Global Wetland Datasets: a Review

Article *in* Wetlands · July 2017

DOI: 10.1007/s13157-017-0927-z

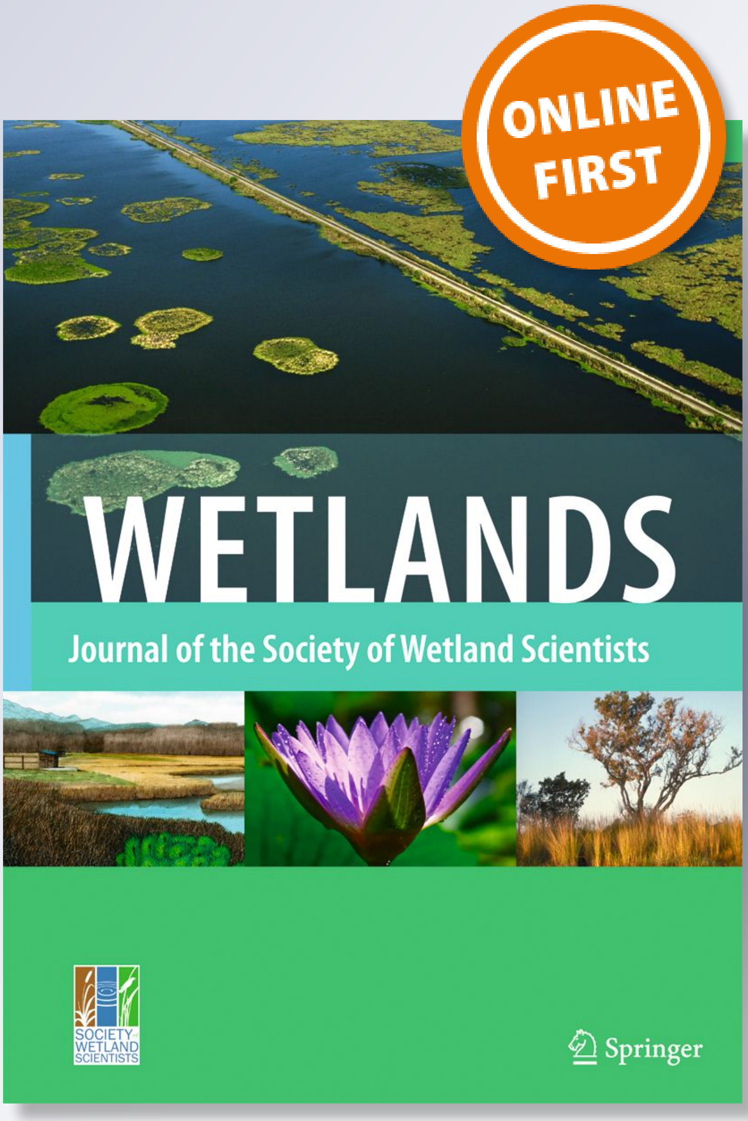
Global Wetland Datasets: a Review

Shengjie Hu, Zhenguo Niu & Yanfen Chen

Wetlands
Official Scholarly Journal of the Society of Wetland Scientists

ISSN 0277-5212

Wetlands
DOI 10.1007/s13157-017-0927-z



Wetlands
DOI 10.1007/s13157-017-0927-z

REVIEW ARTICLE

Global Wetland Datasets: a Review

Shengjie Hu[1] **· Zhenguo Niu**[1] **· Yanfen Chen**[1]

Received: 14 November 2016 / Accepted: 18 June 2017

Abstract Accurate wetland datasets are indispensable for generating polices on wetland conservation and appropriate land uses, global climate change studies, and biodiversity conservation. Although increasing numbers of global wetland-related datasets have been established, the extensive disagreements among these datasets are prominent. In particular, estimates of global wetland area range from 0.54 to 21.26 million km^2 and the class-specific spatial consistency of wetlands is less than 1%. The different definitions of 'wetland' and the intrinsic features of wetlands contribute to this extensive inconsistency. Given the various requirements of wetland-oriented data products, it is important to conduct comprehensive wetland mapping at global scale. The Ramsar wetland definition is recommended and a hierarchical and flexible structure of wetland classification system is preferred for future global wetland datasets. Time-series satellite imagery at 250–1000 m spatial resolution is preferred to characterize wetland dynamics by combination of passive/active SAR data and other ancillary data, such as topography, climate, and soil data. The classification tree method for classification of massive satellite imagery and big-data technology for sample datasets are promising for the enhancement of wetland map product accuracy.

Keywords Global wetland-related datasets · Wetland classification · Wetland inconsistency · Global wetland mapping

✉ Zhenguo Niu
niuzg@radi.ac.cn

[1] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Introduction

Wetlands, distributed in all climate zones ranging from the tropics to the tundra, are among the most productive and valuable ecosystems in the world (Mitsch and Gosselink 1993). The total value of services provided by global wetland ecosystems is nearly half of the total value of services that global ecosystems are estimated to provide (Costanza et al. 1997). Wetlands not only offer fundamental materials for human well-being but also have the potential to generate considerable economic value. According to the Millennium Ecosystem Assessment, capture fisheries in coastal waters alone contribute $34 billion to gross world product annually (Finlayson et al. 2005). Therefore, wetlands can be considered indispensable resources for humans.

With the rapid population growth and economic development, wetland disappearance and degradation has occurred worldwide (Davidson 2014; Dixon et al. 2016). Wetland loss is not only a global environmental and ecological issue; it is also closely related to social development. Reliable wetland datasets would help us to understand the location, area, distribution, and condition of global wetlands, and the decision-makers to develop effective conservation policies and management plans to establish appropriate conservation/management priorities. However, there is little knowledge on recent changes to global wetlands, such as global wetland distribution, loss and status. The typically cited figure of global wetland loss is about 50%. But Davidson (2014) in his research pointed out that the typically cited figure of 50% global wetland loss originated from very limited data from the USA only for the mid-twentieth century and was then erroneously restated as a global figure. By reviewing 189 reports related to changes in wetlands, Davidson (2014) concluded that the long-term loss rate of global wetlands is about 54%–57%. However, there are uncertainties associated with this estimation because there are geographical biases in the numbers of published reports among different regions and different spatial scales and inconsistencies in the available data related to the change of wetland areas.

As a severe environmental issue, climate change is considered to be a threat to species survival, the health of natural systems, and human well-being (Hulme et al. 2005). Wetlands represent a land cover class that is sensitive to climate change. On one hand, through potential or near potential evapotranspiration wetlands exchange water, heat, and energy with atmosphere and exert a direct impact on both global and local/regional climate (Fan et al. 2010). On the other hand, through the sequestration and emission of greenhouse gases wetlands participate in global biogeochemical cycle and regulate both global and local/regional climate indirectly (Russi et al. 2013). As the largest carbon sink and the main natural sources of methane (CH_4) in the world (Takai 1970; Parish and Looi 1999; Russi et al. 2013), wetland type, area, distribution, and the timing and extent of inundation are essential for the calculation of greenhouse gas emissions and carbon storage (Matthews and Fung 1987; Aselmann and Crutzen 1989). Wetland datasets are important input variables in models focusing on the simulation and prediction of global climate change (Boer et al. 2000).

Wetlands provide habitats for a range of birds, fish, and plants, including many threatened and endangered species (Finlayson et al. 2005). The distribution and condition of wetlands is usually associated with the distribution and health of species. Changes to wetland area and distribution would cause shifts in species distribution, changes to migratory routes, and alterations in community structures and species interactions (Erwin 2009). Taking wetland waterfowl as an example, a variety of habitats such as permafrost, forest, marshes, mangroves, and red coral are needed for the annual migration of waterfowl (Browne and Dell 2007). The loss of wetland along the migration route will force large numbers of waterfowl to change their migration routes or even lead to loss of species (Barbier et al. 1997). Wetland datasets are critical in simulating species distribution, exploring the patterns of migration, and establishing natural reserves. Therefore, it is important to develop accurate wetland datasets for wetland management and decision-making policies.

The increasing recognition of the importance of wetlands has stimulated great interest in mapping global wetland distribution. In 1996, the IGBP (*International Geosphere-Biosphere Program*) wetlands workshop developed a wetland parameterization scheme to better incorporate wetlands into global land surface characterization schemes (Sahagian and Melack 1997). In this report, the authors reviewed the available global wetland datasets, including the maps developed by Matthews and Fung (1987) and Aselmann and Crutzen (1989), and found the figure of global wetland extent was uncertain. Finlayson et al. (1999) summarized a selection of international, regional, and national wetland inventories and identified priorities for improving the global wetland inventory. Mitra et al. (2005) reviewed and compared four global databases (Matthews natural wetlands database, ISLSCP (*International Satellite Land Surface Climatology Project*) database, IGBP DISCover (*Data Information System*) database and the Ramsar database) and found that a substantial disagreement on spatial distribution was present among these datasets. However, as the overview of global wetland datasets was just a basic part for the previous researches, these studies analyzed a few of global wetland-related datasets and mainly focused on the differences of the global wetland area. The causes of differences in the current datasets have not been adequately explored either. With the rapid development of Earth observation technologies and other information technologies, an increasing number of global wetland-related datasets have been developed, but the knowledge of the total global wetland area, distribution and their conditions remains limited. To understand the situation of global wetlands, in this paper we (1) summarize the currently available wetland-related datasets at a global scale; (2) examine the inconsistencies among these datasets and their causes; (3) discuss the issues related to global wetland mapping; and (4) suggest some recommendations for the future generation of a comprehensive global wetland dataset.

Global Wetland-Related Datasets

According to the Ramsar Convention, wetlands are "areas of marsh, fen, peat land, or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish, or salty, including areas of marine water the depth of which at low tide does not exceed six meters" (Ramsar Information Bureau 1998). The scope of the Ramsar definition is broader than the plethora of wetlands definitions developed based on different areas of expertise or interest. Therefore, this definition is more suitable for the review of global wetland-related datasets which have different wetland classification systems.

Based on the establishment methods of wetland datasets, the global wetland-related datasets can be divided into three groups: (1) remote-sensed datasets, which map wetlands through remote sensing (RS) classification; (2) compilation datasets, which are developed by aggregating various historical independent datasets; and (3) simulation datasets, which simulate wetlands using models such as hydrological models or Land Surface Models (LSM) (Table 1).

Remote Sensing-Based Datasets

The RS classification has many advantages for wetland inventory at a large scale (Ozesmi and Bauer 2002). For instance,

Table 1 Wetland categories and area in those wetland-related global datasets (unit of area: million km^2)

Type	Methods	Datasets	Total area	Categories	Years of data	Spatial resolution	Reference
1. Remote Sensing	Artificial visual interpretation	GlobalLand 30	7.17	wetlands; water	2000	30-m	Chen et al. (2015)
		GlobalLand 30	7.01	wetlands; water	2010	30-m	
	Supervised Classification	BU-MODIS	3.88	permanent wetlands; water bodies	2000	1-km	Friedl et al. (2002)
		GLCNMO2	4.57	wetlands (wetland; mangrove); water	2008	500-m	Tateishi et al. (2014)
	Unsupervised Classification	DISCover	4.56	permanent wetlands; water bodies	1992–1993	1-km	Loveland et al. (2000)
		GLCC.I	4.56	permanent wetlands; water bodies	1992–1993	1-km	
		GLC2000	6.04	wetlands (tree cover, regularly flooded, fresh water; tree cover, regularly flooded, saline water; regularly flooded shrub and/or herbaceous cover);water bodies	2000	1-km	Bartholomé and Belward (2005)
		GLOBCOVER2005	5.10	wetlands (Closed to open(>15%) broadleaved forest regularly flooded (semi-permanently or temporarily-fresh or brackish water; close(>40%) broadleaves forest or shrubland permanently flooded-saline or brackish water; closed to open (15%) grassland or woody vegetation on regularly flooded or waterlogged soil-fresh, brackish or saline water); water bodies	2005	300-m	Bicheron et al. (2006)
		GLOBCOVER2009	5.00	wetlands (Closed to open(>15%) broadleaved forest regularly flooded (semi-permanently or temporarily-fresh or brackish water; close(>40%) broadleaves forest or shrubland permanently flooded-saline or brackish water; closed to open (15%) grassland or woody vegetation on regularly flooded or waterlogged soil-fresh, brackish or saline water); water bodies	2009	300-m	Sophie and Pierre (2010)
		GIEMS	2.12–5.86	natural wetlands, irrigated rice, and lakes/rivers	1993–2000	1° × 1°	Prigent et al. (2007)
		GIEMS-D15	6.5–17.3	Mean annual minimum; mean annual maximum; long-term maximum	1993–2004	500 m	Fluet-Chouinard et al. (2015)
		Pekel et al.	4.48	surface water	1984–2015	30 m	Pekel et al. (2016)
2. Compilation	Compiling feature datasets	Matthews and Fung	5.30	forested bog; non forested bog; forested swamp; non forested swamp; alluvial formations; water body	1987	1° × 1°	Matthews and Fung (1987)
		Asekmann et al.	5.7	bogs, fens, swamps, marshes, floodplains, shallow lakes and rice paddies	1989	2.5° × 5°	Aselmann and Crutzen (1989)
		GLWD-3	10.7–12.7	wetlands (freshwater marsh, floodplain; swamp forest, flooded forest; coastal wetland; pan, brackish/saline wetland; bog, fen, mire; intermittent wetland/lake; 50–100% wetland; 25–50% wetland; wetland complex(0–25% wetland));water (lake; reservoir; river)	1980s	30 arc-second	Lehner and Döll (2004)
		GLCNMO1	2.82	wetlands (wetland; mangrove); water	2003	1-km	Tateishi et al. (2011)
		Tuanmu	4.31	regularly flooded vegetation; open water	2000–2005	1-km	Tuanmu and Jetz (2014)
3. Model simulation	Hydrological models	Fan	21.26	wetlands	Potential distribution	1-km	Fan et al. (2010, Fan et al. 2013)

Table 1 (continued)

Type	Methods	Datasets	Total area	Categories	Years of data	Spatial resolution	Reference
		Zhu	18.51	wetlands	Potential distribution	30 arc-second	Zhu and Gong (2014)
		Hu	29.83	wetlands with no open water; open water	Potential distribution	1-km	Hu et al. (2017)
	Land surface models	WETCHIMP	2.7–8.17	wetlands	2007–2011	0.5° to 100-km	Melton et al. (2012)

using RS imagery for land cover classification is economic and efficient. It is possible to monitor and inventory wetlands seasonally and annually through repeated observation by satellites. The RS-based method is considered to be the most desirable technique for wetland inventory at global scale.

Early aerial photography and satellite imagery identified wetlands mainly through artificial visual interpretation (Nayak and Sahai 1985). Because of the complexity of wetland ecosystems and the small number of spectral bands of images, visual interpretation based on expert knowledge is more effective and reliable than computer auto-classification. The early wetland inventory of the US was mainly based on visual interpretation of aerial photography (Dahl 1990). Although manual interpretation generally has high precision, this approach is time-consuming and subjective (Phinn et al. 1999; Baker et al. 2006; Wright and Gallant 2007). Dataset updating in this manner is also very expensive and slow. For example, the Global Land Cover (GlobalLand30) products ca. 2000 and 2010 with 30-m spatial resolution (Chen et al. 2015) employed visual interpretation and involved several hundreds of people for four years. However, only two wetland categories (water and wetlands) were included in this dataset. Therefore, this approach is generally limited to local and small regions. Moreover, it is typically employed to collect training and validation samples in line with the development of computerized classification (Gong et al. 2012).

Computerized classification methods comprise supervised and unsupervised approaches. Supervised classification uses training samples to recognize the different classes (Ozesmi and Bauer 2002; Schowengerdt 2012). The advantage of this method is the ability to specify the desired class types. However, it has some limits that the desired classes may not correspond to spectrally unique classes, and the acquisition of training data is time-consuming and expensive (Ozesmi and Bauer 2002). The Moderate Resolution Imaging Spectroradiometer (MODIS) land cover classification algorithm (MLCCA), which used a supervised decision tree classification approach to 5 months of MODIS data, was used to develop the MODIS land cover product (BU-MODIS) at 1-km spatial resolution (Moody and Strahler 1993). Both permanent wetlands of large areal extent and water bodies were contained in this dataset (Friedl et al. 2002). A 500-m global land cover dataset, Global Land Cover by National Mapping Organizations 2008 (GLCNMO Version 2), was developed by the International Steering Committee for Global Mapping (ISCGM). In GLCNMO Version 2, the classification method by using the Tasseled Cap Transformation and supervised decision tree has been employed for six land cover classes, including wetland, water, and mangrove (Tateishi et al. 2014).

The unsupervised classification (clustering) method groups together pixels with similar spectral values and labels clusters with specific classes based on ancillary information (Ozesmi and Bauer 2002; Schowengerdt 2012). This approach eliminates the time-consuming training step and the classes are distinct units. However, the clusters may not correspond to desired class types (Ozesmi and Bauer 2002). A 1-km global land cover database (DISCover) and a 1-km global land cover characteristics database (GLCC.I) were developed for the IGBP-DIS initiative, in which unsupervised classification of monthly Advanced Very High Resolution Radiometer Normalized Difference Vegetation Index (AVHRR-NDVI) of 1992 was used and two wetland categories (permanent wetland and water body) have been identified (Loveland et al. 2000). Another global land cover product for the year 2000 (GLC-2000) has been produced using SPOT/VEGETATION data based on the 'regionally tuned' approach by an international partnership of 30 research groups (Bartholomé and Belward 2005). In this dataset, wetlands were extended from permanent to regularly flooded shrub/herbaceous. Building on the success of the GLC-2000 project, the European Space Agency launched the GlobCover initiative and released higher resolution global land cover (GlobCover2005 and GlobCover2009) product using 300-m resolution ENVISAT/MERIS data for 2005–2006 and 2009 (Bicheron et al. 2006; Sophie and Pierre 2010). Compared with GLC-2000, GlobCover products have not only improved the resolution but also provide more detailed wetland categories (Table 1).

However, optical and infrared RS are unable to penetrate clouds and dense vegetation cover, particularly in tropical or sub-tropical regions, representing a major limitation of this approach. Prigent et al. (2007) merged the passive and active microwave along with visible and infrared observations through an unsupervised clustering technique to detect global inundation which included natural wetland and irrigated rice at 2.5° resolution (GIEMS). Based on this time-series products, Fluet-Chouinard et al. (2015) established a new inundated dataset (GIEMS-D15) using a downscaling method to improve the spatial resolution of the Prigent's result from 2.5° to 500 m. In 2016, by using three million Landsat satellite images and expert systems, Pekel et al. (2016) has quantified changes in global surface water over the past 32 years at 30 m resolution and found that nearly 3 % (4.46 million km^2) of the earths' landmass was under water at some time between 1984 and 2015. Through the Deltares Aqua Monitor tool, Donchyts et al. (2016) has explored the earth's surface water change over the past 30 years. But they only provide the gains and losses of the global surface water area and did not provide the global surface water area for each year.

Compilation Datasets

Another way to establish global wetland datasets is to compile historical data, including independent feature datasets (water, vegetation and soil) and existing wetland-related datasets (wetland maps and land cover products). In order to estimate the area, location and environmental features of global wetlands, Matthews and Fung (1987) combined three independent global digital datasets (vegetation, soil properties and fractional inundation) and developed a global dataset of natural wetland at 1° resolution. By compiling published information and various maps that drew from regional wetland surveys and monographs, Aselmann and Crutzen (1989) created a dataset on the distribution and seasonality of global freshwater wetlands and rice paddy. Another recent global wetland database is Global Lakes and Wetland Database Level-3 (GLWD-3), which was developed in 2004 by combination of the best available sources for lakes and wetland on a global scale with a high spatial resolution of 30 arc-seconds (Lehner and Döll 2004). Because of the difficulty of wetland extraction from MODIS data by the supervised classification, the global land cover dataset, GLCNMO version1, has taken the common wetland areas in GLC2000 and IGBP-DISCover as wetlands (Tateishi et al. 2011). Aiming for biodiversity and ecosystem modeling, Tuanmu and Jetz (2014) integrated four global land cover products (DISCover, GLC2000, BU-MODIS, and GLOBCOVER 2005) and developed a global 1-km consensus land cover product using a generalized classification scheme and an accuracy-based integration approach. Therefore, this dataset inherited two wetland categories: regularly flooded vegetation and open water.

Simulation Datasets

Water is a key factor in wetland occurrence. Topography, climate and soil features have impacts on the spatial-temporal distribution of water simultaneously and their combined effect controls wetland formation and distribution. Some hydrological models have been developed based on this relationship to simulate the areal extent and spatial distribution of wetlands. For example, Fan et al. (2013) and Zhu and Gong (2014) simulated global wetland at 1-km spatial resolution based on the relationship between water table depth and wetland distribution, respectively. Hu et al. (2017) simulated the potential distribution of global wetlands using a new precipitation topographic wetness index.

The increasing awareness of the major role wetlands play in the global climate system has led an increasing number of LSMs which take account of wetlands in their schemes. The Wetland and Wetland CH_4 Inter-comparison of Models Project (WETCHIMP) has investigated wetland CH_4 models for simulations of large-scale wetland characteristics and corresponding CH_4 emissions (Melton et al. 2012). This study showed that these LSMs determined wetland area by prescribing extents, parameterization/forcing with a remotely sensed inundation dataset, or model wetland location via the model's hydrological model. For example, the SDGVM and UVic-ESCM used their internal hydrological model to determine the location of wetlands (Hopcroft et al. 2011; Avis et al. 2011), whereas others such as CLM4Me and DLEM prescribed wetland extent using remotely sensed inundation datasets (Riley et al. 2011; Xu and Tian 2012).

Inconsistencies among Global Wetland-Related Datasets

Inconsistencies of global land cover products have been confirmed by several research studies (Giri et al. 2005; Jung et al. 2006; Herold et al. 2008). As complex earth ecosystems, wetland identification is the most difficult task during land cover mapping (Niu 2015; Gong et al. 2016). In this section, the inconsistencies among various datasets on global wetland areas and spatial distribution are being examined.

Comparison of Wetland-Related Datasets

The area extent of global wetlands varies greatly among different datasets from 0.54 million km^2 to 29.83 million km^2 (Table 1). Even if datasets are produced by the same kind of approach, the wetland areas in these datasets remain considerably different. For example, the wetland area in remote-sensed products ranged from 2.12 to 17.3 million km^2 (Bartholomé and Belward 2005; Tateishi et al. 2011), and in compilation datasets this figure varied from 2.82 to 12.7 million km^2. The difference in wetland area among different kind of datasets is even more prominent. As shown in Table 1, the wetland areas in simulation results are commonly larger than those of other types of datasets. In the 1999 global wetland inventory report, Finlayson et al. (1999) estimated the global wetland area to be between 7.48 and 12.79 million km^2.

The disagreement in spatial distribution among global wetland-related datasets is also prominent. Köchy and Freibauer (2009) compared three different datasets, which include Matthews and Fung, the land cover map from the International Satellite Land Surface Climatology Project (ISLSCP map) and DISCover dataset. The result showed that Matthews and Fung and the ISLSCP map only matched 57%, and the match among Matthews and Fung's map, the ISLSCP map and the DISCover dataset was even lower. Nakaegawa (2012) compared three water-related land cover types (snow and ice, wetland, and open water) in six 1-km global land cover datasets (GLCC.S, GLCC.I, GLC2000, BU-MODIS, GLCNMO Version 1 and GLWD-3) by calculating the class-specific consistency. The result indicated that the agreement for open water is about 67%, but the value for wetlands without surface water is much lower at only about 30%. However, in Nakaegawa (2012)'s study, the class-specific consistency was calculated just between each pair of datasets instead of the six datasets. Since GLCC.S, GLCC.I, and DISCover were all established under the IGBP framework and the GLCNMO version1 has taken the common wetland areas in GLC2000 and DISCover, therefore, in this paper, we have selected four wetland-related datasets (GLCC.I, GLC2000, BU-MODIS, and GLWD-3) whose data resolution were equal and the data collection periods were near 2000 (Table 1) and calculated the class-specific consistency among these four datasets pixel by pixel to check whether the same pixel had the same type at the same time. If a pixel showed identical categories among the four datasets simultaneously, the consistency was 100%; if three datasets indicated the same category in the pixel, the consistency was 75%, and so on. This approach showed a serious inconsistency in that the spatial consistency of water among those four datasets was about 23% and the situation of wetlands without surface water was less than 1% (Fig. 1).

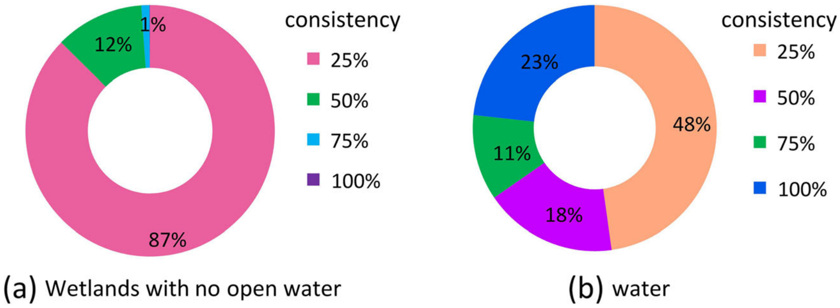
Reasons for the Inconsistencies

Differences in Wetland Definition

Although wetlands can be commonly regarded as the transitional area between terrestrial and aquatic ecosystems (Kent 1996; Klemov 1998), there are numerous different definitions of wetlands such as 'bogs and fens' (peat-accumulating wetlands), 'marshes' (herbaceous, frequently inundated wetlands) or 'swamps' (forested wetlands), and no standardization of these terms (Mitra et al. 2005). Different datasets that serve for different purposes have a different understanding of the 'wetland' concept and apply different wetland classification systems. For example, remote-sensed datasets mainly come from global land cover products, which have been primarily directed towards identifying land surface characteristics (e.g. vegetation features) (Gumbricht 2012) and only permanent wetlands and some regularly flooded areas are considered in the classification system of RS datasets (Table 1). In contrast, compilation datasets developed under the demands of calculating CH_4 emissions have more detailed categories extending from natural wetlands to artificial wetlands. This inconsistency in wetland definition then leads to disagreement on area extent and spatial distribution among different wetland datasets.

Intrinsic Features of Wetlands

Wetlands are characterized by seasonal and annual variation in hydrology (Loveland et al. 2000; Melton et al. 2012). The areal extent and spatial distribution of wetlands may change considerably across different seasons and years (Prigent et al. 2007; Fluet-Chouinard et al. 2015). It is difficult to distinguish seasonal wetlands if only single-data imagery was employed in wetland mapping. Moreover, the chronological inconsistency of input data sources could result in disagreement among different wetland-related datasets (Fluet-Chouinard et al. 2015). Therefore, most of the currently remote-sensed wetland datasets that are static in time could not identify seasonal wetlands well. Global inundated data are time-series datasets and can reflect the dynamic changes of wetlands. However, its coarse resolution and the uncertainties such as the underestimation of GIEMS in some forested regions and low accuracy of GIEMS-D15 in mountainous regions (Prigent et al. 2007;



(a) Wetlands with no open water (b) water

Fig. 1 Class-specific consistency among four global wetland-related datasets

Fluet-Chouinard et al. 2015) have influnced the reliability of its area extent and spatial distribution.

Another feature of wetlands is their transitional environment characteristics, which come from their geospatial location between terrestrial ecosystem and aquatic ecosystem. The spectrum of wetlands is always a spectra mixture of water, soil, and vegetation, which pose great challenges to wetland mapping by RS. The sharp spatial-temporal changes of landscape elements often results in confusion between wetland classes. Therefore, different methods of wetland identification and different input data always produce variable results. For instance, the classification method used by IGBP-DISCover was unsuitable to identify wetlands, which lead to an underestimation of wetland area (Loveland et al. 2000). Furthermore, as optical imageries could not penetrate dense vegetation canopies, the spectrum of swamp and mangrove are similar to the forest spectrum in optical imageries. Underestimation of the wetlands in tropical or sub-tropical areas will occur if only optical imageries are used in wetland mapping (Gumbricht 2012).

Discussions for Global Wetland Mapping

In terms of quantity it seems that datasets related to global wetlands are adequate; however, in terms of quality, accurate datasets specifically for wetlands are scarce. And these currently available datasets are widely dispersed among land cover products (Loveland et al. 2000; Friedl et al. 2002; Bartholomé and Belward 2005), water products (Sharma et al. 2015; Donchyts et al. 2016; Min et al. 2016; Pekel et al. 2016), peatland datasets (Van Engelen and Huting 2002; Kleinen and Brovkin 2013) and mangrove maps (Spalding et al. 1997), which are either static in time or coarse in spatial resolution and have extensive disagreement of wetlands among them (Ozesmi and Bauer 2002; Finlayson et al. 2005; Nakaegawa 2012). Given the importance of wetland dataset in policy decision-making, global climate change studies, and biodiversity conservation, it is necessary to develop an accurate and specific dataset for global wetlands. While wetland thematic mapping is not a new topic, it is more often performed at local scale. Global wetland thematic mapping, which will have more higher accuracy of wetland categories and can provide more detailed ecological and environmental features of the wetland ecosystem, requires more attention on the following aspects.

Wetland Conception and Classification System

The wetland concept and classification system are essential prerequisites for wetland thematic mapping capable of encompassing the diversity of wetland subclasses. As mentioned above, different wetland concepts were adopted in those global wetland-related databases, most of which merely focused on wetland area and distribution, and only comprised a few wetland categories, such as regularly flooded grass and open water. This obviously cannot meet the requirements of wetland conservation and other applications. The classification system of wetlands and deep-water habitats developed for the U.S. Fish and Wildlife Service (Cowardin et al. 1989) is widely regarded as the one of the most comprehensive and versatile wetland classification systems (Finlayson and van der Valk 1995); however, it is too complicated to be operational at a large scale (Scott 2010). Up to now there is no corresponding global dataset to this wetland classification system. By contrast, the Ramsar wetland definition has been accepted by many organizations (Blanco et al. 2013; Navid 2014). Although its wetland classification system has been criticized for being too embracing (Mitsch and Gosselink 1993; Paul 2000; Mitra et al. 2003; Yin et al. 2014), it is suitable for wetland management. The aim of a comprehensive global wetland dataset is well suited to this definition because most wetland categories in those existing global wetland-related datasets are included. Furthermore, the global wetland classification system should be hierarchical and have a flexible structure in which (besides the traditional information such as location, size, and distribution), more information on biogeochemical features of wetlands can be considered.

Mapping Method

Construction of wetland datasets through compiling historical data could make full use of existing information and represent the best choice when there is no other specific wetland data.

This kind of dataset can help cross-validate a new wetland map. However, it is challenging to reduce the uncertainties within each dataset that could be inherited by a wide variety of datasets, which have different application purposes, different wetland definitions, different mapping methods and different dates. Wetland modeling represents an efficient way to simulate wetlands, especially when physically based models can reflect the formation mechanisms of wetlands. One of the most important advantages of wetland models is that not only can these models trace back the historical wetland distribution, but they can also predict the future changes under different scenarios. Simulation wetland data provide a potential wetland distribution area. However, the major obstacles are that all models are simplifications of reality because of the complexity of wetlands and the human interference is not well considered in the models. Though these two kinds of approaches are not recommended to be used at current stage in the global wetland thematic mapping, they are helpful to understand global wetland distribution and to develop a new comprehensive global wetland dataset.

The RS classification is the most promising method for global wetland mapping in consideration of the cost of imagery processing. A variety of classification algorithms have been proposed in those successful global landcover mapping datasets (Friedl et al. 2002; Bartholomé and Belward 2005; Hansen et al. 2007; Prigent et al. 2007; Tateishi et al. 2011; Sadeghi et al. 2012; Blanco et al. 2013; Gong et al. 2016). Among these, the decision trees classification method has the ability to integrate a wide variety of input data, high flexibility and computerized efficiency, and has been widely used from regional/local scales to the global scale. This method is preferred, given the complexity of global wetlands, which are distributed across all climate zones from the tropics to the tundra, characterized by high hydro-dynamics, and encompass all vegetation types (herbaceous, shrub, and forest).

Input Data Sources

Compared with the classification algorithms, input data sources have a larger impact on the accuracy of classification, and the accuracy differences caused by the various input data are larger than those caused by different algorithms (Augusteijn and Warrender 1998; Li et al. 2013; Yu et al. 2014). Therefore, further efforts should be made to include new features for improving wetland thematic mapping accuracies (Li et al. 2013). Multi-source data, time-series data and auxiliary data have proved to be successful for wetland thematic mapping. For example, microwave data are less affected by weather and is sensitive to moisture. Hyperspectral data contains detailed spectral information. These multi-source data could make up the deficiencies of optical data. The combination of these approaches usually has a better result than the use of a single approach alone (Stankiewicz et al. 2003; Gumbricht 2012; Gong et al. 2016). However, because of the poor availability, high expense, relatively coarse spatial resolution and complicated process of these data, their combined usage at the global scale remains limited at present. At the same time, the dynamic characteristics of wetland make it difficult to identify the seasonal wetlands and detect the seasonal variation of wetlands with the single-date images. Time-series satellite imagery can provide more detailed information on wetland dynamics, which will help to improve the accuracy of wetlands classification. Other ancillary information such as topography, climate and soil data, which are closely related to wetland occurrence, are also useful in wetland mapping. For global wetland mapping initiatives, time-series multi-spectral imager, Synthetic Aperture Radar (SAR) could be more appropriate than other satellite data sources, such as high-resolution and hyperspectral imagery.

Sample Dataset

The sample dataset is very important for the automatic classification of RS imagery because the accuracy of samples can directly influence the accuracy of the classified result. For example, the supervised classification is based on the training sample and the reliability of the classification is verified by the validation sample. However, the selection of samples is a labor-intensive and time-consuming process, especially when developing a global sample dataset, which mainly relies on human experience through field investigation and high-resolution imagery.

The development of a tool of sample sharing and make full use of existing results is a practical way to solve this problem (Klemas 2005; Li et al. 2013; Zhao et al. 2014). The emerging big-data technology, which can help search for and analyze massive wetland data that are stored on the Internet and in all formats, is promising to help construct a global wetland sample dataset. Future research is also needed to develop discriminative method of precise wetland samples. The transformation of samples from different wetland classifications systems is also to be solved. Moreover, in consideration of the dynamic variation of wetland and the changes in land cover, a sample dataset comprising the seasonal dynamics of wetland would be useful for mapping wetland dynamics (Li et al. 2013). Developing a universal sample dataset with spatial and temporal representation is challenge, but it is essential for global wetland mapping.

Conclusions

Accurate wetland datasets are indispensable in generating policies on wetland conservation and their appropriate land use, global climate change studies, and biodiversity conservation. According to the development methods of global datasets, the global wetland-related datasets can be divided into three groups: remote-sensed datasets from satellite imagery through RS classification, compilation datasets by aggregating various independent datasets, and simulation datasets by employing hydrological models or LSMs. However, because of the diversity of wetland definition and intrinsic features of wetland ecosystem, there are major inconsistencies between the existing global wetland-related datasets in terms of wetland area and spatial distribution.

Given the various requirements of wetland-oriented data products, it is important to conduct comprehensive global wetland mapping. Because of the broader usages of wetland datasets, the Ramsar wetland definition is recommended to be adopted in the development of comprehensive global wetland dataset, since most wetland categories in existing global wetland-related datasets are included within it. At the same time, a hierarchical and flexible structure of global wetland classification system is preferred, in which (besides the traditional information such as location, size, and distribution), more information on biogeochemical features of wetlands can be considered.

Considering the limitations and strengths of current wetland-related data products at the regional and global scale, we suggest that the synthetic method of wetland mapping should be applied. For instance, the simulation of wetland datasets can provide the maximum boundary of wetland distribution, whereas a compilation wetland dataset could help cross-validation of the resultant wetland map. Among the computer automatic classification methods, the classification tree method is promising because of its flexibility as it can deal with various conditions of wetland across the world and adopt various data sources and approaches. To address the dynamics of wetlands at global scale, time-series satellite imagery with 250–1000 m spatial resolution is preferred. In addition, in order to avoid the optical imagery limitations of cloud cover and forested wetland canopy cover, the time-series satellite imagery could be combined with passive/active SAR data. Though sample selection is labor-intensive and time-consuming, it is essential for wetland mapping. Big-data technology is promising for the development of global wetland sample datasets.

Acknowledgements This research was supported by National Natural Science Fund of China (41271423). We also greatly appreciated the professional comments and advice of the anonymous reviewers.

References

Aselmann I, Crutzen PJ (1989) Global distribution of natural freshwater wetlands and rice paddies, their net primary productivity, seasonality and possible methane emissions. Journal of Atmospheric Chemistry 8:307–358

Augusteijn MF, Warrender CE (1998) Wetland classification using optical and radar data and neural network classification. International Journal of Remote Sensing 19:1545–1560

Avis C, Weaver AJ, Meissner KJ (2011) Reduction in areal extent of high-latitude wetlands in response to permafrost thaw. Nature Geoscience 4:444–448

Baker C, Lawrence R, Montagne C et al (2006) Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models. Wetlands 26:465–474

Barbier EB, Acreman MC, Knowler D. Economicvaluation of wetlands: a guide for policy makers and planners. Ramsar Convention Bureau, Gland, 1997

Bartholomé E, Belward AS (2005) GLC2000: a new approach to global land cover mapping from earth observation data. International Journal of Remote Sensing 26:1959–1977

Bicheron P, Leroy M, Brockmann C et al (2006) Globcover: a 300 m global land cover product for 2005 using ENVISAT MERIS time series. Proceedings of the Recent Advances in Quantitative Remote Sensing Symposium, Valencia, pp 538–542

Blanco PD, Colditz RR, Saldaña GL et al (2013) A land cover map of Latin America and the Caribbean in the framework of the SERENA project. Remote Sensing of Environment 132:13–31

Boer G, Flato G, Ramsden D (2000) A transient climate change simulation with greenhouse gas and aerosol forcing: projected climate to the twenty-first century. Climate Dynamics 16:427–450

Browne DM, Dell R (eds) (2007) Conserving waterfowl and wetlands amid climate change. Ducks Unlimited, Inc., Memphis

Chen J, Chen J, Liao A et al (2015) Global land cover mapping at 30m resolution: a POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing 103:7–27

Costanza R, D'Arge R, Groot RD et al (1997) The value of the world's ecosystem services and natural capital. Ecological Economics 387: 253–260

Cowardin LM, Carter V, Golet FC, et al. 1989. Classification of wetlands and deepwater habitats of the United States. U. S. Department of the Interior, fish and Wildlife Service, Washington, D.C. Jamestown: northern prairie Wildlife research Center home page. http://www.Npwrc.Usgs.Gov/resource/1998/classwet/classwet.Htm (Version 04DEC98)

Dahl TE (1990) Wetlands losses in the United States, 1780's to 1980's. Report to the Congress, National Wetlands Inventory, St. Petersburg

Davidson NC (2014) How much wetland has the world lost? Long-term and recent trends in global wetland area. Marine and Freshwater Researh 65:936–941

Dixon MJR, Loh J, Davidson NC et al (2016) Tracking global change in ecosystem area: the wetland extent trends index. Biological Conservation 193:27–35

Donchyts G, Baart F, Winsemius H et al (2016) Earth's surface water change over the past 30 years. Nature Climate Change 6:810–813

Erwin KL (2009) Wetlands and global climate change: the role of wetland restoration in a changing world. Wetlands Ecology and Management 17:71–84

Fan Y, Miguez-Macho G (2010) A simple hydrologic framework for simulating wetlands in climate and earth system models. Climate Dynamics 37:253–278

Fan Y, Li H, Miguezmacho G (2013) Global patterns of groundwater table depth. Science 339:940–943

Finlayson CM, van der Valk AG (1995) Wetlands classification and inventory: a summary. Vegetatio 118:185–192

Finlayson CM, Davidson NC, Spiers AG et al (1999) Global wetland inventory – current status and future priorities. Marine and Freshwater Research 50:717–727

Finlayson M, Cruz RD, Davidson N et al (2005) Millennium ecosystem assessment: ecosystems and human well-being: wetlands and water synthesis. World Resources Institute, Washington, DC 656:87–98

Fluet-Chouinard E, Lehner B, Rebelo L-M et al (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

Friedl MA, McIver DK, Hodges JC et al (2002) Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83:287–302

Giri C, Zhu Z, Reed B (2005) A comparative analysis of the global land cover 2000 and MODIS land cover data sets. Remote Sensing of Environment 94:123–132

Gong P, Wang J, Yu L et al (2012) Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. International Journal of Remote Sensing 34:2607–2654

Gong P, Yu L, Li C et al (2016) A new research paradigm for global land cover mapping. Annual GIS:1–16

Gumbricht T. Mapping global tropical wetlands from earth observing satellite imagery. Center for International Forestry Research (CIFOR), Bogor, 2012

Hansen M, Dubayah R, Defries R (2007) Classification trees: an alternative to traditional land cover classifiers. International Journal of Remote Sensing 17:1075–1081

Herold M, Mayaux P, Woodcock CE et al (2008) Some challenges in global land cover mapping: an assessment of agreement and accuracy in existing 1 km datasets. Remote Sensing of Environment 112: 2538–2556

Hopcroft PO, Valdes PJ, Beerling DJ (2011) Simulating idealized Dansgaard-Oeschger events and their potential impacts on the global methane cycle. Quaternary Science Revies 30:3258–3268

Hu SJ, Niu ZG, Chen YF et al (2017) Global wetlands: potential distribution, wetland loss, and status. Science of the Total Environment. doi:10.1016/j.scitotenv.2017.02.001

Hulme M, Doherty R, Ngara T et al (2005) Global warming and African climate change: a reassessment. Climate Change:29–40

Jung M, Henkel K, Herold M et al (2006) Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment 101:534–553

Kent, Donald M. Defining wetlands. In: Kent, Donald M, et al. Applied wetlands science and technology. Lewis Publishers, Boca Raton, 1996

Kleinen T, Brovkin V (2013) Modelling global peatlands from LGM to Anthropocene[C]// AGU fall meeting abstracts

Klemas V (2005) Remote sensing Techniques for studying coastal ecosystems: an overview. Journal of Coastal Research 27(1):2–17

Klemov, Kenneth M (1998) Wetland mapping. In: Majumdar SK, Miller EW et al (eds) Ecology of wetlands and associated systems. The Pennsylvania Academy Science

Köchy M, Freibauer A (2009) Global distribution of wetlands. Coordination action carbon observation system. http://literatur.Thuenen.De/digbib_extern/dk043226.Pdf, 14 Sept 2016, data last accessed

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

Li C, Wang J, Wang L et al (2013) Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. Remote Sensing 6:964–983

Loveland T, Reed B, Brown J et al (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing 21:1303–1330

Matthews E, Fung I (1987) Methane emission from natural wetlands: global distribution, area, and environmental characteristics of sources. Global Biogeochemical Cycles 1:61–86

Melton JR, Wania R, Hodson EL et al (2012) Present state of global wetland extent and wetland methane modelling: methodology of a model intercomparison project (WETCHIMP). Geoscientific Model Development Discussion 5:4071–4136

Min F, Joseph OS, Saurabh C et al (2016) 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 9:113–133

Mitra S, Wassmann R, Paul LG. Global Inventory of Wetlands and their Role in the Carbon Cycle. ZEF – Discussion Papers On Development Policy No. 64. Center for Development Research, Bonn, 2003, pp. 44

Mitra S, Wassmann R, Vlek PLG (2005) An appraisal of global wetland area and its organic carbon stock. Current Science 88:25–35

Mitsch WJ, Gosselink JG (1993) Wetlands, Second edn. Van Nostrand Reinhold, New York

Moody A, Strahler A (1993) MODIS land cover product algorithm technical basis document. NASA Mission to Planet Earth Program, NASA, Greenbelt

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

Navid D (2014) International law of migratory species: the Ramsar Convention. Doklady Earth Sciences 455:296–298

Nayak SR, Sahai B (1985) Coastal morphology -a case-study of the Gulf of Khambhat (Cambay). International Journal of Remote Sensing 6: 559–567

Niu ZG (2015) Mapping China's wetlands and recent changes with remotely sensed data. In: Tiner RW, Lang MW, Klemas VV (eds) Remote sensing of wetlands: applications and advances. CRC Press, Boca Raton, pp 473–490

Ozesmi SL, Bauer ME (2002) Satellite remote sensing of wetlands. Wetlands Ecology and Management 10:381–402

Parish F, Looi C (1999) Wetlands, biodiversity and climate change. Options and needs for enhanced linkage between the Ramsar Convention on wetlands, Convention on Biological Diversity and UN Framework Convention on Climate Change. Global Environment Network

Paul AK (2000) Wetland Ecology: principles and conservation, 2nd edn. Cambridge University Press, Cambridge

Pekel JF, Cottam A, Gorelick N et al (2016) High-resolution mapping of global surface water and its long-term changes. Nature 540:418–422

Phinn S, Hess L, Finlayson CM (1999) An assessment of the usefulness of remote sensing for wetland inventory and monitoring in Australia. In: Finlayson CM, Speirs AG (eds) Techniques for enhanced wetland inventory and modelling, supervising scientist report. Supervising Scientist, Canberra, pp 44–81

Prigent C, Papa F, Aires F et al (2007) Global inundation dynamics inferred from multiple satellite observations, 1993–2000. Journal of Geophysical Research-Atmospheres 112:1103–1118

Ramsar Information Bureau (1998) Report of allocations 1992–97, the Ramsar Convention small grants fund. Ramsar, Gland, http://www.iucn.org/themes/Ramsar

Riley WJ, Subin ZM, Lawrence DM et al (2011) Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. Biogeosciences 8:1925–1953

Russi D, Brink PT, Badura T et al (2013) The economics of ecosystems and biodiversity for water and wetlands. IEEP, London

Sadeghi R, Zarkami R, Sabetraftar K et al (2012) Use of support vector machines (SVMs) to predict distribution of an invasive water fern *Azolla filiculoides* (lam.) in Anzali wetland, southern Caspian Sea, Iran. Ecological Modelling 244:117–126

Sahagian D, Melack J (1997) Global wetland distribution and functional characterization: trace gases and the hydrologic cycle. Environmental Policy Collection

Schowengerdt RA (2012) Techniques for image processing and classifications in remote sensing. Academic, New York

Scott DA (2010) A directory of wetlands. The Middle East Iucn, Gland, Switzerland and International Waterfowl Research Bureau, Slimbridge, United Kingdom

Sharma R, Tateishi R, Hara K et al (2015) Developing superfine water index (SWI) for global water cover mapping using MODIS data. Remote Sensing 7:13807–13841

Sophie B, Pierre D (2010) GlobCover 2009 products description and validation report. European Space Agency, Paris

Spalding MD, Blasco F, Field CD (1997) World mangrove atlas. The International Society for Mangrove Ecosystems, Okinawa

Stankiewicz K, DabrowskaZielinska K, Gruszczynska M et al (2003) Mapping vegetation of a wetland ecosystem by fuzzy classification of optical and microwave satellite images supported by various ancillary data. Proceedings of SPIE - The International Society for Optical Engineering 4879:352–361

Takai Y (1970) The mechanism of methane formation in flooded paddy soil. Soil Science & Plant Nutrition 16:238–244

Tateishi R, Uriyangqai B, Al-Bilbisi H et al (2011) Production of global land cover data – GLCNMO. International Journal of Digital Earth 4:22–49

Tateishi R, Hoan NT, Kobayashi T et al (2014) Production of global land cover data - GLCNMO2008. Journal of Geography and Geology 6:99

Tuanmu MN, Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Global Ecology and Biogeography 23:1031–1045

Van Engelen V, Huting J (2002) Peatlands of the world. An interpretation of the world soil map. ISRIC, Wageningen, unpublished. GPI project 29 GPI 1

Wright C, Gallant A (2007) Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. Remote Sensing of Environment 107:582–605

Xu X, Tian H (2012) Methane exchange between marshland and the atmosphere over China during 1949–2008. Global Biogeochemical Cycles 26:1–14

Yin SB, Bing LI, Fang S (2014) Review on studies of wetland definition. Wetland Science 12(4):504–514

Yu L, Liang L, Wang J et al (2014) Meta-discoveries from a synthesis of satellite-based land-cover mapping research. International Jouranl of Remote Sensing 35:4573–4588

Zhao Y, Gong P, Yu L et al (2014) Towards a common validation sample set for global land-cover mapping. International Jouranl of Remote Sensing. 35:4795–4814

Zhu P, Gong P (2014) Suitability mapping of global wetland areas and validation with remotely sensed data. Science China Earth Sciences 57(10):2283–2292