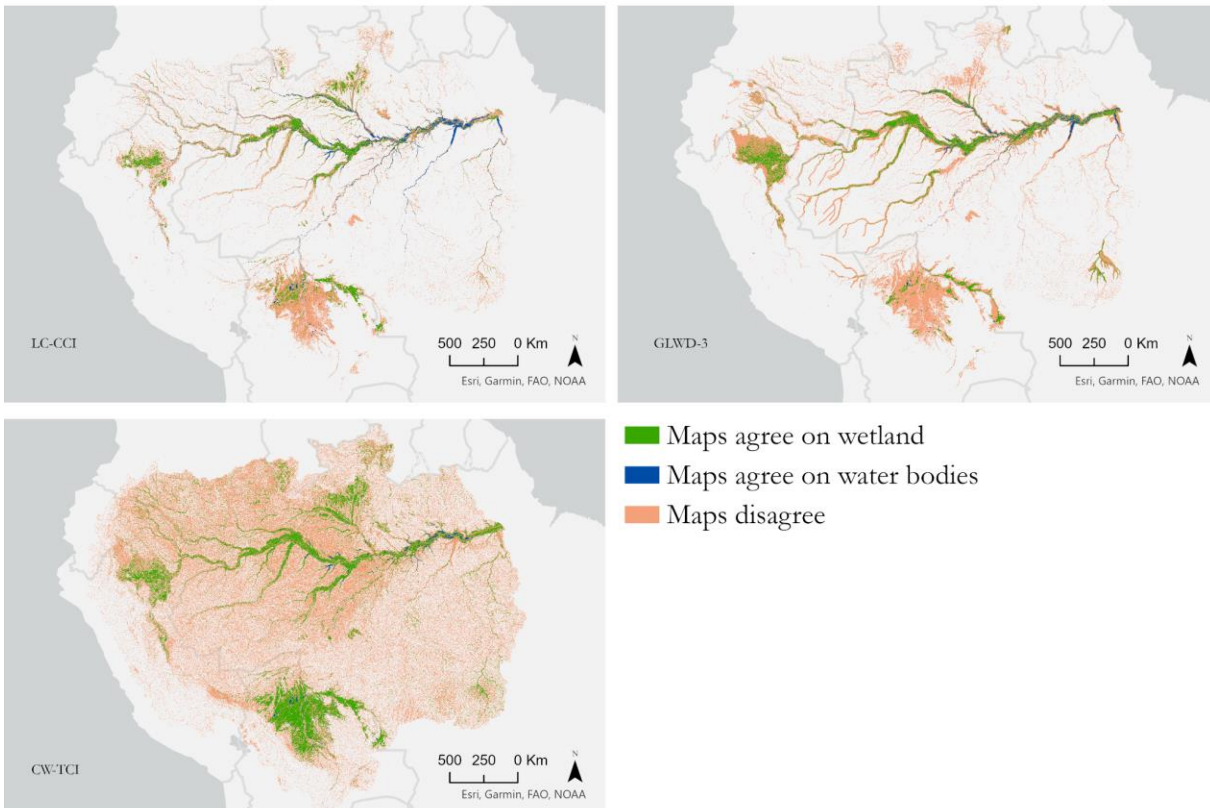


Figure 11: Disagreement map between Amazon Basin regional dataset and global datasets

The bar charts in Figure 12 summarize the percentages of the harmonized classes of each global product when compared to the areas of disagreement with the regional product, while - Table A8- Table A9- Table A10 in the Appendix present a detailed account of the percentages of the original classes for each global and regional dataset.

Wetlands class has the highest disagreement for the CW-TCI with 93.9%, which is consistent with the global comparison seen in Chapter 3.1. The original class that contributes the most to this disagreement is “groundwater-driven wetlands”, followed by “intersection of RFWs and GDW-TCI15%” and “regularly flooded wetlands”. The GLWD-3 has the second-highest disagreement of wetlands with 34%, deriving mainly from the original class named “swamp forest, flooded forest”. Lastly, the LC-CCI dataset presents the lowest disagreement of wetlands with 11.7% represented mostly by the original classes named “tree cover, flooded, fresh or brackish water” and “shrub or herbaceous cover, flooded, fresh-saline or brackish water”.

Water bodies classes do not have a high disagreement overall, with 9.5% being the highest percentage for the GLWD-3 dataset mainly represented by the original classes of "river" and "lake". On the other hand, the LC-CCI and the CW-TCI datasets have 6.4% and 0.3%, respectively.

Lastly, the class "other" presents a high disagreement for the LC-CCI dataset with 81.9%, represented mainly by the original classes categorized as "tree cover". The GLWD-3 dataset has a 56.6% of disagreement, and as for the global consistency evaluation, the class was created by turning NoData pixels into a class, since the original map does not have classes that could be harmonized into the class "other". The CW-TCI dataset shows the lowest disagreement with 5.8%.

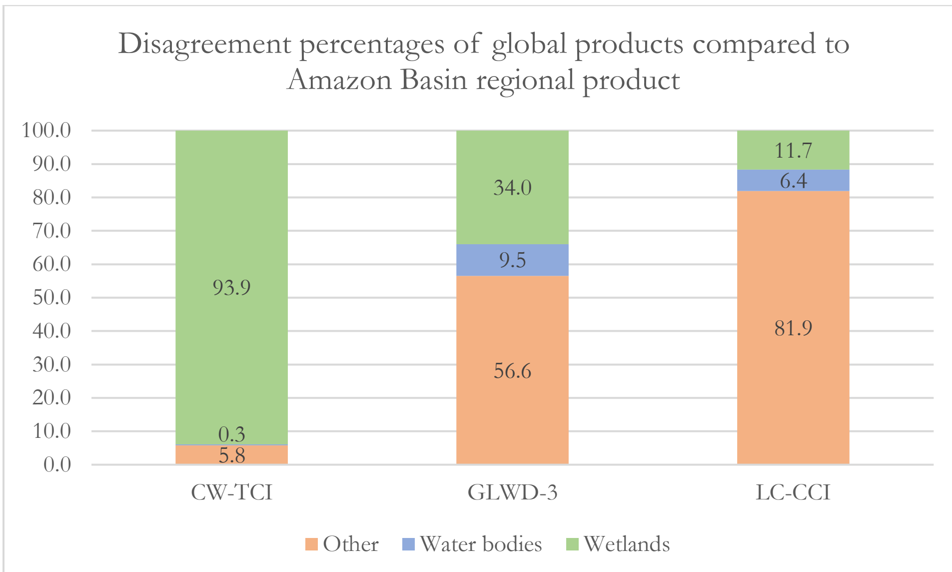


Figure 12: Disagreement percentages of global products when compared to Amazon Basin product LBA-ECO LC-07

Congo Basin

The CW-TCI map shows the highest disagreement areas for the Congo Basin. On the other hand, the LC-CCI and the GLWD-3 maps present different patterns of disagreement with the regional dataset. The first map appears to have a more fragmented agreement on wetlands, while the latter has a more consistent agreement on areas classified as "swamp forest, flooded forest", while disagreement is more evident for the class "other", which was created from NoData since the original legend did not have any class that could be included in the harmonized class "other".

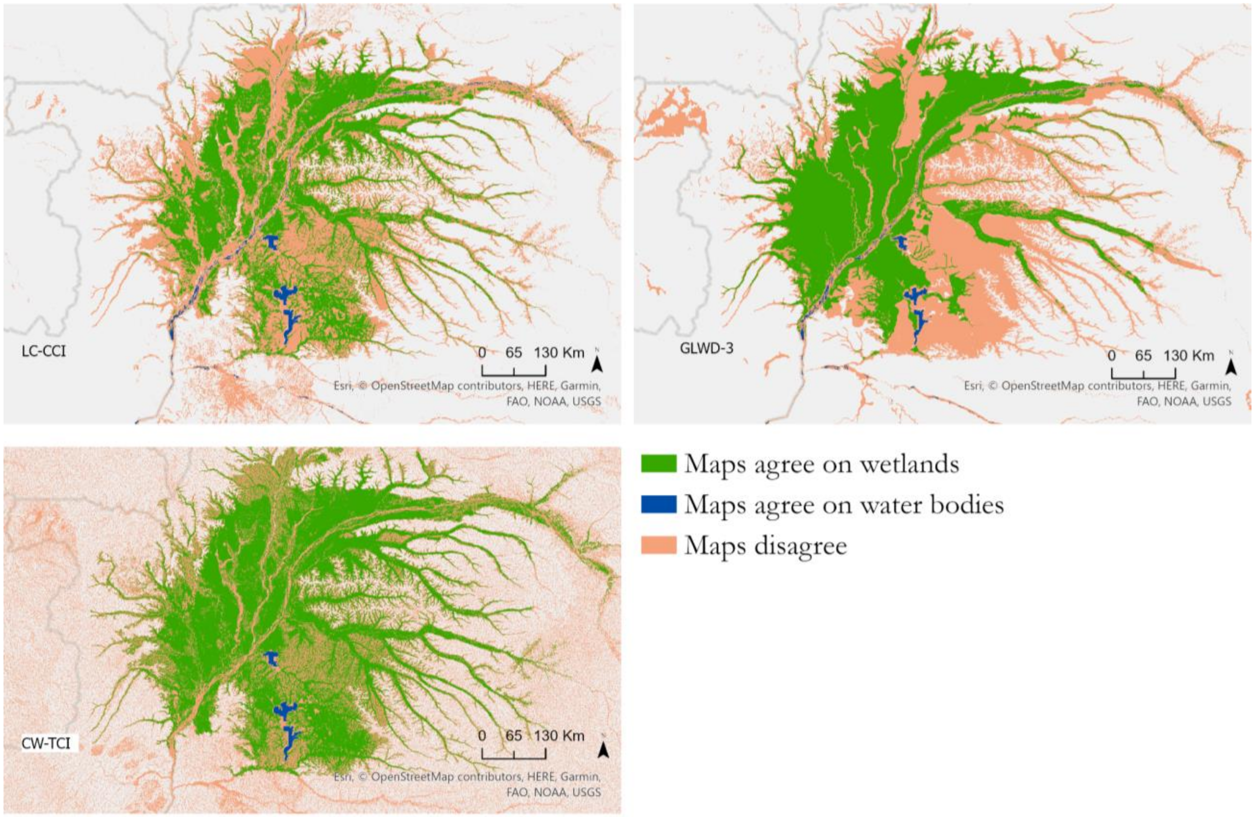


Figure 13: Disagreement map between Congo Basin regional dataset and global datasets

Figure 14 summarizes the percentages of harmonized classes of each global product when compared to the areas of disagreement in the Congo Basin region. Furthermore, - Table A11,- Table A12- Table A13 (Appendix) present the percentages of the original classes for each global and regional dataset.

For the Congo Basin region, water bodies do not have a high disagreement percentage. LC-CCI and GLWD-3 have a similar percentage with 9.8% and 9.2%, respectively. The original classes of GLWD-3 that contribute the most disagreement are "river" and "lake". CW-TCI has the lowest disagreement reaching only 1.0%.

Wetlands class presents the highest disagreement for all three datasets. CW-TCI has 92.0%, with "groundwater-driven wetlands" being the original class that contributes the most to the percentage. GLWD-3 has the second-highest disagreement of 62.3%, mainly due to the original class "freshwater marsh, floodplain". Lastly, LC-CCI shows 45.8%, originating mostly from "shrub or herbaceous cover, flooded, fresh-saline or brackish water" and "tree cover, flooded, fresh or brackish water".

LC-CCI and GLWD-3 datasets have the highest disagreement for classes harmonized as "other" with 44.4% and 28.4%, respectively. "Tree cover" is the class that contributes the most to the disagreement for the LC-CCI map. On the other hand, CW-TCI has 6.1% of disagreement of class "other".

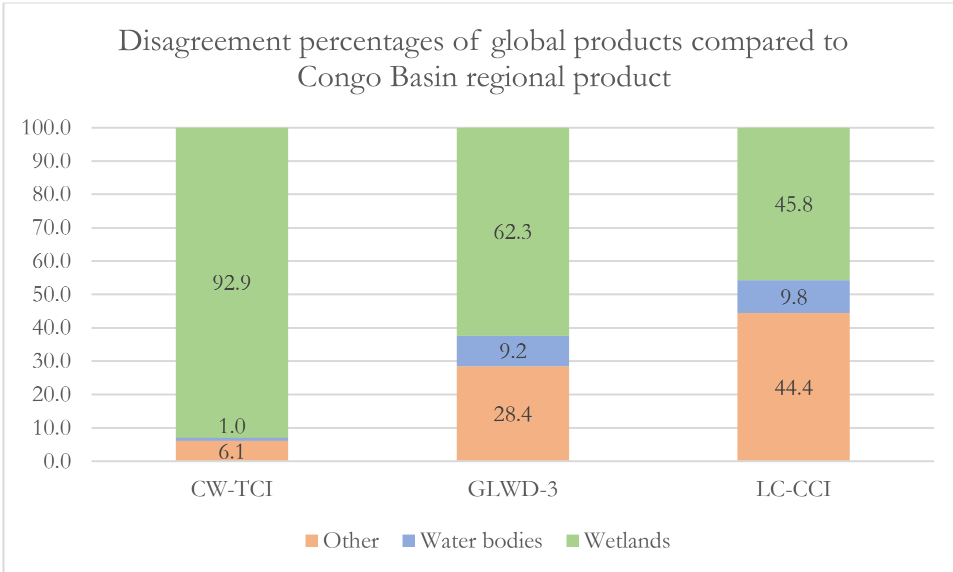


Figure 14: Disagreement percentages of global products when compared to Congo Basin regional product

4. Discussion

This chapter presents the main findings on the results of the research. Chapter 4.1 outlines the findings on global consistency, and Chapter 4.2 discusses the accuracies of the global products. Finally, consistency on a regional level is highlighted in Chapter 4.3.

4.1 Global consistency

The results indicate that the CW-TCI dataset has the highest percentage of disagreement for the harmonized class wetlands at 87.7%, while the LC-CCI and GLWD-3 datasets have a relatively low percentage of disagreement for wetlands. Taken the classification accuracy of CW-TCI into account, it can be inferred that the inconsistency could also be caused by the low quality of CW-TCI because it has a lot of misclassifications. To explain the high percentage of CW-TCI, it is important to remind that this map has the largest extent of wetlands reported at 21.6% in comparison to the other two maps (6.2-7.6% for GLWD-3 and 4.7% for LC-CCI). Its significant extent is due to the topographic-climate wetness index (TCI), which is one of the data sources implemented to create the dataset, which allows incorporating small wetlands that are not detected with remote sensing techniques (Tootchi et al., 2019). Thus, the elevated disagreement can be explained by the high difference in wetlands extent.

Definition inconsistencies are a cause of disagreement. For example, the Makgadikgadi wetlands in Botswana presented high percentages of disagreement on seasonal and intermittent saline/brackish lakes (Arntzen, 2018). The three global maps do not agree on the classification of this area, but GLWD-3 has the most suitable classification with the original class being "pan, brackish/saline wetland". On the other hand, CW-TCI classifies them as water bodies and LC-CCI as "other". As a result, the disagreement occurs due to a problem of definition inconsistency, which is related to the type of dataset and data source.

Partial agreement occurs when two maps agree on a class and a third one disagrees. CW-TCI has the highest value of the agreement for the wetlands class, followed by GLWD-3. The two datasets have the highest consistency for the wetlands class compared to the low percentage for LC-CCI. On the other hand, LC-CCI stands out for the percentage of the water bodies, while GLWD-3 and CW-TCI have a similar percentage. Lastly, the agreement on the class "other" occurs more between LC-CCI and GLWD-3, with the first one having the highest percentage, while CW-TCI presents a high percentage of wetlands class in disagreement since the dataset overestimated and wrongly classified the class as explained in Chapter 3.2. An example of partial disagreement occurs with the LC-CCI original class "cropland, irrigated, or post-flooding", which was classified into the harmonized class "other" (Figure 7, Map C). The area that includes Bangladesh and the north-east part of India is classified mainly as wetlands by GLWD-3 and CW-TCI, while the LC-CCI assigns it to the class "other". On the other hand, the zone of north-east Pakistan at the border with India presents a wetland area for the CW-TCI map, a non-wetland one for GLWD-3, and the original class "cropland, irrigated, or post-flooding" for LC-CCI. These two examples show that agreement for the class wetlands is not consistent among the same two maps when a third map disagrees. Therefore, assigning the original class "cropland, irrigated, or post-flooding" to the harmonized class wetland would have created disagreement with any of the two options, which can be related to both data source and wetland definition. Indeed, the cropland class of LC-CCI does not refer to areas of considerable inundation that could be considered wetlands, while a more specific class such as "paddy rice field" would have been classified as wetlands.

As reported by Tootchi et al. (2019), the GLWD-3 dataset classifies areas prone to water accumulation, but with limited information on the presence of wetlands, as fractional classes. This can be observed in the Prairie Pothole Region in North America and in the Tibetan Plateau in Asia, which have a similar outcome in the disagreement map. LC-CCI and CW-TCI datasets classify them with classes that were harmonized as "other", while the original classes of GLWD-3 are "wetland complex (0-25% wetland)" for the Tibetan Plateau and "25-50% wetland" for the Prairie Pothole Region. Thus, these classes represent areas where wetlands are overestimated by GLWD-3 when compared to the other maps. This is consistent for all areas classified with one of the three fractional classes of the dataset, which can also be observed in Alaska, Canada, and India.

Lastly, temporal consistency needs to be mentioned as the products present different temporal ranges. GLWD-3 is the most outdated dataset of the three with its data sourcing from the 1980s. The product showed some disagreement and accuracy issues, which indicates that newer products should be chosen instead. CW-TCI data sources were collected between 1984 and 2015 and this product performed poorly both for consistency and accuracy, even though the time frame was fairly recent compared to the validation datasets that refer to data collected between 2014 and 2016. Lastly, LC-CCI, a recent product with data from 2015, had better consistency than the other maps and revealed to be the most reliable for wetland mapping.

4.2 Global accuracy

The confusion matrix assessed class-specific accuracies of wetlands class for LC-CCI at 34.6% (producer) and 17.7% (user). On the other hand, the product report indicated the class-accuracies for each original class: "Tree cover, flooded, fresh or brackish water" at 86%, "Tree cover, flooded, saline water" at 86%, and "Shrub or herbaceous cover, flooded, fresh/saline/brackish water" at 24% for the producer's accuracy, while the user's accuracy was at 26%, 75%, and 53%, respectively (Defourny et al., 2017). Overall, both results report an overestimation of wetlands, but there are some differences among the outcomes of the original classes, which are not present for the harmonized class since it reduced thematic detail into one generalized land cover class (Herold et al., 2008). Furthermore, the LC-CCI classified most of the missed samples as "other" (552), which signifies omission of wetlands, an error that occurs for GLWD-3 as well with a similar number of points (554).

CW-TCI and GLWD-3 do not have reported class-accuracies and the assessment of this study determined their producer's accuracy for wetlands at 69.2% and 37.0%, respectively, while the user's accuracy was 4.4% and 6.5%. CW-TCI has the lowest accuracy given the prominent overestimation of wetlands, while GLWD-3 has a less degree of overestimation compared to the other map.

Overall, LC-CCI is the dataset that performs better for class wetlands in the confusion matrix, followed by GLWD-3 with a similar outcome. Lastly, CW-TCI is the dataset that overestimates wetlands the most. However, it is relevant to mention that the validation dataset is composed of land cover reference datasets, which could be potential bias towards LC-CCI, a land cover product.

4.3 Regional consistency

CW-TCI overrepresents the class wetlands since the disagreement is above 90% in both Amazon and Congo Basin regions. The map classifies wetlands mostly at the cost of the class "other", and the overestimation is in line with what occurs at a global extent and confirmed by the confusion matrix. The outcome for the water bodies class is positive, with only 0.3% and 1.0% of disagreement in the Amazon and Congo Basin, respectively. Compared to the results on a global extent, the accuracy is quite consistent with the regional results. Overall, CW-TCI overrepresents wetlands in both regions and the regional results are consistent with the global ones.

GLWD-3 has a different outcome when compared to the regional products. The comparison with the Amazon Basin product shows a higher disagreement on the class "other", followed by wetlands. Furthermore, it appears that some wetland areas are missed, and others are wrongly classified. On the other hand, the comparison with the Congo Basin product presents a different outcome since GLWD-3 overestimates wetland areas with 62.3% in disagreement with the regional product. Compared to the global wetland representation, GLWD-3 underrepresents the class in the Amazon Basin region, while the same class is overestimated in the Congo Basin region. Therefore, GLWD-3 presents inconsistencies in wetland mapping in regions that are considered hotspots, while the outcome at a global extent can be considered suitable.

LC-CCI dataset overestimates the class "other" in the Amazon Basin area, but the map also underrepresents wetlands, as the regional product indicates 57.5% of wetlands class being part of the disagreement areas (Table A9). Regarding the Congo Basin region, the global product overestimates 45.8% of wetland areas and misses 40.3% according to the regional product (Table A12). The class "other" presents a similar account, which means that LC-CCI does not classify the area correctly and it overestimates the class. On the other hand, the "water bodies" class has a lower disagreement, which is not what occurs at a global level because it overestimates the class. As a result, LC-CCI presents a diverging outcome for wetland mapping at the regional hotspots.

Overall, the class "other" and wetlands are the ones involved in the misclassification most often at a regional level, while the water bodies class has a lower misclassification, which is in contrast with the global result for GLWD-3 and LC-CCI. Lastly, CW-TCI is consistent in overestimating wetlands both at regional and global scales.

5. Conclusions and recommendations

5.1 Research questions

This research aimed to evaluate the consistency and accuracy of the global wetland products GLWD-3, LC-CCI, and CW-TCI. This was accomplished with the creation of disagreement maps on a global and regional scale, and by assessing the overall accuracy and class-specific accuracies of each global product. The main research question and the sub-questions are answered below.

> *"To what extent are the three selected global wetland products consistent and accurate: CW-TCI, LC-CCI, and GLWD-3?"*

The three global products are not consistent due to the high variation among them both at global and regional extents. CW-TCI dataset disagreed the most on wetlands with the regional datasets and presented high percentages of disagreement on a global scale. Furthermore, its low classification accuracy makes the product unreliable. GLWD-3 showed better definition consistency than the other two datasets, but the accuracy assessment result reveals an overestimation of wetlands for this product as well, even if the regional comparison showed better agreement than the CW-TCI dataset, but with areas that were not correctly classified. LC-CCI has definition inconsistencies due to a lack of aquatic classes, but the product showed a better outcome from accuracy assessment and the regional comparison in contrast with the other datasets. Even though it missed areas of wetlands that the regional datasets mapped, the product showed less disagreement with the Amazon and Congo products. Overall, all datasets tend to overestimate wetlands, and accuracies for wetlands and water is low.

1. *To what extent are they spatially, thematically, and temporally consistent among each other on a global extent?*

From a spatial point of view, CW-TCI and GLWD-3 datasets are more consistent with each other on the mapping of wetlands. However, CW-TCI is not reliable on wetland extent, due to its high disagreement for wetlands class in respect to the other maps. On the other hand, LC-CCI is less consistent on wetlands compared to the other datasets, while water bodies appear to be more spatially consistent with CW-TCI and GLWD-3. Thematically, LC-CCI showed some legend inconsistencies with the other maps, which was due to the lack of specific aquatic classes. CW-TCI and GLWD-3 have both aquatic-oriented legends, however, CW-TCI revealed problems related to the data source that integrated lakes into the final dataset. Regarding the temporal consistency, the products are not consistent temporally given the different reference periods of the data sources used.

2. *How accurate are they when compared with an integrated validation dataset on a global extent?*

Even though the datasets are not temporally consistent, the accuracy assessment provided an understanding of how suitable the datasets are for contemporary studies. LC-CCI is the most recent dataset and the validation confirmed a better wetland accuracy for this product, followed by GLWD-3 and CW-TCI datasets. Even though all maps overestimate wetlands, LC-CCI has a better representation of wetlands compared to the other products.

3. *What insights can the Amazon Basin and Congo Basin regions provide to the global extent consistency and accuracy analysis?*

The regional analysis confirms the overestimation of wetlands for all maps, and it reveals areas that were missed or wrongly classified as wetlands, mostly due to the harmonized class "other". Furthermore, the regional comparison shows the same trends for the global products on the global extent, such as high disagreement for the class wetlands for CW-TCI, and lower disagreement for GLWD-3 and LC-CCI.

In conclusion, wetlands are overestimated by all products, and despite the overrepresentation, LC-CCI is found to be better suitable for contemporary studies, while GLWD-3 can be considered outdated and CW-TCI unreliable. However, choosing one product over the others depends on the requirements needed for its application. Given the inconsistency assessed at both Amazon Basin and Congo Basin by the datasets, it is not recommended to use the products at regional extents for applications related to wetlands, while LC-CCI and GLWD-3 were found to be better suitable at a global extent. Furthermore, users should choose the product according to the wetland definition most compatible with the one used in the dataset, which implies a better understanding of different definitions a priori. A possible solution could be using the LCCS-based aquatic land cover characterization framework by Xu et al. (2020) which proposes a flexible approach capable of adapting different layers to meet the users' needs. On the other hand, researchers should continue to assess wetland products to determine their suitability over time and evaluate new improved products to highlight potential disagreements and verify their consistency.

5.2 Limitations and recommendations

To better address the inconsistencies revealed among the products, future studies can address the variable of changes of wetlands since the maps that were considered for this research were static. Indeed, there is a lack of availability for datasets that include changes in wetlands. Therefore, a product comparison that takes change into account would provide further insights into the reliability of products, which is particularly significant for wetland areas that are prone to seasonal fluctuations. Considering LC-CCI, it is recommended that future global land cover products include a wider variety of aquatic classes to better represent their complex ecosystems. Moreover, a further validation for classes that include vegetation and open water in their description, such as "intermittent lake/wetland" or "coastal wetlands", will provide a better understanding of these wetland ecosystems that are more prone to water changes. Lastly, the validation was performed with land cover datasets, which could have been biased towards LC-CCI. Thus, a more unbiased choice is recommended for future assessments, which will be inclusive of wetland types such as mangroves and peatlands, especially in regional areas. Furthermore, GLWD-3 and CW-TCI lacked a rigorous assessment of their quality, so it is suggested that future products provide an assessment for clear information needed by potential users.

6. References

Aires, F., Prigent, C., Fluet-Chouinard, E., Yamazaki, D., Papa, F., & Lehner, B. (2018). Comparison of visible and multi-satellite global inundation datasets at high-spatial resolution. *Remote Sensing of Environment*, *216*, 427–441. https://doi.org/10.1016/j.rse.2018.06.015

Arntzen, J. (2018). Makgadikgadi Wetlands (Botswana): Planning for Sustainable Use and Conservation. In C. M. Finlayson, G. R. Milton, R. C. Prentice, & N. C. Davidson (Eds.), *The Wetland Book: II: Distribution, Description, and Conservation* (pp. 487–494). https://doi.org/10.1007/978-94-007-4001-3_24

Defourny, P., Bontemps, S., Lamarche, C., Brockmann, C., & Boettcher, M. (2017). *Land Cover CCI: Product User Guide Version 2.0*.

Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, *339*(6122), 940–943. https://doi.org/10.1126/science.1229881

Fritz, S., & See, L. (2008). Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, *14*(5), 1057–1075. https://doi.org/10.1111/j.1365-2486.2007.01519.x

Gallant, A. L. (2015). The challenges of remote monitoring of wetlands. *Remote Sensing*, *7*(8), 10938–10950. https://doi.org/10.3390/rs70810938

Gao, H., & Jia, G. (2013). Assessing disagreement and tolerance of misclassification of satellite-derived land cover products used in WRF model applications. *Advances in Atmospheric Sciences*, *30*(1), 125–141. https://doi.org/10.1007/s00376-012-2037-4

Grantham, H. S., Shapiro, A., Bonfils, D., Gond, V., Goldman, E., Maisels, F., … Rickenbach, O. (2020). Spatial priorities for conserving the most intact biodiverse forests within Central Africa. *Environmental Research Letters*, *15*(9). https://doi.org/10.1088/1748-9326/ab9fae

Gregorio, A. Di. (2016). Land Cover Classification System. In *October*. Retrieved from http://www.fao.org/docrep/003/x0596e/x0596e00.htm

Harrison, I. J., Brummett, R., & Stiassny, M. L. J. (2016). The Wetland Book. *The Wetland Book*, 1–18. https://doi.org/10.1007/978-94-007-6173-5

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, *112*(5), 2538–2556. https://doi.org/10.1016/j.rse.2007.11.013

Hess, L. L., Melack, J. M., Novo, E. M. L. M., Barbosa, C. C. F., & Gastil, M. (2015). *LBA-ECO LC-07 JERS-1 SAR Flooded Wetlands and Vegetation, Amazon Basin: 1995-1996*. https://doi.org/10.3334/ORNLDAAC/1284

Hu, S., Niu, Z., & Chen, Y. (2017). Global Wetland Datasets: a Review. *Wetlands*, *37*(5), 807–817. https://doi.org/10.1007/s13157-017-0927-z

Lamarche, C., Santoro, M., Bontemps, S., d'Andrimont, R., Radoux, J., Giustarini, L., … Arino, O. (2017). Compilation and validation of sar and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community. *Remote Sensing*, *9*(1). https://doi.org/10.3390/rs9010036

Land Product Validation Subgroup. (2003). Retrieved November 25, 2020, from https://lpvs.gsfc.nasa.gov/

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. https://doi.org/10.1016/j.jhydrol.2004.03.028

Nakaegawa, T. (2012). Comparison of water-related land cover types in six 1-km global land cover datasets. *Journal of Hydrometeorology*, *13*(2), 649–664. https://doi.org/10.1175/JHM-D-10-05036.1

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., … Woodcock, C. E. (2006). Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps. In *academia.edu*. Retrieved from http://www.fao.org/gtos/gofc-gold

Tiner, R. W. (2009). Global Distribution of Wetlands. *Encyclopedia of Inland Waters*, 526–530. https://doi.org/10.1016/B978-012370626-3.00068-5

Tootchi, A., Jost, A., & Ducharne, A. (2019). Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth System Science Data*, *11*(1), 189–220. https://doi.org/10.5194/essd-11-189-2019

Tsendbazar, N. E., de Bruin, S., & Herold, M. (2017). Integrating global land cover datasets for deriving user-specific maps. *International Journal of Digital Earth*, *10*(3), 219–237. https://doi.org/10.1080/17538947.2016.1217942

Wittmann, F., & Junk, W. J. (2016). The Wetland Book. *The Wetland Book*, 1–20. https://doi.org/10.1007/978-94-007-6173-5

Wu, Q. (2017). GIS and Remote Sensing Applications in Wetland Mapping and Monitoring. In *Comprehensive Geographic Information Systems* (Vol. 3). https://doi.org/10.1016/B978-0-12-409548-9.10460-9

Xu, P., Herold, M., Tsendbazar, N. E., & Clevers, J. G. P. W. (2020). Towards a comprehensive and consistent global aquatic land cover characterization framework addressing multiple user needs. *Remote Sensing of Environment*, *250*(July), 112034. https://doi.org/10.1016/j.rse.2020.112034

Zhao, Y., Gong, P., Yu, L., Hu, L., Li, X., Li, C., … Wang, X. (2014). Towards a common validation sample set for global land-cover mapping. *International Journal of Remote Sensing*, *35*(13), 4795–4814. https://doi.org/10.1080/01431161.2014.930202

Zheng, Y., Niu, Z., Gong, P., & Wang, J. (2017). *A database of global wetland validation samples for wetland mapping*. *428*. https://doi.org/10.1007/s11434-014-0717-4

Appendix

Appendix 1 - Table A1

Thematic harmonization of wetland products

Harmonized legend		LC-CCI			CW maps		GLWD	
Class value	**LC class**	**LCCS-code**	**Class value**	**LC class**	**Class value**	**LC class**	**Class value**	**LC class**
1	Wetlands	A24	160, 170, 180	Tree cover, flooded, fresh or brackish water; tree cover, flooded, saline water; shrub or herbaceous cover, flooded, fresh/saline/brackish water	1, 2, 3	Groundwater-driven wetlands; regularly flooded wetlands; intersection of RFWs and GDW-TCI15%	4, 5, 7, 8, 10, 11, 12	Freshwater marsh, floodplain; swamp forest, flooded forest; pan, brackish/saline wetland; bog, fen, mire; 50-100% wetland; 25-50% wetland; Wetland complex (0-25% wetland)
2	Water bodies	B27, B28	210	Water bodies	4	Lakes	1, 2, 3	Lake; reservoir; river
3	Other	A11, A12, B15, B16, A24, B27, B28	0:150, 190:202, 220	Cropland, rainfed, cropland, irrigated or post-flooding; mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%); tree cover, broadleaved, evergreen, closed to open (>15%); tree cover, needleleaved, evergreen, closed to open (>15%); tree cover, needleleaved, deciduous, closed to open (>15%); tree cover, mixed leaf type (broadleaved and needleleaved); mosaic tree and shrub (>50%) / herbaceous cover (<50%); mosaic herbaceous cover (>50%) / tree and shrub (<50%); shrubland; grassland; lichens and mosses; sparse vegetation (tree, shrub, herbaceous cover) (<15%); urban areas; bare areas; permanent snow and ice	0	Non-wetland	6, 9	Coastal wetland (mangrove, estuary, delta, lagoon); intermittent wetland/lake

Appendix 2 - Table A2

Thematic harmonization of reference datasets

Harmonized legend			Global Wetland Validation Samples (GWVS)		Integrated GLC map		GLC validation sample set	
Value	**LC class**	**LCCS**	**Value**	**LC class**	**Value**	**LC class**	**Value**	**LC class**
1	Wetlands	A24	0, 1, 6, 7, 8	Coastal swamp; coastal marsh; swamp; marsh; peatland	5	Wetland vegetation	51	Wetlands/Marshland
2	Water bodies	B27, B28	3, 4, 5, 9, 10, 12, 13	Estuary; lagoon; shallow water; river; lake; water storage area; salt field and fish farm	9	Water	61, 62, 63	Water bodies/lake; water bodies/reservoir-pond; water bodies/river
3	Other	A11, A12, B15, B16, B27, B28	2, 11	Coastal mudflats; mudflats	1, 2, 3, 4, 6, 7, 8	Forest; shrubland; grassland; cropland (incl. mixtures); urban/built-up; snow and ice	11, 52	Croplands/rice fields; wetlands/mudflats

Appendix 3 - Table A3

Thematic harmonization of regional datasets

Harmonized legend		**Amazon Basin product**			**Congo Basin product**	
Value	**LC class**	**LCCS**	**Value**	**LC class**	**Value**	**LC class**
1	Wetlands	A24	13, 23, 33, 45, 55, 67, 77, 89, 99	Flooded herbaceous; flooded shrub; flooded woodland; flooded forest	67, 69, 119	Marantaceae; swamp forest; mangrove
2	Water bodies	B27, B28	11, 21, 41, 51	Open water	256	Water bodies
3	Other	A11, A12, B15, B16	0, 1, 44, 77, 88, 200	Non-wetland within Amazon Basin; non-flooded shrub; non-flooded woodland; non-flooded forest	83; 133, 138; 183, 195; 255	Dense evergreen rainforest; semi-deciduous rainforest; evergreen and deciduous rainforest; semi-deciduous rainforest with pioneer; open forest

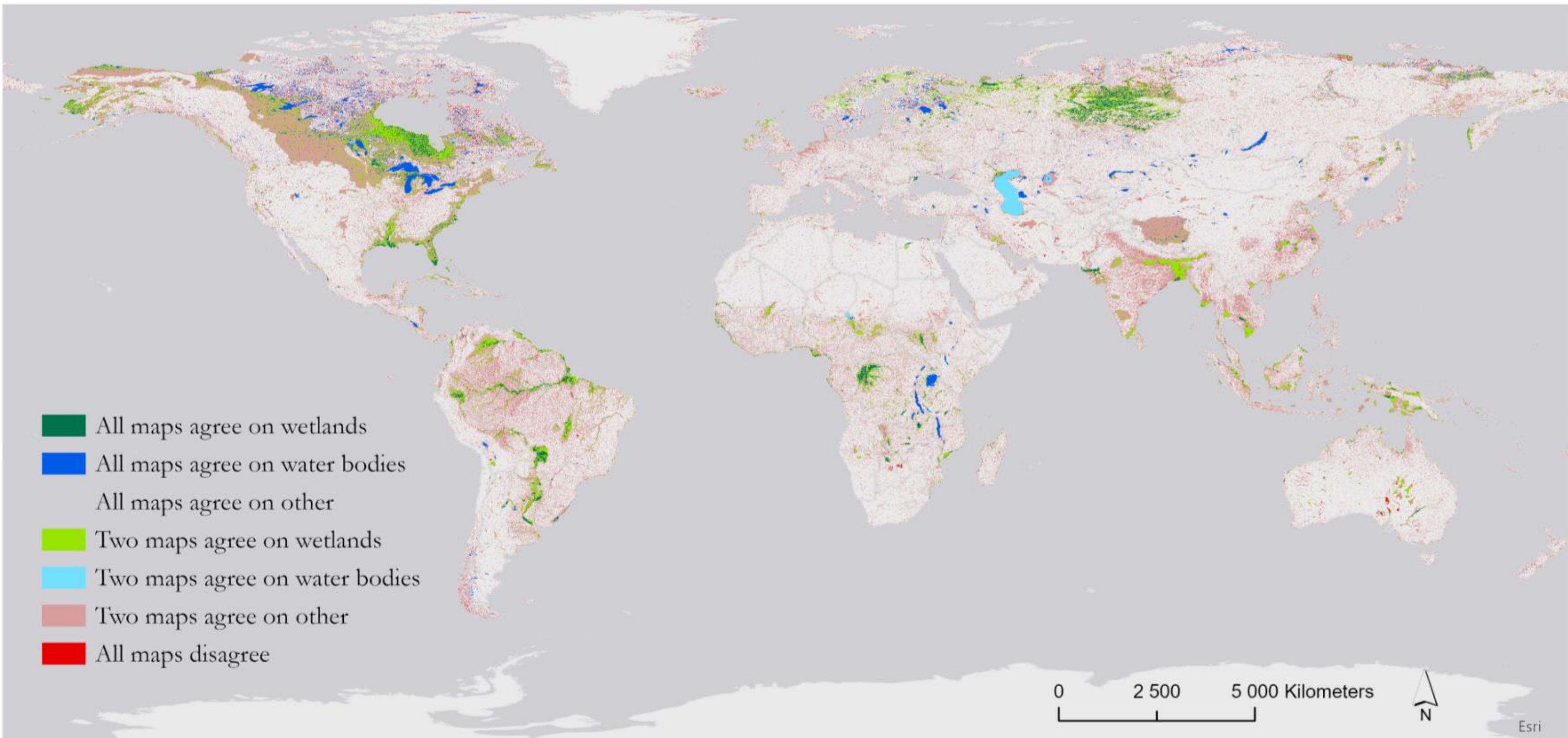
Appendix 4 - Figure A1

R code to calculate the confusion matrices for each global product

```
library(raster)
require(rgdal)


mapValues <- readOGR(dsn = "C:\\...\\Updated_hamornization_9112020.gdb", layer =
"Extract_values_GLWD_3")
refdata_df1<-data.frame("rc0" = mapValues@data[,6], "rc1" = mapValues@data[,5])


rc0_cmatrix <- as.matrix(table(refdata_df1$rc0, refdata_df1$rc1))
oa_world <- sum(diag(rc0_cmatrix))/sum(rc0_cmatrix)
name_cont = "world"
names(oa_world) <- paste("overall",name_cont)

userAccuracy <- diag(rc0_cmatrix)/rowSums(rc0_cmatrix)*100

producerAccuracy <- diag(rc0_cmatrix)/colSums(rc0_cmatrix)*100


mapValues <- readOGR(dsn = "C:\\...\\Updated_hamornization_9112020.gdb", layer = "Extract_values_LC_CCI"
)
refdata_df1<-data.frame("rc0" = mapValues@data[,6], "rc1" = mapValues@data[,5])


rc0_cmatrix <- as.matrix(table(refdata_df1$rc0, refdata_df1$rc1))
oa_world <- sum(diag(rc0_cmatrix))/sum(rc0_cmatrix)
name_cont = "world"
names(oa_world) <- paste("overall",name_cont)

userAccuracy <- diag(rc0_cmatrix)/rowSums(rc0_cmatrix)*100

producerAccuracy <- diag(rc0_cmatrix)/colSums(rc0_cmatrix)*100


refdata_df1<-data.frame("rc0" = mapValues@data[,6], "rc1" = mapValues@data[,5])
mapValues <- readOGR(dsn = "C:\\...\\Updated_hamornization_9112020.gdb", layer = "Extract_values_CW_TCI"
)
refdata_df1<-data.frame("rc0" = mapValues@data[,6], "rc1" = mapValues@data[,5])


rc0_cmatrix <- as.matrix(table(refdata_df1$rc0, refdata_df1$rc1))
oa_world <- sum(diag(rc0_cmatrix))/sum(rc0_cmatrix)
name_cont = "world"
names(oa_world) <- paste("overall",name_cont)

userAccuracy <- diag(rc0_cmatrix)/rowSums(rc0_cmatrix)*100

producerAccuracy <- diag(rc0_cmatrix)/colSums(rc0_cmatrix)*100
```

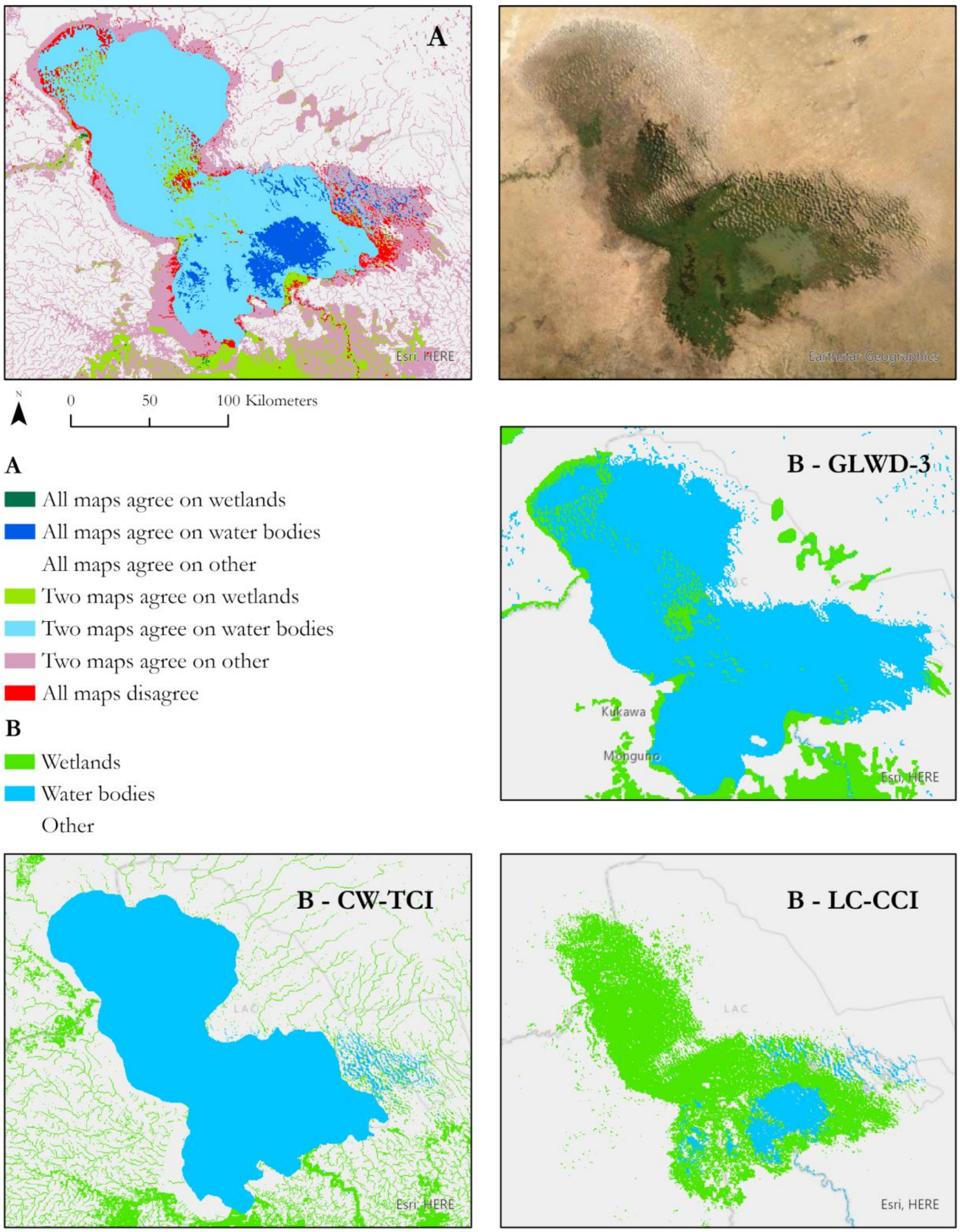
Appendix 5 - Figure A2

Global disagreement map



Appendix 6 - Figure A3

Lake Chad



Appendix 7 - Table A4

Disagreement percentages of land cover classes for each global product

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	4	Freshwater marsh, floodplain	0.7	160	Tree cover, flooded, fresh or brackish water	0.5	1	Groundwater-driven wetlands	17.4
	5	Swamp forest, flooded forest	0.1	170	Tree cover, flooded, saline water	0.1	2	Regularly flooded wetlands	51.3
	6	Coastal wetland (mangrove, estuary, delta, lagoon)	0.5	180	Shrub or herbaceous cover, flooded, fresh-saline or brackish water	3.0	3	Intersection of RFWs and GDW-TCI15%	18.9
	7	Pan, brackish/saline wetland	2.4						
	8	Bog, fen, mire	0.9						
	10	50-100% wetland	1.6						
	11	25-50% wetland	2.4						
	12	Wetland complex (0-25%)	0.1						
Water bodies	1	Lake	19.2	222	Water bodies	50.4	4	Lakes	8.5
	2	Reservoir	1.8						
	3	River	9.8						
	9	Intermittent wetland/lake	8.5						
Other	13	Other	52.0	10-40	Cropland	6.3	0	Non-wetlands	3.5
				50-100	Tree cover	17.6	5	Other	0.3
				110, 130	Grassland	2.7			
				120, 121, 122	Shrubland	2.6			

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
				140, 150, 153, 200, 201, 202	Bare/sparse vegetation	13.3			
				190	Urban	0.3			
				220	Permanent snow and ice	0.0			
				221	Ocean	3.1			

Appendix 8 - Table A5

Partial agreement percentages of each global product for class wetlands

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	4	Freshwater marsh, floodplain	18.8	160	Tree cover, flooded, fresh or brackish water	6.4	1	Groundwater-driven wetlands	27.9
	5	Swamp forest, flooded forest	7.0	170	Tree cover, flooded, saline water	1.6	2	Regularly flooded wetlands	44.6
	6	Coastal wetland (mangrove, estuary, delta, lagoon)	3.9	180	Shrub or herbaceous cover, flooded, fresh-saline or brackish water	23.5	3	Intersection of RFWs and GDW-TCI15%	23.3
	7	Pan, brackish/saline wetland	1.4						
	8	Bog, fen, mire	5.4						
	10	50-100% wetland	15.9						
	11	25-50% wetland	17.7						
	12	Wetland complex (0-25%)	2.6						
Water bodies	1	Lake	0.7	222	Water bodies	4.6	4	Lakes	0.4
	2	Reservoir	0.0						
	3	River	0.7						
	9	Intermittent wetland/lake	0.5						
Other	13	Other	25.3	10-40	Cropland	15.2	0	Non-wetlands	3.8
				50-100	Tree cover	31.4	5	Other	0.0
				110, 130	Grassland	4.6			
				120, 121, 122	Shrubland	6.3			

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
				140, 150, 152, 153, 200, 201, 202	Bare/sparse vegetation	5.1			
				190	Urban	0.6			
				220	Permanent snow and ice	0.0			
				221	Ocean	0.7			

Appendix 9 - Table A6

Partial agreement of each global product for class water bodies

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	4	Freshwater marsh, floodplain	1.8	160	Tree cover, flooded, fresh or brackish water	0.1	1	Groundwater-driven wetlands	3.9
	5	Swamp forest, flooded forest	0.3	170	Tree cover, flooded, saline water	0.0	2	Regularly flooded wetlands	8.5
	6	Coastal wetland (mangrove, estuary, delta, lagoon)	0.4	180	Shrub or herbaceous cover, flooded, fresh-saline or brackish water	1.1	3	Intersection of RFWs and GDW-TCI15%	5.8
	7	Pan, brackish/saline wetland	0.2						
	8	Bog, fen, mire	1.6						
	10	50-100% wetland	2.6						
	11	25-50% wetland	3.7						
	12	Wetland complex (0-25%)	0.2						
Water bodies	1	Lake	39.7	222	Water bodies	91.4	4	Lakes	49.9
	2	Reservoir	8.7						
	3	River	8.5						
	9	Intermittent wetland/lake	1.7						
Other	13	Other	30.6	10-40	Cropland	0.5	0	Non-wetlands	10.6
				50-100	Tree cover	2.4	5	Other	21.3
				110, 130	Grassland	0.5			
				120, 121, 122	Shrubland	0.2			

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
				140, 150, 152, 153, 200, 201, 202	Bare/sparse vegetation	3.4			
				190	Urban	0.0			
				220	Permanent snow and ice	0.0			
				221	Ocean	0.3			

Appendix 10 - Table A7

Partial agreement percentages of each global product for class other

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	4	Freshwater marsh, floodplain	3.1	160	Tree cover, flooded, fresh or brackish water	0.4	1	Groundwater-driven wetlands	44.2
	5	Swamp forest, flooded forest	1.1	170	Tree cover, flooded, saline water	0.0	2	Regularly flooded wetlands	19.1
	6	Coastal wetland (mangrove, estuary, delta, lagoon)	0.9	180	Shrub or herbaceous cover, flooded, fresh-saline or brackish water	0.7	3	Intersection of RFWs and GDW-TCI15%	7.1
	7	Pan, brackish/saline wetland	0.7						
	8	Bog, fen, mire	2.3						
	10	50-100% wetland	4.7						
	11	25-50% wetland	8.6						
	12	Wetland complex (0-25%)	2.1						
Water bodies	1	Lake	1.2	222	Water bodies	1.0	4	Lakes	0.8
	2	Reservoir	0.2						
	3	River	0.2						
	9	Intermittent wetland/lake	1.3						
Other	13	Other	73.4	10-40	Cropland	24.3	0	Non-wetlands	27.5
				50-100	Tree cover	39.0	5	Other	1.2
				110, 130	Grassland	9.1			
				120, 121, 122	Shrubland	7.8			

	GLWD-3			LC-CCI			CW-TCI		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**	**Value**	**LC class**	**%**
				140, 150, 152, 153, 200, 201, 202	Bare/sparse vegetation	14.1			
				190	Urban	1.0			
				220	Permanent snow and ice	0.6			
				221	Ocean	2.1			

Appendix 11 - Table A8

Disagreement percentages of Amazon Basin product compared to GLWD-3 product

	GLWD-3			LBA-ECO LC-07		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	4	Freshwater marsh, floodplain	1.7	13	Aquatic macrophyte (flooded herbaceous)	0.2
	5	Swamp forest, flooded forest	32.3	23	Aquatic macrophyte (flooded herbaceous)	7.8
	7	Pan. brackish/saline wetland	0.0	33	Aquatic macrophyte (flooded herbaceous)	5.4
				45	Flooded shrub	3.8
				55	Flooded shrub	0.1
				67	Flooded woodland	1.9
				77	Flooded woodland	5.1
				89	Flooded forest	16.7
				99	Flooded forest	8.1
Water bodies	1	Lake	2.0	11	Open water	3.9
	2	Reservoir	0.0	21	Open water	1.5
	3	River	7.4	41	Open water	0.3
	9	Intermittent wetland/lake	0.0	51	Open water	0.2
Other	13	Other	56.6	1	Non-wetland within Amazon Basin	30.7
				44	Non-flooded shrub	0.1
				77	Flooded woodland	5.1
				88	Non-flooded forest	13.5
				200	Elevation >= 500 m, in Basin	0.8

Appendix 12 - Table A9

Disagreement percentages of Amazon Basin product compared to LC-CCI product

	LC-CCI			LBA-ECO LC-07		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	160	Tree cover, flooded, fresh or brackish water	6.9	13	Aquatic macrophyte (flooded herbaceous)	0.4
	170	Tree cover flooded, saline water	0.0	23	Aquatic macrophyte (flooded herbaceous)	10.4
	180	Shrub or herbaceous cover, flooded, fresh-saline or brackish water	4.8	33	Aquatic macrophyte (flooded herbaceous)	7.3
				45	Flooded shrub	5.0
				55	Flooded shrub	0.1
				67	Flooded woodland	2.4
				77	Flooded woodland	6.5
				89	Flooded forest	22.9
				99	Flooded forest	11.4
Water bodies	222	Water bodies	6.4	11	Open water	2.4
				21	Open water	1.7
				41	Open water	0.3
				51	Open water	0.2
Other	10-40	Cropland	6.6	1	Non-wetland within Amazon Basin	13.7
	50-100	Tree cover	64.9	44	Non-flooded shrub	0.1
	110, 130	Grassland	5.4	77	Flooded woodland	6.5
	120	Shrubland	4.7	88	Non-flooded forest	14.6
	150, 153, 200	Bare/sparse vegetation	0.1	200	Elevation >= 500 m, in Basin	0.5
	190	Urban	0.1			
	220	Permanent snow and ice	0.0			
	221	Ocean	0.1			

Appendix 13 - Table A10

Disagreement percentages of Amazon Basin product compared to CW-TCI product

	CW-TCI			LBA-ECO LC-07		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	1	Groundwater-driven wetlands	84.3	13	Aquatic macrophyte (flooded herbaceous)	0.1
	2	Regularly flooded wetlands	4.6	23	Aquatic macrophyte (flooded herbaceous)	0.7
	3	Intersection of RFWs and GDW-TCI15%	5.0	33	Aquatic macrophyte (flooded herbaceous)	0.8
				45	Flooded shrub	0.7
				55	Flooded shrub	0.0
				67	Flooded woodland	0.2
				77	Flooded woodland	1.2
				89	Flooded forest	3.2
				99	Flooded forest	1.4
Water bodies	4	Lakes	0.3	11	Open water	1.6
				21	Open water	0.4
				41	Open water	0.1
				51	Open water	0.1
Other	0	Non-wetlands	5.8	1	Non-wetland within Amazon Basin	80.4
				44	Non-flooded shrub	0.0
				77	Flooded woodland	1.2
				88	Non-flooded forest	4.7
				200	Elevation >= 500 m, in Basin	4.5

Appendix 14 - Table A11

Disagreement percentages of Congo Basin product compared to GLWD-3 product

	GLWD-3			Congo Basin product		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	4	Freshwater marsh, floodplain	54.8	67	Marantaceae	0.4
	5	Swamp forest, flooded forest	5.9	69	Swamp forest	26.6
	6	Coastal wetland	1.5	119	Mangrove	0.3
	7	Pan, brackish, saline wetland	0.1			
Water bodies	1	Lake	3.9	256	Water bodies	4.4
	2	Reservoir	0.4			
	3	River	4.3			
	9	Intermittent wetland/lake	0.6			
Other	13	Other	28.4	83	Dense evergreen rainforest	0.7
				133	Semi-deciduous rainforest	4.9
				138	Evergreen and deciduous rainforest	1.5
				183	Semi-deciduous rainforest with pioneer	1.1
				195	Open forest	5.7
				255	Other	54.2

Appendix 15 - Table A12

Disagreement percentages of Congo Basin product compared to LC-CCI product

	LC-CCI			Congo Basin product		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	160	Tree cover, flooded, fresh or brackish water	17.0	67	Marantaceae	0.6
	170	Tree cover, flooded, saline water	1.3	69	Swamp forest	39.2
	180	Shrub or herbaceous cover, flooded, fresh-saline or brackish water	27.4	119	Mangrove	0.4
Water bodies	222	Water bodies	9.8	256	Water bodies	5.9
Other	10-40	Cropland	5.2	83	Dense evergreen rainforest	1.0
	50-100	Tree cover	37.4	133	Semi-deciduous rainforest	2.5
	110, 130	Grassland	0.5	138	Evergreen and deciduous rainforest	0.9
	120, 122	Shrubland	1.1	183	Semi-deciduous rainforest with pioneer	1.2
	150, 153, 200, 201, 202	Bare/sparse vegetation	0.0	195	Open forest	14.5
	190	Urban	0,1	255	Other	33.7
	221	Ocean	0,2			

Appendix 16 - Table A13

Disagreement percentages of Congo Basin product compared to CW-TCI product

	CW-TCI			Congo Basin product		
	Value	**LC class**	**%**	**Value**	**LC class**	**%**
Wetlands	1	Groundwater-driven wetlands	75.4	67	Marantaceae	0.1
	2	Regularly flooded wetlands	9.9	69	Swamp forest	5.2
	3	Intersection of RFWs and GDW-TCI15%	7.6	119	Mangrove	0.0
Water bodies	4	Lakes	1.0	256	Water bodies	1.5
Other	0	Non-wetlands	6.1	83	Dense evergreen rainforest	0.8
	5	Other	0.0	133	Semi-deciduous rainforest	15.7
				138	Evergreen and deciduous rainforest	3.0
				183	Semi-deciduous rainforest with pioneer	3.1
				195	Open forest	13.5
				255	Other	57.1