



**HAL**  
open science

# Satellite remote sensing of the variability of the continental hydrology cycle in the lower Mekong basin over the last two decades

Binh Pham-Duc

► **To cite this version:**

Binh Pham-Duc. Satellite remote sensing of the variability of the continental hydrology cycle in the lower Mekong basin over the last two decades. Hydrology. Sorbonne Université, 2018. English. NNT : 2018SORUS024 . tel-02109003

**HAL Id: tel-02109003**

**<https://theses.hal.science/tel-02109003v1>**

Submitted on 24 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# SORBONNE UNIVERSITÉ

ÉCOLE DOCTORALE DES SCIENCES DE L'ENVIRONNEMENT D'ÎLE DE FRANCE

---

## SATELLITE REMOTE SENSING OF THE VARIABILITY OF THE CONTINENTAL HYDROLOGY CYCLE IN THE LOWER MEKONG BASIN OVER THE LAST TWO DECADES

---

*PRÉSENTÉE PAR*

**Binh PHAM-DUC**

*THÈSE DE DOCTORAT DE SCIENCES DE L'ENVIRONNEMENT*

DIRECTRICE DE RECHERCHE: Catherine PRIGENT

CO-DIRECTEUR DE RECHERCHE: Filipe AIRES

PRÉSENTÉE ET SOUTENUE PUBLIQUEMENT LE 6 FÉVRIER, 2018

DEVANT UN JURY COMPOSÉ DE:

Agnès DUCHARNE	METIS/CNRS-UPMC	Présidente
Stéphane JACQUEMOUD	U. PARIS DIDEROT	Rapporteur
Thanh NGO-DUC	USTH/VAST	Rapporteur
Sylvain BIANCAMARIA	LEGOS/CNRS	Examineur
Selma CHERCHALI	CNES	Examinatrice
Catherine PRIGENT	LERMA/CNRS	Directrice de thèse
Filipe AIRES	LERMA/CNRS	Co-Directeur de thèse

THÈSE PRÉPARÉE À L'OBSERVATOIRE DE PARIS, AU LERMA

---

---

# Résumé

Les eaux superficielles sont nécessaires à toute forme de vie en tant que parties intégrantes de tout processus de vie sur Terre. Quantifier les eaux de surface et suivre leurs variations est primordiale en raison du lien direct qui existe entre les variables hydrologiques et le changement climatique. La télédétection par satellite, de l'hydrologie continentale offre l'opportunité unique d'étudier, depuis l'espace, les processus hydrologiques à différentes échelles (régionale et globale). Dans cette thèse, différentes techniques ont été développées afin d'étudier les variations des eaux superficielles ainsi que d'autres variables hydrologiques, au niveau du bassin inférieur du Mékong (entre le Vietnam et le Cambodge) et ce en utilisant plusieurs estimations satellitaires différentes. Cette thèse s'articule autour de quatre points principaux.

Premièrement, l'utilisation d'observations satellitaires dans le visible et dans l'infrarouge (MODIS) est étudiée et comparée afin d'évaluer les eaux de surface au niveau du bassin inférieur du Mékong. Quatre méthodes de classification ont été utilisées afin de différencier les types de surface (inondés ou pas) dans le bassin. Les différentes méthodes ont donné des cartes d'eaux de surface aux résultats semblables en terme de dynamique saisonnière. La classification la plus adaptée aux régions tropicales a été ensuite choisie pour produire une carte des eaux de surface à la résolution de 500 m entre janvier 2001 et aujourd'hui. La comparaison des série temporelles issues de cette carte et de celles issues du produit de référence MODIS donne une forte corrélation temporelle (> 95%) pour la période 2001-2007.

Deuxièmement, l'utilisation des observations issues du satellite SAR Sentinel-1 est examinée à des fins identiques. L'imagerie satellitaire optique est ici remplacée

---

par les images SAR qui grâce aux longueurs d'ondes utilisées dans le micro-onde, permettent de « voir » à travers les nuages. Un jeu d'images Landsat-8-sans-nuage est alors utilisé pour entraîner un Réseau de Neurones (RN) afin de restituer des cartes d'eaux de surface par l'utilisation d'un seuillage sur les sorties du modèle RN. Les cartes sont à la résolution spatiale de 30 m et disponibles depuis janvier 2015. Comparées aux cartes de référence Landsat-8-sans-nuage, les sorties de modèles RN montrent une très grande corrélation (90%) ainsi qu'une détection "vraie" à 90%. Les cartes restituées d'eaux de surface utilisant la technologie SAR sont enfin comparées aux cartes d'inondation issues de données topographiques. Les résultats montrent une fois encore une très grande consistance entre les deux cartes avec 98% des pixels considérés comme inondés dans les cartes SAR se trouvant dans les régions de très grande probabilité d'inondation selon la topographie (>60%).

Troisièmement, la variation volumique des eaux de surface est calculée comme le produit de l'étendue de la surface avec la hauteur d'eau. Ces deux variables sont validées à l'aide d'autres produits hydrologiques et montrent de bons résultats. La hauteur d'eau superficielle est linéairement interpolée aux régions non inondées afin de produire des cartes mensuelles à la résolution spatiale de 500 m. La hauteur d'eau est ensuite analysée pour estimer les variations volumiques. Ces résultats montrent une très bonne corrélation avec la variation volumique induite par la mesure du contenu en eau du satellite GRACE (95%) ainsi qu'avec la variation des mesures in situ de débit des rivières.

Finalement, deux produits globaux et multi-satellites d'eaux superficielles sont comparés à l'échelle régionale et globale sur la période 1993-2007: GIEMS et SWAMPS. Lorsqu'elles existent, les données auxiliaires sont utilisées afin de renforcer l'analyse. Les deux produits montrent une dynamique similaire, mais 50% des pixels inondés dans SWAMPS se trouvent le long des côtes. Pour les bassins de l'Amazonie et de l'Orénoque, GIEMS and SWAMPS montrent de très grandes corrélations (respectivement 95% et 99%), mais le maximum d'étendue d'eau de surface est moitié moins dans SWAMPS que dans GIEMS et/ou SAR. De plus SWAMPS ne prend pas compte de la dynamique des eaux de surface pour le bassin du Niger et son estimation de la saison s'y trouve déphasée aussi bien avec les estimations issues de GIEMS et MODIS qu'avec les données in situ de débit du Niger.

---

**Mots clés:** Télédétection par satellite, surveillance des eaux de surface, hauteur des eaux de surface, volume d'eau de surface, précipitations, MODIS, Landsat, Sentinel-1, ENVISAT, Delta du Mékong



# Abstract

Surface water is essential for all forms of life since it is involved in almost all processes of life on Earth. Quantifying and monitoring surface water and its variations are important because of the strong connections between surface water, other hydrological components (groundwater and soil moisture, for example), and the changing climate system. Satellite remote sensing of land surface hydrology has shown great potential in studying hydrology from space at regional and global scales. In this thesis, different techniques using several types of satellite estimates have been made to study the variation of surface water, as well as other hydrological components in the lower Mekong basin (located in Vietnam and Cambodia) over the last two decades. This thesis focuses on four aspects.

First, the use of visible/infrared MODIS/Terra satellite observations to monitor surface water in the lower Mekong basin is investigated. Four different classification methods are applied, and their results of surface water maps show similar seasonality and dynamics. The most suitable classification method, that is specially designed for tropical regions, is chosen to produce regular surface water maps of the region at 500 m spatial resolution, from January 2001 to present time. Compared to reference data, the MODIS-derived surface water time series show the same amplitude, and very high temporal correlation for the 2001-2007 period ( $> 95\%$ ).

Second, the use of SAR Sentinel-1 satellite observations for the same objective is studied. Optical satellite data are replaced by SAR satellite data to benefit the ability of their microwave wavelengths to pass through clouds. Free-cloud Landsat-8 satellite imagery are set as targets to train and optimize a Neural Network (NN).



---

Predicted surface water maps (30 m spatial resolution) are built for the studied region from January 2015 to present time, by applying a threshold (0.85) to the output of the NN. Compared to reference free-cloud Landsat-8 surface water maps, results derived from the NN show high spatial correlation ( $\sim 90\%$ ), as well as true positive detection of water pixels ( $\sim 90\%$ ). Predicted SAR surface water maps are also compared to floodability maps derived from topography data, and results show high consistency between the two independent maps with 98% of SAR-derived water pixels located in areas with a high probability of inundation ( $>60\%$ ).

Third, the surface water volume variation is calculated as the product of the surface water extent and the surface water height. The two components are validated with other hydrological products, and results show good consistencies. The surface water height are linearly interpolated over inundated areas to build monthly maps at 500 m spatial resolution, then are used to calculate changes in the surface water volume. Results show high correlations when compared to variation of the total land surface water volume derived from GRACE data (95%), and variation of the in situ discharge estimates (96%).

Fourth, two monthly global multi-satellite surface water products (GIEMS & SWAMPS) are compared together over the 1993-2007 period at regional and global scales. Ancillary data are used to support the analyses when available. Similar temporal dynamics of global surface water are observed when compared GIEMS and SWAMPS, but  $\sim 50\%$  of the SWAMPS inundated surfaces are located along the coast line. Over the Amazon and Orinoco basins, GIEMS and SWAMPS have very high water surface time series correlations (95% and 99%, respectively), but SWAMPS maximum water extent is just a half of what observed from GIEMS and SAR estimates. SWAMPS fails to capture surface water dynamics over the Niger basin since its surface water seasonality is out of phase with both GIEMS- and MODIS-derived water extent estimates, as well as with in situ river discharge data.

**Keywords:** Satellite remote sensing, surface water monitoring, surface water height, surface water volume, discharge, precipitation, MODIS, Landsat, Sentinel-1, ENVISAT, Mekong Delta.

# Publications

- **Binh Pham-Duc**, Catherine Prigent, Filipe Aires, Fabrice Papa, *Comparisons of global terrestrial surface water datasets over 15 years*. Journal of Hydrometeorology 04/2017. DOI:10.1175/JHM-D-16-0206.1
- Filipe Aires, Leo Miolane, Catherine Prigent, **Binh Pham-Duc**, Etienne Fluet-Chouinard, Bernhard Lehner, Fabrice Papa, *A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations*. Journal of Hydrometeorology 05/2017. DOI:10.1175/JHM-D-16-0155.1
- **Binh Pham-Duc**, Catherine Prigent, Filipe Aires, *Surface water monitoring within Cambodia and the Vietnamese Mekong Delta over a year, with Sentinel-1 SAR observations*. Water 05/2017. DOI:10.3390/w9060366



# Acknowledgment

The last three years have been a time of tremendous learning and growth for me professionally and personally. This thesis would not have been possible without the support, guidance and motivation from my supervisors, colleagues, thesis committee, friends, and family.

I owe my deepest gratitude to my supervisor, Dr. Catherine Prigent, for the continuous support of my PhD study and related researches, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. As well, I would like to express my sincere thanks and immense pleasure to my co-supervisor, Dr. Filipe Aires, for his scientific guidance, consistent support, and encouragement throughout the duration of my PhD.

I would like to thank to my colleagues whom I have been working with over the years, especially, Frédéric Frappart, Fabrice Papa, Victor Pellet, Xiaoni Wang, Jordane Mathieu, Lise Kilic, Die Wang, Evan Eames, and Duy Ba Nguyen. Their comments, corrections and suggestions contributed greatly to my work. It has been also a great pleasure working with them.

I am deeply grateful to my thesis committee: Prof. Stéphane Jacquemoud and Assoc. Prof. Thanh Ngo Duc, for their insightful comments and encouragement. The quality of my thesis has been largely improved thank to their interest and valuable scientific questions and suggestions.

To all my friends, thank you for wonderful time we shared that made my years in Paris become unforgettable.

---

To my family, I cannot express my gratitude for always supporting me for all the years. To my parents and in-laws, thank you for your unconditional love that inspire me to follow my dreams. To little my brother, thank you for being there taking care of our parents when I was not available. I owe my special thanks to my beloved wife and little daughter for their continued and unfailing love, understanding, support and kindness. A few words cannot express how grateful I am to my family, but this thesis would not have been possible without their support.

This thesis is dedicated to my wife and my little daughter.

**Binh PHAM-DUC**

# Contents

<b>Abstract</b>	<b>v</b>
<b>Publications</b>	<b>vi</b>
<b>Acknowledgment</b>	<b>ix</b>
<b>Contents</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Land surface remote sensing from satellites . . . . .	3
1.1.1 The need for satellite observations . . . . .	3
1.1.2 The atmospheric transmission windows . . . . .	4
1.2 Hydrological cycle from space . . . . .	6
1.2.1 Estimation of continental surface water extent from space . . . . .	7
1.2.1.1 Visible and Near-Infrared (NIR) observations . . . . .	7
1.2.1.2 Active microwave observations . . . . .	8
1.2.1.3 Passive microwave observations . . . . .	9
1.2.2 Estimation of surface water height from space . . . . .	9
1.2.3 Estimation of other hydrological components from space . . . . .	10
1.3 The study area: The Vietnam Mekong Delta and the Tonle Sap Lake in Cambodia . . . . .	11
1.3.1 Presentation of the region . . . . .	11
1.3.2 The satellite remote sensing studies so far over the region . . . . .	14
1.4 Thesis objectives . . . . .	14
<b>2 Surface Water Monitoring within the Mekong Delta and Cambodia using Visible and Infrared MODIS Satellite Observations</b>	<b>17</b>
2.1 Introduction . . . . .	19
2.2 Different methodologies using MODIS/Terra observations for surface water monitoring . . . . .	23
2.2.1 MODIS/Terra observations . . . . .	23
2.2.2 Sakamoto et al. [2007] methodology . . . . .	25
2.2.3 Bergé-Nguyen and Crétaux [2015] methodology . . . . .	27

---

2.2.4	Unsupervised and supervised classification techniques . . . .	28
2.2.4.1	Unsupervised classification . . . . .	28
2.2.4.2	Supervised classification . . . . .	30
2.2.4.3	Hybrid classification . . . . .	31
2.3	MOD09A1 pre-processing . . . . .	32
2.3.1	MOD09A1 quality and reliability checking . . . . .	32
2.3.2	Smoothing method selection . . . . .	35
2.4	Results . . . . .	36
2.4.1	Comparing Sakamoto et al. [2007] and Bergé-Nguyen and Crétau [2015] methodologies . . . . .	36
2.4.2	Comparing unsupervised and supervised methodologies . . . . .	38
2.4.3	Expanding Sakamoto methodology . . . . .	41
2.5	Discussions and conclusions . . . . .	42
<b>3</b>	<b>Surface Water Monitoring within the Mekong Delta and Cambodia using SAR Sentinel-1 Satellite Observations</b>	<b>45</b>
3.1	Introduction . . . . .	47
3.2	Sentinel-1 SAR data and the ancillary datasets . . . . .	50
3.2.1	Sentinel-1 SAR data . . . . .	50
3.2.2	Ancillary datasets . . . . .	52
3.2.2.1	Inundation maps derived from Landsat-8 data . . . . .	52
3.2.2.2	Inundation maps derived from MODIS/Terra data . . . . .	53
3.3	Methodology . . . . .	55
3.3.1	Surface water information from the Sentinel-1 SAR images . . . . .	55
3.3.2	A Neural Network-based classification . . . . .	57
3.3.3	NN sensitivity tests . . . . .	60
3.3.3.1	Selection of an optimized threshold for the NN output . . . . .	61
3.3.3.2	Equalization of water and non-water pixel number . . . . .	62
3.3.3.3	Analyzing the weight of each NN satellite input . . . . .	63
3.4	Results and comparisons with other surface water products . . . . .	66
3.4.1	Evaluation of the SAR NN classification method with Landsat-8 images . . . . .	67
3.4.2	Evaluation using a topography-based floodability index . . . . .	70
3.4.3	Comparisons with MODIS/Terra-derived inundation maps . . . . .	72
3.5	Improvement of the Neural Network . . . . .	74
3.5.1	Performance comparisons between NN5 and NN7 . . . . .	76
3.6	Conclusions and perspectives . . . . .	79
<b>4</b>	<b>Toward the analyses of the change in surface water volume within the lower Mekong Delta</b>	<b>83</b>
4.1	Introduction . . . . .	85

---

4.2	Datasets	86
4.2.1	Satellite altimetry data	86
4.2.2	GRACE data	87
4.2.3	In situ water level and river discharge data	88
4.2.4	Precipitation data	89
4.2.5	The El Niño-Southern Oscillation index	89
4.3	Evaluations of the MODIS-derived surface water extent	90
4.3.1	Comparison with SAR-derived surface water extent	91
4.3.2	Comparison with precipitation data	93
4.3.3	Comparison with in situ water level data	94
4.3.4	Comparison with the Multivariate ENSO Index	95
4.4	Surface water height from altimetry data	96
4.4.1	The principle of radar satellite altimeter	96
4.4.2	Satellite-based surface water height estimations along the Mekong River	98
4.4.3	Evaluations of satellite-based surface water height data	100
4.4.3.1	Comparison with other studies	100
4.4.3.2	Comparison with in situ water level data	102
4.4.4	Interpolated surface water height maps	104
4.5	Surface water volume variations	106
4.5.1	Comparison between MODIS-derived surface water extent and satellite-based surface water height	106
4.5.2	Methodology	107
4.5.3	Results	107
4.6	Discussions and conclusions	108
<b>5</b>	<b>Comparisons between Global Terrestrial Surface Water Datasets</b>	<b>111</b>
5.1	Introduction	113
5.2	Data	115
5.2.1	GIEMS	115
5.2.2	SWAMPS	117
5.2.3	GLWD	117
5.2.4	Other ancillary datasets	118
5.3	Comparisons of the surface water datasets	119
5.3.1	Global comparisons	119
5.3.2	Basin-scale comparisons	124
5.3.2.1	Comparisons over the lower Mekong basin	124
5.3.2.2	Comparisons over the Amazon basin	126
5.3.2.3	Comparisons over the Orinoco basin	128
5.3.2.4	Comparisons over the Niger basin	130
5.3.2.5	Comparisons over the Ganges basin	134
5.3.2.6	Comparisons over the Ob basin	135



---

5.3.2.7	Comparisons over the Mississippi basin . . . . .	137
5.4	Discussions . . . . .	139
5.5	Conclusions . . . . .	141
<b>6</b>	<b>Conclusions and Perspectives</b>	<b>143</b>
6.1	Conclusions . . . . .	145
6.1.1	Surface water monitoring with visible/infrared MODIS satellite observations . . . . .	145
6.1.2	Surface water monitoring with SAR Sentinel-1 satellite observations . . . . .	146
6.1.3	Toward the variation of the surface water volume . . . . .	147
6.1.4	Comparisons between two global terrestrial surface water datasets . . . . .	148
6.1.5	Other activities during my PhD . . . . .	149
6.1.5.1	Participation to another paper . . . . .	149
6.1.5.2	Teaching experience . . . . .	149
6.2	Perspectives . . . . .	150
6.2.1	Surface water monitoring over the Vietnam Red River Delta . . . . .	150
6.2.2	Closure of the water cycle over the Mekong Delta and Cambodia . . . . .	153
6.2.3	The future with the SWOT and JV-LOTUSat satellites . . . . .	154
<b>A Paper A:</b>		<b>157</b>
<b>B Paper B:</b>		<b>181</b>
<b>Bibliography</b>		<b>199</b>

# **Chapter 1**

## **Introduction**



## 1.1 Land surface remote sensing from satellites

### 1.1.1 The need for satellite observations

Earth remote sensing, by definition, is the science of acquiring information about the Earth without actually being in contact with it. This is done by sensing and recording reflected or emitted energy and processing, analyzing, and applying that information [CCMEO, 2013]. Satellite remote sensing started in 1960s after Sputnik-1, the world first man-made satellite, that was launched into space by the Soviet Union on 4 October 1957. Since then, satellite technology has developed quickly, especially over the last few decades, providing valuable knowledge for numerous fields of Earth sciences (for example, geography, oceanography, glaciology, forestry, agriculture, and hydrology).

Applications of satellite remote sensing focusing on land surface hydrology began with the successful launch of the first Earth Resources Technology Satellite (ERTS-1, later renamed as Landsat-1) on 23 July 1972. It opened a new era for hydrological studies since scientists had a new source of valuable satellite observations useful for researches of hydrological dynamics and processes [Pietroniro and Prowse, 2002]. Compared to hydrological data provided by in situ measurements, satellite remote sensing observations have obvious advantages. First, in situ measurements can provide longer data records, however, they contain limited information about the spatial dynamics of hydrological parameters (surface water, for instance). In contrast, satellite observations can provide not only regional, but also global observations at different wavelengths and spatial resolutions, with uniform quality and rapid data acquisition. Second, the distribution of the stream gauge network is not uniform, very dense in developed countries (in the US, for example), but very sparse in developing countries (especially in African countries), or even not available in remote areas [Alsdorf et al., 2007]. Satellite observations can access to remote or inaccessible areas, and provide regular observations all over the globe. Third, in situ measurements are often unavailable for scientific purposes due to geopolitical reasons, but satellites can acquire information over countries that are not willing to share data (that is very important for military applications). For these reasons, nowadays, in situ data

are normally used for calibration and validation of methodologies using land surface satellite observations. Despite all these advantages, satellite remote sensing has limitations. A satellite system is very expensive, and it takes a long and complicated process to develop, build, test, launch and operate the system. As a consequence, many countries have their own satellites with the ability to provide observations, but only a few countries can master satellite technology (the US, Russia, Japan, France, German, Italy, India, etc). Different satellites provide different spatial resolutions, and the applications can be affected due to large uncertainties and errors in the measurements. It is also difficult for data interpretation, especially with radar images.

In land surface satellite remote sensing, there are two types of sensors: passive and active. Depending on the observed surfaces, the interactions between the surfaces and passive/active sensors will be different. Passive sensors measure sunlight radiation reflected from the Earth surfaces, or thermal radiation naturally emitted from objects. Active sensors create their own electromagnetic energy that is transmitted from the sensor toward the target to be investigated. The incoming energy interacts with the target producing a backscattered energy that is reflected back to the sensor for measurements. Observations acquired by measuring surface reflectance from the Sun is known as optical observations. Therefore, optical imagery relies on the Sun, and can be subjected to the presence of clouds. Optical observations are often used for applications related to surface water monitoring, disaster monitoring, inundation mapping, or land surface change detection [Owe et al., 2001]. Active sensors emit radar beams that are not blocked by clouds, and radar imagery can be acquired at any time, under all weather conditions and are independent from the Sun [Owe et al., 2001].

### **1.1.2 The atmospheric transmission windows**

The presence of different types of molecules in the Earth's atmosphere puts limitations on the spectral regions that can be used in land surface remote sensing. Each molecule has its absorption bands in various parts of the electromagnetic spectrum. As a consequence, only wavelengths outside the main atmospheric absorption bands can be used in remote sensing of the Earth surface, and these

wavelength regions are known as the atmospheric transmission windows. Some wavelengths easily pass through the atmosphere, other wavelengths are blocked or absorbed partly to totally by the atmosphere. Figure 1.1 shows details of the atmospheric transmission windows from radio to X-ray wavelengths. High energy wavelengths (Ultraviolet, X-rays, and Gamma-rays) are absorbed by the ozone in the Earth's upper atmosphere. The visible wavelengths are not blocked by the Earth's atmosphere, but they can be scattered by dust and clouds. In the infrared ranges, some wavelengths are blocked by the atmosphere, but others can pass. Similarly in the microwave ranges, the Earth's atmosphere is transparent at some wavelengths, but not at others. Finally, in the radio ranges, the Earth's atmosphere is totally transparent to most of its wavelengths. Based on these atmospheric transmission windows, satellite remote sensing instruments are designed to operate in one or more windows where wavelengths can pass through the Earth's atmosphere to observe the Earth surface. More details of the atmospheric transmission windows can be found in remote sensing books, for example [Elachi and van Zyl \[2006\]](#).

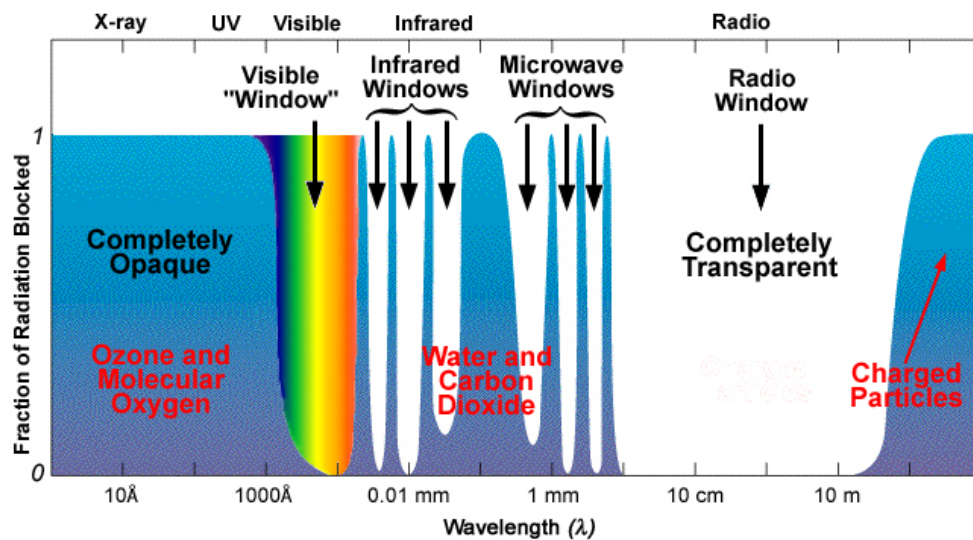


FIGURE 1.1: The atmospheric transmission windows from radio to X-ray wavelengths. Figure created by NASA (<https://earthobservatory.nasa.gov/>).

## 1.2 Hydrological cycle from space

Satellite observations are very useful to study land surface hydrological cycle over large areas from space (Figure 1.2 provides an overview of the hydrological cycle). It is possible to estimate and monitor different hydrological parameters (precipitation, evaporation, groundwater, soil moisture, wetlands, and surface water over rivers or lakes, for example) based on different published techniques. Estimation of surface water extent or precipitation can be done from both visible/infrared and passive/active satellite observations. Active satellite radar altimeters are effective to estimate surface water height over rivers and lakes. Passive microwave satellite estimates are used to extract river flow and discharge information. Satellite-based products are also used to determine variations of surface water extent or inland water storage. A quick summary of remote sensing techniques, mainly focusing on estimation of surface water extent and height, is presented in the following sections.

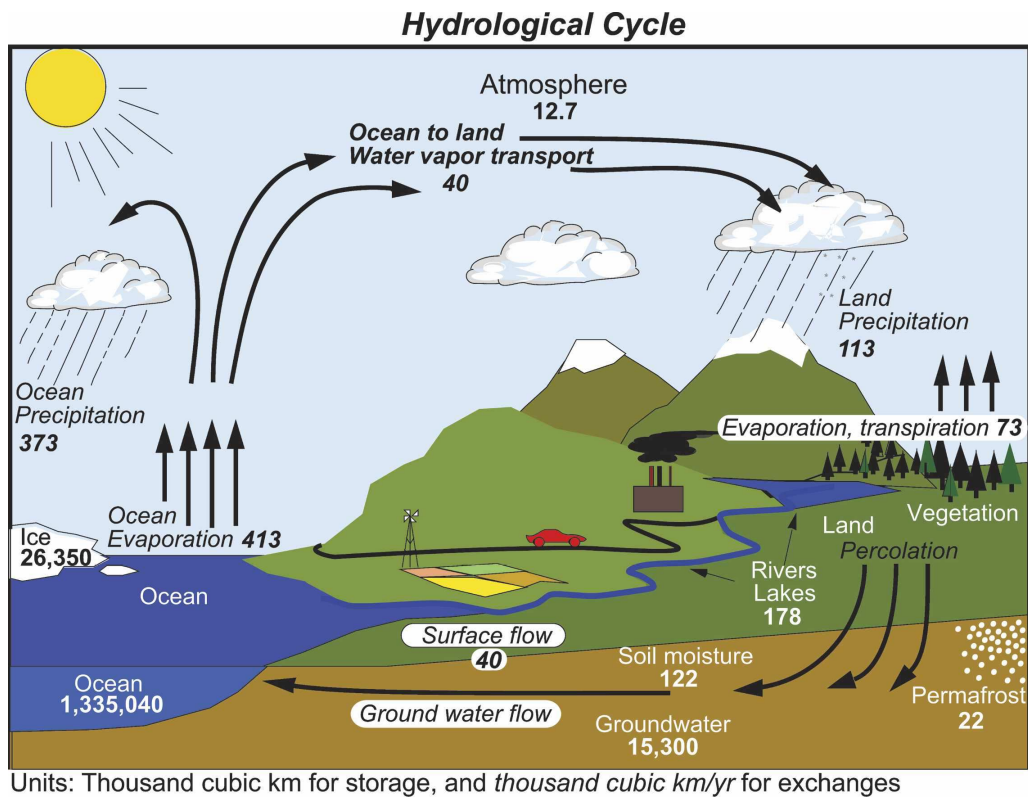


FIGURE 1.2: The hydrological cycle [Trenberth et al., 2007].

## 1.2.1 Estimation of continental surface water extent from space

Surface waters can be defined as water found on the surface of the Earth, such as in lakes, rivers, streams, ponds, wetlands, and natural watercourses. Surface waters are replenished by precipitation and by recruitment from groundwater. They are lost by evaporation to the atmosphere, seeping into the ground, running off into the ocean, or used by plants, animals, as well as humans. Remote sensing of surface waters is of interest because they have strong impacts on all forms of life, as well as on the environment. Surface waters are essential to plants for photosynthesis, to animals for hydration, and to humans for their consumption in living, agriculture and industries. In term of impacts on the environment, surface waters play a key role in the biogeochemical and hydrological cycles, in biodiversity, and in climate variability. Surface waters are the world's largest natural source of methane ( $\text{CH}_4$ ), producing approximately 20%-40% of the world's total annual methane emission to the atmosphere [Houweling et al., 1999]. Hence, the variation of surface water extent contributes to changes of methane surface emissions [Bousquet et al., 2006], that affects the global warming. As a consequence, monitoring surface water extent and their seasonal dynamics, as well as understanding global surface water distribution are necessary for water management, and climate change studies [Alsdorf et al., 2007]. In the literature, different methodologies have been developed for surface water extent estimation, using visible and Near-Infrared (NIR), as well as passive and active microwave satellite observations.

### 1.2.1.1 Visible and Near-Infrared (NIR) observations

Methodologies using visible and NIR satellite imagery for the estimation of surface water extent are well established, and are considered to be reliable at local and regional scales as long as the surface water target is not covered by clouds or dense vegetation [Smith, 1997]. In principle, all water detection methodologies are developed based on the absorption of water bodies in the visible and in the NIR wavelengths [Smith, 1997]. Water bodies absorb almost all incoming energy in near and shortwave infrared wavelengths, causing low to no reflection to the



satellite sensors when the wavelength increases [Smith, 1997]. Based on this phenomenon, reflectance data in the NIR wavelengths can be used as an indicator for water bodies [Bryant and Rainey, 2002; Bergé-Nguyen and Crétaux, 2015], or it can be combined with reflectance data in the visible bands to create water indices, such as the Normalized Difference Vegetation Index (NDVI) [Tucker, 1979], the Normalised Difference Water Index (NDWI) [McFeeters, 1996], the Modified Normalised Difference Water Index (MNDWI) [Xu, 2006], and the Automated Water Extraction Index (AWEI) [Feyisa et al., 2014]. Using these water indices for surface water extent and flood estimation has been tested with different types of visible/infrared satellite imagery, such as NOAA/AVHRR [Bryant and Rainey, 2002; Jain et al., 2006], MODIS [Martinez et al., 2008; Sakamoto et al., 2007; Bergé-Nguyen and Crétaux, 2015], Landsat [Hallberg et al., 1973; Du et al., 2014; Ji et al., 2015; Yamazaki et al., 2015; Pekel et al., 2016], and Sentinel-2 [Du et al., 2016; Yang et al., 2017]. Results showed high classification accuracy.

### 1.2.1.2 Active microwave observations

Using active microwave estimates derived from Synthetic Aperture Radar (SAR) satellite instruments for surface water extent and flood estimation appeared later, but it shows great potential. Microwaves can penetrate clouds, and provide day and night operation that are extremely useful for flood extent measurements. Active microwaves also have the ability to penetrate vegetation, to a certain extent, and this possibility increases when the wavelength increases [Prigent et al., 2016]. The principle of surface water mapping using active microwave satellite observations is based on very low backscatter coefficient over water bodies when the surface is observed off nadir. Energy sent from the active instruments will be reflected in the specular directions, thus very low backscatter coefficient will be received back at the satellite sensors when the incidence angle is off nadir. This makes water bodies appear very dark in the SAR imagery. Based on this principle, many flood events all over the globe have been estimated using SAR observations from different satellites: ERS-2 [Nguyen and Bui, 2001], JERS-1 [Wang, 2004], ENVISAT [Hess et al., 2003; Bartsch et al., 2008; Kuenzer et al., 2013], RADARSAT-1 [Brisco et al., 2009], COSMO-SkyMed [Pierdicca et al., 2013], TerraSAR-X [Voor-mansik et al., 2014], and recently Sentinel-1 [Amitrano et al., 2014]. Despite of

these advantages, global mapping of water bodies using SAR observations is limited due to the lack of global estimations. Compared to visible/infrared observations, SAR observations are more difficult to interpret because the backscatter coefficients are affected by complicated mechanisms. For example, the effect of wind on backscatter coefficient over water surfaces [Smith, 1997; Alsdorf et al., 2007; Kuenzer et al., 2013], the double-bounce scattering mechanism [Hess et al., 1990], or the similarity between water surfaces and flat dry surfaces [Prigent et al., 2015].

### 1.2.1.3 Passive microwave observations

Surface water extent estimation using passive microwave satellite observations has been shown useful for a long time [Giddings and Choudhury, 1989]. The principle of detecting surface water extent using passive microwave satellite estimates is based on the low emissivity of water bodies in both linear polarizations due to differences in dielectric properties of water and soil or vegetation, especially at lower frequencies. Passive microwave can provide global coverage, but the spatial resolution is very coarse (10-50 km), and prevents their use for applications requiring higher spatial resolutions [Prigent et al., 2016]. In addition, at this low spatial resolution, there are large ambiguities in mixture regions of open water bodies and other types of environments (such as soil, or vegetation). To achieve better accuracy, some authors combined passive microwave with active microwave and visible/NIR satellite estimates to develop methodologies producing global surface water maps over long periods [Prigent et al., 2001, 2007, 2012; Schroeder et al., 2015].

## 1.2.2 Estimation of surface water height from space

Satellite radar altimetry technique was originally developed for oceanographic applications, and has been proved to be the most successful method for measuring continental surface water elevations from space since the early 1990s [Alsdorf et al., 2007]. Then this technique has been found to be a valuable source to monitor surface water variations over large lakes [Birkett, 1995] and rivers [Birkett, 1998].

Many authors have shown studies using satellite altimetry data for monitoring water height of inland water bodies [Birkett et al., 2002; Coe and Birkett, 2004; Crétaux and Birkett, 2006; Frappart et al., 2006a,b; Papa et al., 2015]. Satellite radar altimeters are nadir-viewing instruments primarily designed to measure the two-way time delay of radar pulses emitted from satellite sensors and reflected from the surfaces as the satellite orbits around the Earth [Birkett et al., 2002], that gives the distance between the satellite sensors and the reflected surface [Crétaux et al., 2016]. This distance measurement combined with the accurate knowledge of the satellite position (using Laser Retro-Reflector, GPS, or DORIS systems), enables calculation of the surface water height with respect to a reference ellipsoid [Birkett et al., 2002; Frappart et al., 2006b; Crétaux et al., 2016]. Although radar altimetry is independent from weather conditions and has the unique ability to provide data at remote or ungauged locations, there are limitations. Elevation calculations from satellite altimeters and its accuracy can be affected by topography, size and surface roughness of lakes or rivers. To obtain reliable accuracy in surface water level, corrections have to be taken into account (instrument corrections, propagation corrections, atmospheric corrections, geophysical corrections, ect). Current radar altimeters only estimate elevations at the nadir direction along the satellite orbit track, therefore, it cannot provide data all over the globe [Alsdorf et al., 2007]. However, the concept of using a constellation of several satellites at the same time may enable monitoring surface water variation for a large area [Crétaux et al., 2015]. In addition, the future joint US-French Surface Water and Ocean Topography (SWOT) mission can provide altimetry estimates on both sides of the satellite track (up to 50 km) due to its ability to observe the Earth surface in near nadir incidence angles ( $0.6^{\circ}$ - $3.9^{\circ}$ ) [Fjrtoft et al., 2014].

### 1.2.3 Estimation of other hydrological components from space

Other hydrological components can be estimated from space to some extent, using different satellite data and techniques.

Rainfall estimation from low-Earth and geostationary orbit satellites has a long history since 1970s, with different techniques using wavelength ranges from visible/infrared to microwave (both passive and active) to access to different

cloud layers. A complete overview of rainfall estimation from space can be found in a book edited by [Levizzani et al. \[2007\]](#). Nowadays, there are several satellite-based datasets that provide near-real time global precipitation estimations: the gauge-calibrated Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis (TMPA, 3B42 V7), or the Global Precipitation Climatology Project (GPCP).

The Gravity Recovery and Climate Experiment (GRACE) satellite, launched in 2002, provides data to monitor global inland water storage. Detailed information on the GRACE mission, as well as the processes to extract terrestrial water storage anomalies can be found in review papers, for example [Wouters et al. \[2014\]](#) or [Wahr \[2015\]](#). Many studies have successfully used GRACE data to quantify water storage changes in many regions all over the globe [[Tapley et al., 2004](#); [Wahr et al., 2004](#); [Leblanc et al., 2009](#); [Famiglietti et al., 2011](#); [Feng et al., 2013](#)].

The surface water volume variations over some river basins in the world have been studied by combining different inundation and altimetry satellite products [[Frappart et al., 2005, 2008, 2010, 2011](#); [Papa et al., 2008a, 2013](#)]. Recently, river flow and discharge information of major river basins in all continents have been estimated [[Brakenridge and Nghiem, 2017](#)] based on a methodology that uses mainly historical and current passive microwave satellite data [[Brakenridge et al., 2005](#)].

## **1.3 The study area: The Vietnam Mekong Delta and the Tonle Sap Lake in Cambodia**

### **1.3.1 Presentation of the region**

Most parts of this thesis focus on monitoring surface water extent, surface water height, and surface water volume over the lower Mekong basin using satellite observations. The Mekong river is one of the longest rivers in the world, ranking 12<sup>th</sup>, at ~4,350 km in length [[MRC, 2011](#)]. The river starts from the Himalayas at an elevation of about 5,000 km, then it runs through China, Myanmar, Laos,

### Mekong overview



FIGURE 1.3: Overview of the Mekong River and its catchment [MRC, 2011].

Thailand, Cambodia and Vietnam before entering the East Sea [MRC, 2011]. The Mekong basin is commonly divided into two parts: the upper Mekong basin, mostly located in China, and the lower Mekong basin, located in other countries [MRC, 2011]. The lower Mekong basin is important because there are nearly 61 million people living within the catchment of the river. This corresponds to about 35% of the total population of the four countries. It is even more important to Cambodia and Laos because 81% and 89% of their population live in the catchment of the Mekong river [MRC, 2011]. This study focuses on the Vietnam Mekong Delta and the Cambodia Tonle Sap Lake as they are two regions affected the most by climate change and global warming. The Vietnam Mekong Delta in Southeast Asia (one of the largest deltas in the world) is a vast triangular plain of approximately 55,000 km<sup>2</sup>, most of it lower than 5 m above the sea. The Mekong Delta is bounded by the Gulf of Thailand in the southwest, the East Sea in the east, and Cambodia in the northwest. The seasonal variation in water level results in rich and extensive wetlands. For instance, the Mekong Delta region covers only 12% of Vietnam but produces ~50% of the annual rice (with two or three harvests per year depending on the provinces), represents ~50% of the fisheries, and ~70% of the fruit production. The Tonle Sap Lake (also called the Great Lake), is the largest freshwater lake in Southeast Asia. It is located in the middle of Cambodia. The lake and its floodplain are very important for the ecosystems because it is home to nearly 150 fish species, reptiles and birds. The lake is also important to local inhabitants since it supports daily food and livelihood for nearly 3 million people through their fishing and farming activities. In the Mekong Delta and the Tonle Sap Lake, the dry season extends from November to May and the rainy season from June to October. During the rainy seasons, the Tonle Sap Lake is filled with water flowing from the Tonle Sap River, making its water height and flooded surface areas being maximum (10-12 m and ~ 10,000 km<sup>2</sup>, respectively). It is several times higher than that at their minimum states during the dry seasons when water throws out from the lake to the Mekong Delta (1-2 m for water height, and ~ 3,000 km<sup>2</sup> for flooded surface areas). An overview of the Mekong River and its catchment is shown in Figure 1.3.

### **1.3.2 The satellite remote sensing studies so far over the region**

The Vietnam Mekong Delta and Cambodia are an attractive region to many authors to study the ability of satellite observations in monitoring land surface hydrology. [Sakamoto et al. \[2007\]](#) used visible/infrared MODIS data while [Nguyen and Bui \[2001\]](#) and [Kuenzer et al. \[2013\]](#) used SAR data to detect temporal changes of floods in this region. Water height estimation with satellite altimetry data, as well as water volume variations over the Mekong Delta were introduced in [Frappart et al. \[2006b\]](#). Daily discharge records over the main stream of the Mekong river can be found in [Brakenridge and Nghiem \[2017\]](#).

The use of observations from the VNREDSAT-1 satellite, the first optical Earth observing satellite of Vietnam launched in May 2013, is still limited for hydrological monitoring applications over large areas because of the small satellite swath (17.5 km). In addition, long processes are required to be able to get access to VNREDSAT-1 data. The JV-LOTUSat-1, the first Vietnam radar satellite (working in X-band), is planned to be launched in 2020 with support from Japan. JV-LOTUSat-1 data are expected to be used for many applications (agriculture, land use/land cover, topography, and disaster management).

## **1.4 Thesis objectives**

Monitoring and quantifying the variation of surface water extent is essential to study the water cycle. This thesis is developed to benefit from publicly available satellite observations for water management. The lower Mekong basin in Vietnam and Cambodia is chosen as the studied area for this thesis. Results from the thesis are distributed for further studies over the Mekong basin (e.g. flood analysis and possibly flood warning, studies on the effects of the climate change on the Mekong delta). Thesis results can be used as input for regional climate models to predict the variation of hydrological components over the short- and long-term. Scientific results from this study can be useful for managers, decision makers or insurance companies to understand what is happening in the region, to predict

the changes in the future, and to build plans for mitigation of damages caused by extreme weather events like floods.

In this thesis, four main objectives will be pursued:

- The first objective is to build surface water extent maps over the Mekong Delta and Cambodia based on visible/infrared MODIS/Terra satellite observations. Several classification methodologies will be studied and compared to a reference dataset to find the most suitable methodology for the area. Then, the chosen classification method will be applied to provide regular surface water estimates in the region, from January 2001 until the present time.
- The second objective is similar to the first one, but with SAR Sentinel-1 satellite observations to benefit from the advantages of SAR data over highly cloudy areas. Free-cloud optical Landsat-8 satellite data are used as targets to train an optimized Neural Network (NN) method. Regular SAR Sentinel-1 surface water extent estimates in the region can be produced since February 2015 until the present time.
- In the third objective, a direct application of the surface water extent in the previous chapters is studied. That is the calculation of the variation of the surface water volume over the Mekong Delta and Cambodia. It is calculated as the product of the MODIS-derived surface water extent and the surface water height derived from satellite altimetry data. The validation of the two components, the surface water extent and height, is performed first to check their consistency, then results are evaluated with the variation of the total land surface water volume from GRACE data, as well as with the variation of the in situ discharge estimates of the Mekong River.
- The last objective is to analyze the variation of global surface water extent by comparing mainly two global multi-satellite surface water datasets with monthly time series over the 1993-2007 period. Global comparisons between the two datasets are performed, followed by regional analyses over the 23 largest basins in the world. Ancillary data (river discharge, for example) are used for comparisons when available.



The structure of the thesis closely follows these objectives. Chapter 2 and 3 present the methodologies using optical and SAR satellite observations, respectively, to monitor the surface water extent over the lower Mekong basin in Vietnam and Cambodia. Chapter 4 focuses on calculation of the monthly surface water volume changes in the region. Chapter 5 focuses on comparisons between two monthly multi-satellite surface water datasets globally and regionally. Chapter 6 concludes this thesis and provides scientific perspectives for future works.

## **Chapter 2**

# **Surface Water Monitoring within the Mekong Delta and Cambodia using Visible and Infrared MODIS Satellite Observations**



## 2.1 Introduction

In this chapter, the use of visible/infrared satellite observations for surface water monitoring will be studied. Detecting surface water using visible/infrared satellite observations began in the early 1970s with the use of Earth Resources Technology Satellite (ERTS-1, later renamed as Landsat-1) data for flood mapping in several locations (in Iowa [Hallberg et al., 1973], in Arizona [Morrison and Cooley, 1973], in Virginia [Rango and Salomonson, 1974], and in Australia [Robinove, 1978], for example). Since then, many satellites equipped with optical instruments have been launched, providing regional and global imagery at different spatial and temporal resolutions. Observations from some satellites are free of charge for final users (NOAA/AVHRR, Landsat, MODIS, and Sentinel-2, for example), and have been used widely for flood and surface water monitoring applications. The Advanced Very High Resolution Radiometer (AVHRR) instrument onboard the National Oceanic and Atmospheric Administration (NOAA) satellites was first launched in 1978. The AVHRR instrument was designed to monitor clouds and to measure the thermal emission of the Earth, but then its data were discovered to be suitable for flood detection applications [Islam and Sado, 2000; Nyborg and Sandholt, 2001; Domenikiotis et al., 2003; Jain et al., 2006]. Similar to the AVHRR instrument, the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument onboard the Terra and Aqua satellites has provided regular long-term remote sensing data (since 2000) for surface water monitoring applications [Martinez et al., 2008; Peng et al., 2009; Sakamoto et al., 2007; Bergé-Nguyen and Crétaux, 2015]. Although NOAA/AVHRR and MODIS instruments have a high temporal resolution (twice a day), global coverage during a long period of time, their spatial resolutions are rather coarse (1100 m for NOAA/AVHRR, and maximum 250 m for MODIS) and limit their data to applications requiring low spatial resolution. Data from the Landsat satellite series, and recently the Sentinel-2 satellite can be a better choice for high resolution applications. Landsat satellites provide global observations at 30 m spatial resolution, but their temporal resolution is 16 days. Landsat-8 is the latest satellite in the series and was launched in 2013. Observations provided by the Operational Land Imager (OLI) instrument onboard the Landsat-8 satellite have been used widely to detect surface water in

many studies [Du et al., 2014; Ji et al., 2015; Acharya et al., 2016]. Long-term Landsat series data are also used for creation of surface water maps at continental scale [Mueller et al., 2016] as well as at global scale [Yamazaki et al., 2015; Pekel et al., 2016]. The newest satellite that provides freely visible observations is Sentinel-2 that was launched in 2015. It is a pair of satellites with 5-day temporal resolution (10 days with only one satellite). The MultiSpectral Instrument (MSI) onboard the Sentinel-2 satellites provides observations with different spatial resolutions ranging from 10 m to 60 m. Some authors started using Sentinel-2 data for mapping surface water bodies [Du et al., 2016; Yang et al., 2017]. Detailed information of these four "free" satellites are shown in Table 2.1.

TABLE 2.1: Four commonly used satellites providing freely visible/infrared observations. Sources: NASA & ESA.

Satellites	Agency	Number of bands	Temporal Res. (day)	Spatial Res. (m)	Swath (km)
NOAA/AVHRR (1978-present)	NOAA	4-6	0.5	1100	2800
MODIS (2000-present)	NASA	36	0.5	250-1000	2330
Landsat (1972-present)	NASA	4-9	16	30-80	185
Sentinel-2 (2015-present)	ESA	13	5	10-60	290

The principle of detecting surface water using visible/infrared satellite observations is based on the very low reflectance of water compared to other land surface types. Surface reflectance is the amount of light reflected by the Earth surface, and it is calculated by the ratio between surface radiance and surface irradiance. It typically has values between 0 and 1. When a body is illuminated by electromagnetic radiation, at the surface of the body, energy will be partly transmitted, absorbed or reflected, depending on the wavelength and the incidence angle of the incoming radiation, and on the material and roughness of the surface of the body. Objects having different surface features transmit, absorb or reflect radiation in different ways, therefore, they have different spectral signatures. Therefore, by studying these spectral signatures, it is possible to identify different surface features or materials.

Figure 2.1 shows the spectral signatures from visible to mid-infrared (MIR) of three common natural surfaces on the Earth: dry bare soil, green vegetation and clear water bodies. Taking clear water bodies as an example, they absorb effectively all wavelengths longer than the visible range. It causes a very low surface reflectivity in the range of infrared wavelengths. Green vegetation has a very special spectral signature since its surface reflectance is low in the visible range but higher for the green light than for the blue and red light due to absorption of chlorophyll before it increases dramatically in the domain of the near-infrared (NIR) wavelengths. In case of the dry bare soil, its surface reflectance rises gradually as a function of the wavelengths, reaching the peak in the middle of the MIR range. Based on these basic physical principles, surface reflectance data of NIR wavelengths are often used to detect open water bodies such as lakes or rivers because these water surfaces appear very dark in NIR reflectance images. Bryant and Rainey [2002] applied a threshold on NIR surface reflectances of the NOAA/AVHRR satellite data to delineate lakes. However, using a single band to detect surface water usually yields large errors because of misclassification between water pixels and other land surface types [Du, 2012]. A more effective way to detect surface water with higher accuracy is to use different water indices that are calculated from two or more bands. Several water indices have already been proposed.

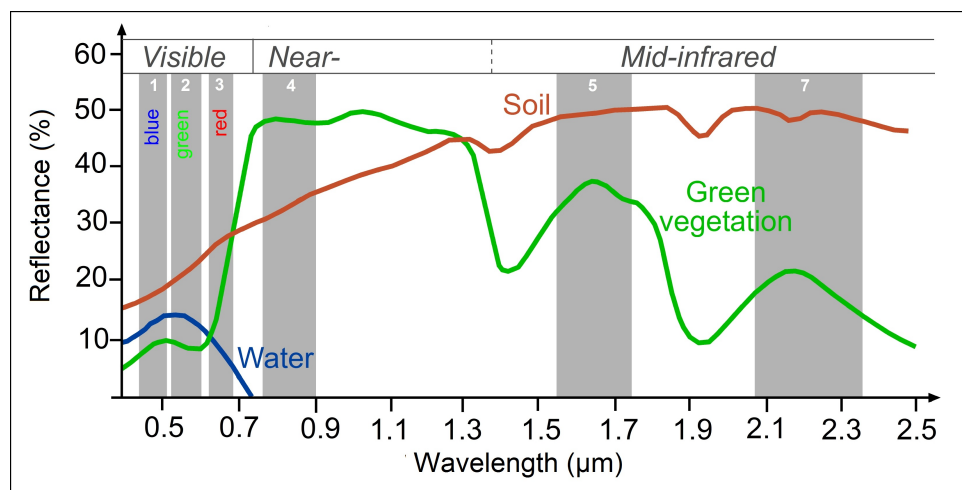


FIGURE 2.1: Spectral signatures of soil, vegetation and clear water as a function of wavelengths from visible to MIR. Numbers 1-7 indicate visible, NIR and MIR wavelength ranges, respectively. Source: SEOS project (<http://www.seos-project.eu/home.html>.)

McFeeters [1996] introduced a technique using the Normalized Difference Water Index (NDWI) to classify surface water bodies. NDWI is calculated as the ratio between the green and the NIR wavelengths (equation 2.1) that makes surface water have positive values.

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (2.1)$$

Later, Xu [2006] replaced the NIR band by the MIR band to make the Modified of Normalized Difference Water Index (MNDWI). Surface water still has positive values, but this MNDWI is reported to reduce errors between water and build-up land pixels [Rokni et al., 2014]. The Normalized Difference Vegetation Index (NDVI) is the ratio between the NIR and the red wavelengths (equation 2.7). This index was developed to study the vegetation canopy, but it can work as a water index since surface water has negative NDVI values [Rouse Jr. et al., 1974; Domenikiotis et al., 2003; Rokni et al., 2014]. NDVI data are also combined with NIR reflectance data to classify water bodies [Cretaux et al., 2011]. Several indices (the Enhanced Vegetation Index (EVI), the Land Surface Water Index (LSWI), and the Difference Values between EVI and LSWI (DVEL)), are also combined together for flood mapping and flood monitoring [Sakamoto et al., 2007]. Recently, a new water index called the Automated Water Extraction Index (AWEI) was introduced by Feyisa et al. [2014] that improves classification accuracy in shadow and dark surfaces. AWEI has two indices used for shadow and non-shadow situations (equations 2.2 and 2.3), and surface water has positive AWEI values.

$$AWEI_{nsh} = 4 \times (GREEN - MIR1) - (0.25 \times NIR + 2.75 \times MIR2) \quad (2.2)$$

$$AWEI_{sh} = BLUE + 2.5 \times GREEN - 1.5 \times (NIR + MIR1) - 0.25 \times MIR2 \quad (2.3)$$

Comparisons between different water indices in different environments using Landsat data are performed in some studies [Yang et al., 2011; Rokni et al., 2014; Zhai et al., 2015]. It is concluded that the accuracy of each water index depends on studied areas and the reference information. Therefore, there is no best index among all of them, and the selection of the most suitable water index can be different depending on specific situations. Although using water indices can provide water surface maps with a good accuracy, cloud contamination is still

the greatest limitation for the use of water indices since nearly 60% of the Earth surface is cloud covered at any moment. This makes their applications are only suitable for cloud-free conditions, which is very restrictive in some regions (e.g., in the Tropics). Vegetation can also mask the surface water partly or totally, making the water detection difficult or even impossible under canopy. In addition, the NIR reflectance over highly turbid water can be higher than over the visible reflectance, introducing confusions in the indices used for the water detection. Despite these limitations, flood and surface water mapping using visible/infrared satellite observations can be efficient, and provide good agreements with results derived from aerial photography [Smith, 1997]. Among all available visible/infrared satellite data, Landsat and MODIS data are widely used because their data are now free of charge, quick to download, and cover a long period of time.

This chapter focuses on using visible/infrared MODIS/Terra observations for detecting and monitoring surface water extent over the Vietnamese Mekong Delta and Cambodia, from 2001 to present time. MODIS images have moderate spatial resolution (500 m) compared to Landsat one (30 m), but the temporal resolution is much higher (twice a day compared to 16 days) making MODIS data more suitable for monitoring applications. Section 2.2 provides more details of the MODIS/Terra observations, as well as 4 different methodologies that use MODIS/Terra observations for surface water detection. Pre-processing steps applied to the data are described in Section 2.3. Results are shown in Section 2.4. Section 2.5 concludes this chapter.

## **2.2 Different methodologies using MODIS/Terra observations for surface water monitoring**

### **2.2.1 MODIS/Terra observations**

The Moderate Resolution Imaging Spectroradiometer (MODIS) instrument is one of the key instruments onboard the Terra satellite (launched on 18 December 1999), and the Aqua satellite (launched on 4 May 2002) from NASA. The two satellites are sun-synchronous, near-polar, and are orbiting at an altitude of  $\sim 705$



km. The two satellites were designed so that the Terra satellite passes the equator from north to south at 10.30 am (local time) and the Aqua satellite passes the equator from south to north at 1.30 pm (local time). MODIS instrument provides global Earth observations in 36 spectral bands ranging from visible (0.4  $\mu\text{m}$ ) to infrared wavelengths (14.4  $\mu\text{m}$ ). MODIS instrument provides observations at different spatial resolutions (250 m, 500 m, and 1 km depending on the wavelength). The two satellites can cover the entire Earth's surface every 1-2 days. MODIS Terra atmospherically corrected surface reflectance 8-day Level-3 (L3) Global 500 m (MOD09A1) products used in this study are provided freely through the NASA's Earth Observing System Data and Information System (EOS-DIS; <https://reverb.echo.nasa.gov/reverb/>). MOD09A1 data provide atmospherically corrected surface reflectance information of seven different bands from optical to infrared wavelengths in two native spatial resolutions. Band 1 (red) and band 2 (NIR) are at the highest spatial resolution of 250 m, and bands 3-7 are at 500 m spatial resolution. MODIS data are available from 20 February 2000 until present. Each MOD09A1 image was created by selecting the best Level-2 gridded (L2G) observation during an 8-day period on the basis of high observation coverage, low view angle, absence of clouds or cloud shadow, and aerosol loading [Vermote, 2015]. Table 2.2 shows details of the MOD09A1 seven bands and their main applications. Each MOD09A1 image has the dimension of 2400 $\times$ 2400 (rows/columns) at 500 m spatial resolution. One image covers an area of approximately 10 $^{\circ}$  $\times$ 10 $^{\circ}$ , projected on the sinusoidal projection.

TABLE 2.2: MOD09A1 seven bands and their main applications. Source: NASA.

Band number	Wavelength range ( $\mu\text{m}$ )	Resolution (m)	Key applications
1	0.620-0.670	250	Absolute Land cover Transformation Vegetation Chlorophyll
2	0.841-0.876	250	Cloud Amount Vegetation Land cover Transformation
3	0.459-0.479	500	Soil/Vegetation Differences
4	0.520-0.600	500	Green Vegetation
5	1.230-1.250	500	Leaf/Canopy Differences
6	1.628-1.652	500	Snow/Cloud Differences
7	2.105-2.155	500	Cloud Properties, Land Properties

### 2.2.2 Sakamoto et al. [2007] methodology

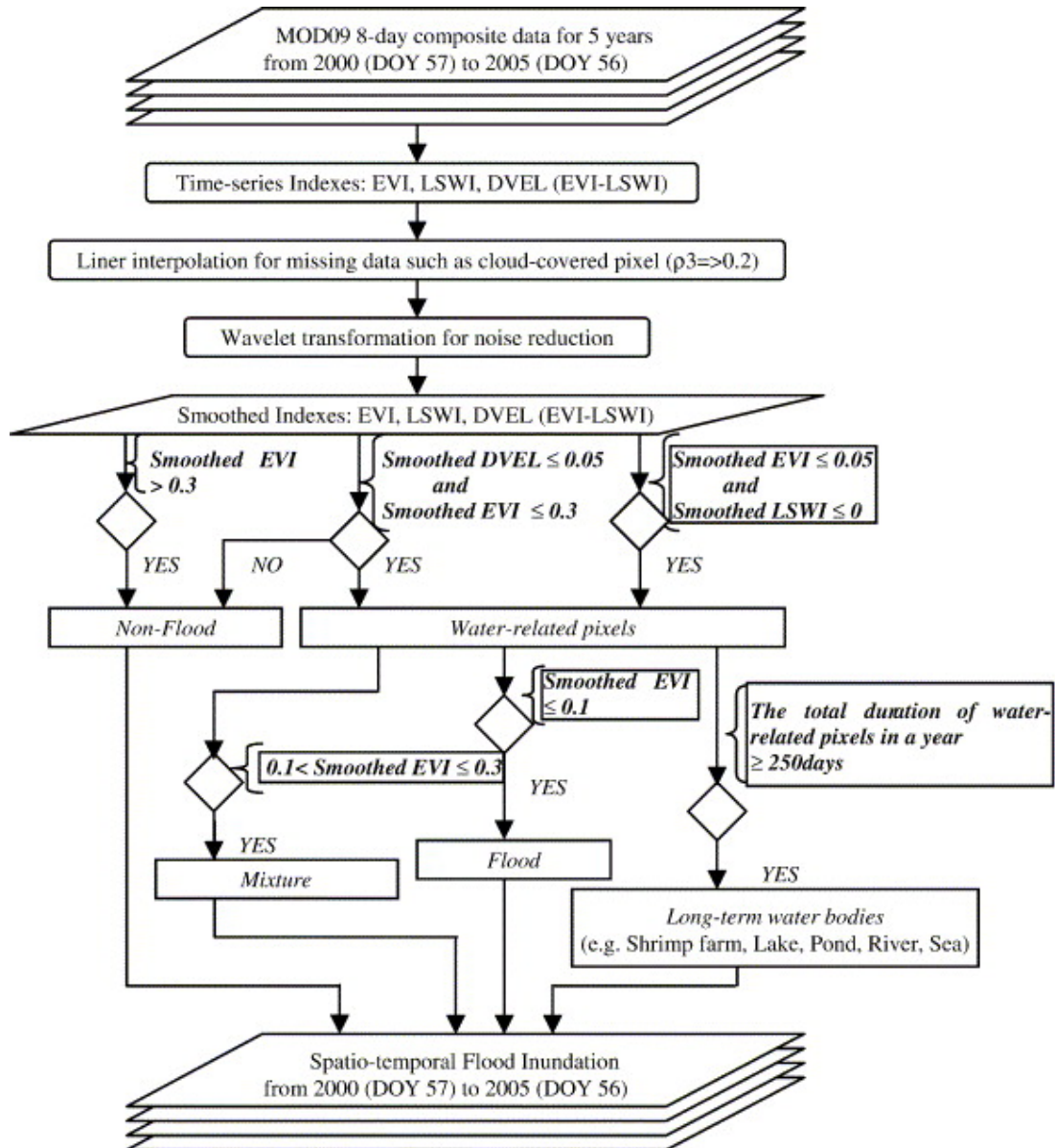


FIGURE 2.2: Flowchart describing the methodology for detecting the spatio-temporal changes of flood inundation by Sakamoto et al. [2007].

Sakamoto et al. [2007] presented a methodology for detecting the spatio-temporal changes in the extent of annual flooding over the Mekong River Delta in Vietnam using MOD09A1 time series imagery. Details of the algorithm are shown in Figure 2.2. First, the Enhanced Vegetation Index (EVI), the Land Surface Water Index (LSWI), and the Different Value between EVI and LSWI (DVEL) are

calculated at pixel level, from the original atmospherically corrected surface reflectance data derived from the MOD09A1 8-day composite products. The EVI and the LSWI are obtained using equation 2.4 and equation 2.5, while the DVEL is calculated from equation 2.6:

$$EVI = 2.5 * \frac{NIR - RED}{NIR + 6 * RED - 7.5 * BLUE + 1} \quad (2.4)$$

$$LSWI = \frac{NIR - MIR}{NIR + MIR} \quad (2.5)$$

$$DVEL = EVI - LSWI \quad (2.6)$$

where RED, NIR, BLUE and MIR are the surface reflectance values of visible band 1 (red band; 0.62 - 0.67  $\mu\text{m}$ ), NIR band 2 (0.841 - 0.876  $\mu\text{m}$ ), visible band 3 (blue band; 0.459 - 0.479  $\mu\text{m}$ ) and MIR band 6 (1.628 - 1.652  $\mu\text{m}$ ), respectively. Next step, a linear interpolation is used to deal with missing data such as cloud-covered pixels (pixels where surface reflectance values of the blue band are greater than 0.2), then a wavelet transformation is applied to reduce noise for the indices. After this step, all the EVI, LSWI, and DVEL indices are smoothed and ready for the classification step.

For the classification, all pixels with smoothed EVI values greater than 0.3 are classified as non-flood pixels. Water-related pixels are marked when the smoothed DVEL values are smaller than or equal to 0.05 and the smoothed EVI values are smaller than or equal to 0.3. If a pixel has the smoothed EVI value smaller than or equal to 0.05 and the smoothed LSWI value smaller than or equal to 0, it is also marked as a water-related pixel. At the final step, a threshold of the smoothed EVI values is used to distinguish between mixture pixels and fully inundated pixels from water-related pixels. Mixture pixels are defined as pixels that are partly inundated. The smoothed EVI values in open water bodies such as lakes or the ocean are normally low, therefore, if the smoothed EVI values are smaller than or equal to 0.1, these water-related pixels are set as the fully inundated pixels. If the smoothed EVI values are greater than 0.1 and smaller than or equal to 0.3, then these water-related pixels are marked as the mixture pixels. Based on these conditions, inundation maps over the Mekong River Delta can be re-built. Results are shown in section 2.4.

### 2.2.3 Bergé-Nguyen and Crétaux [2015] methodology

Bergé-Nguyen and Crétaux [2015] also developed a methodology for flood and wetland mapping over the Inner Niger Delta (IND), based on the same MOD09A1 time series imagery. This methodology is simpler than Sakamoto et al. [2007] one since it uses mainly surface reflectance information derived from visible band 1 (red band; 0.62 - 0.67  $\mu\text{m}$ ), NIR band 2 (0.841 - 0.876  $\mu\text{m}$ ) and NIR band 5 (1.23 - 1.25  $\mu\text{m}$ ) to classify surfaces into five different types, including open water, mix water/dry land, aquatic vegetation, vegetation, and dry land. Details of this methodology are described in Table 2.3. First, the Normalized Difference Vegetation Index (NDVI) for each single pixel is calculated using equation 2.7. Since open water bodies absorb most of radiation in the infrared wavelengths, values of surface reflectance of NIR band 5 (1.23 - 1.25  $\mu\text{m}$ ) is selected to determine open water bodies. When band 5 surface reflectance value of a pixel is smaller than 0.12, that pixel is marked as a fully inundated pixel. When that value increases to the threshold value of 0.27, that pixel is marked as a water-related pixel, and when it overcomes the threshold of 0.27 that pixel is marked as a non water-related pixel. For these water-related pixels, the NDVI is used to distinguish between mix of water/dry land and aquatic vegetation. If the NDVI of a water-related pixel is smaller than 0.4, the pixel is considered as a mix water/dry land and if the NDVI is greater than 0.4, the pixel is classified as an aquatic vegetation pixel. For these non water-related pixels, the NDVI is tested to detect the presence of vegetation on dry land. If the NDVI of one pixel in this class is smaller than 0.4, this pixel is called a dry land pixel, and if the NDVI is greater than 0.4, then this pixel is considered to be covered by vegetation. Due to its simple criterion, this methodology is reported to be more suitable for the arid and semi-arid regions (like the Inner Niger Delta), but less precise for the equatorial and boreal regions where the cloud contamination is high [Bergé-Nguyen and Crétaux, 2015]. However, this methodology is still applied over the Mekong River Delta and Cambodia for comparison to results obtained from the Sakamoto et al. [2007] methodology. More details will be shown in the section 2.4.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (2.7)$$

TABLE 2.3: Threshold values used for classifying terrain surface into five different classes by [Bergé-Nguyen and Crétaux \[2015\]](#).

Index	Open Water	Mix Water/ Dry Land	Aquatic Vegetation	Vegetation	Dry Land
Band 5	<0.12	>0.12 & <0.27	>0.12 & <0.27	>0.27	>0.27
NDVI	No test	<0.4	>0.4	>0.4	<0.4

## 2.2.4 Unsupervised and supervised classification techniques

Beside using criterion and indices to classify, image classification can be done based on two major techniques: unsupervised and supervised classifications. Each technique has its own advantages and disadvantages, and they are used for different purposes. More information on classification techniques can be found in remote sensing books, such as [Jensen \[1995\]](#). Unsupervised and supervised are performed here to calibrate classification results obtained from the two approaches described in the two previous sections.

### 2.2.4.1 Unsupervised classification

The principle of unsupervised classification methodology is to group pixels within an image that have similar spectral signature together into the same groups. The most popular algorithms used in unsupervised classification are K-means and ISODATA methods. For the K-means classification, the user selects the number of classes to obtain,  $n$  classes, for example. At the very first step,  $n$  pixels in the image are chosen randomly as  $n$  bary-centers. The second step, each single pixel in the image is associated with one bary-center by calculating and comparing the Euclidean distance from all pixel to each  $n$  bary-centers. The third step, new  $n$  bary-centers are re-calculated based on classification results in the previous step, and again each single pixel in the image is associated with a bary-center. This process runs several times until all pixels in the image stop changing their classes. Finally, all pixels in the image are classified into  $n$  classes (or groups) as expected. The ISODATA algorithm has the same principle as the K-means one, but the number of classes may be varied by splitting a cluster into smaller ones or merging

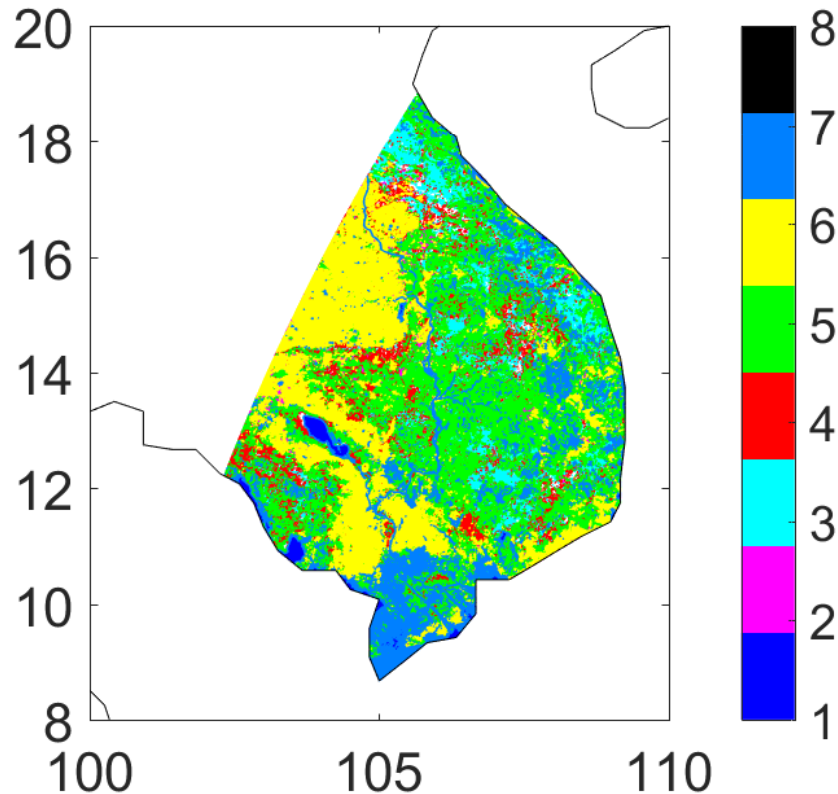


FIGURE 2.3: Classification results with 8 clusters after applying K-means algorithm to a MODIS image over land areas of Vietnam and Cambodia in February 2007.

several clusters together to create a bigger one. If the distance between two clusters is smaller than a given threshold, they will be merged together making a bigger cluster. And if the standard deviation in any dimension is greater than a given threshold for a cluster, it will be split into two different ones. In addition, if a cluster has the total number of individual pixel smaller than a given number, it could be deleted from the calculation. The unsupervised classification method is easy to perform, however, this method is often used when the users do not know very well about the environment of the studied areas.

An unsupervised classification is designed to classify each MODIS image into three clusters (water, mixed and non-water). First, using the K-means algorithm, the image is classified into 8 clusters. Several tests have been made to conclude that classifying into 8-10 clusters is optimal for this purpose. Water-related clusters are determined based on the low values of their bary-centers, and the spatial

structure of the classification map (see Figure 2.3, colors 1 and 7). Then, pixels associated to the other 6 clusters are marked as non-water pixels. Next, the K-means algorithm is used another time to classify the remaining water-related pixels into mixed-pixel or water-pixel clusters. Finally, the image is completely classified, as expected. Bary-centers of the two clusters at the seven MODIS bands derived from the unsupervised classification method are shown in Table 2.4.

#### 2.2.4.2 Supervised classification

The supervised classification is guided by a training dataset provided by the user. The user defines "training sites" in the image to be classified. "Training sites" are regions known as a particular land cover type. The algorithm calculates and defines the spectral signature of all the pixels within each "training site". Then, based on the spectral signature of each pixel in the image, the algorithm determines possibilities belonging to each land cover type in the training dataset. By comparing its possibilities, a pixel is associated with a land cover type it matches most closely. Defining "training sites" is very important for supervised classification because it strongly affects the accuracy of the classification. Classification accuracy increases with accuracy of the training dataset. Supervised classification method is often used when the user knows very well the studied area.

A supervised classification method using the minimum distance algorithm is designed for the same objective to classify MODIS imagery into three classes, as described above. Four surface water maps created based on the Sakamoto et al. [2007] method are used as training dataset. The four images are selected in January, April, July and October 2007, to improve the ability to capture the wetland dynamics in a year. The training dataset contains  $\sim 9$  million pixels already classified into three classes. First, mixed pixels and water pixels in the training dataset are merged into one class, called the water-related class. At this step, the training dataset only has two clusters (non-water and water-related). Then, each pixel in the MODIS image is classified into either non-water group or water-related group, using the minimum distance algorithm. Next, the water-related pixels are classified one more time to put them into the mixed-pixel or the water-pixel

group. The bary-centers of the two clusters derived from the supervised classification technique are also shown in Table 2.4.

Bary-centers of mixed-pixel and water-pixel classes derived from both unsupervised and supervised classification techniques are very close, meaning that surface water maps derived from the two techniques are also close. More comparisons will be shown in section 2.4.

TABLE 2.4: Bary-centers of surface reflectance at the MODIS seven bands for mixed-pixel and water-pixel clusters derived from unsupervised and supervised classification techniques.

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
<b>Mixed pixels</b>							
Unsupervised	0.08	0.20	0.06	0.08	0.19	0.13	0.07
Supervised	0.10	0.21	0.07	0.10	0.19	0.13	0.07
<b>Water pixels</b>							
Unsupervised	0.08	0.08	0.05	0.08	0.07	0.05	0.02
Supervised	0.09	0.09	0.06	0.09	0.07	0.05	0.03

### 2.2.4.3 Hybrid classification

Unsupervised and supervised classifications can also be used together in a process called hybrid classification. Unsupervised classification can be used first to determine the number of land cover types that can be defined, then supervised classification can be used to classify the image into the land cover classes of interest. However, in this chapter, unsupervised and supervised classification techniques are applied separately to build surface water maps. The same MOD09A1 products used in Sakamoto et al. [2007] and Bergé-Nguyen and Crétaux [2015] are used to run both unsupervised and supervised classifications.



## 2.3 MOD09A1 pre-processing

### 2.3.1 MOD09A1 quality and reliability checking

85%-95% of the Mekong River Delta is often covered by thick clouds during the rainy seasons [Leinenkugel et al., 2013], causing problems to optical sensors like MODIS because these observations cannot penetrate clouds. The relationship between the presence of clouds and the variability of MOD09A1 surface reflectance needs to be checked to quantify the quality as well as the reliability of the MOD09A1 products. Surface reflectance pixels located over the ocean are removed from the original data using a land/ocean mask to obtain only surface reflectances over land area. Figure 2.4 shows land surface reflectance maps from the seven MODIS bands over Cambodia and Vietnam, after applying the land/ocean mask.

A MOD09A1 data file contains a quality assurance flag (4 states) and two cloud flags. The first cloud flag is the internal cloud algorithm flag (two states: cloud or no cloud), and the second one is the MOD35 cloud flag (four states: clear, cloud, mixed, or not set and assumed clear). Figure 2.5 shows cloud covered

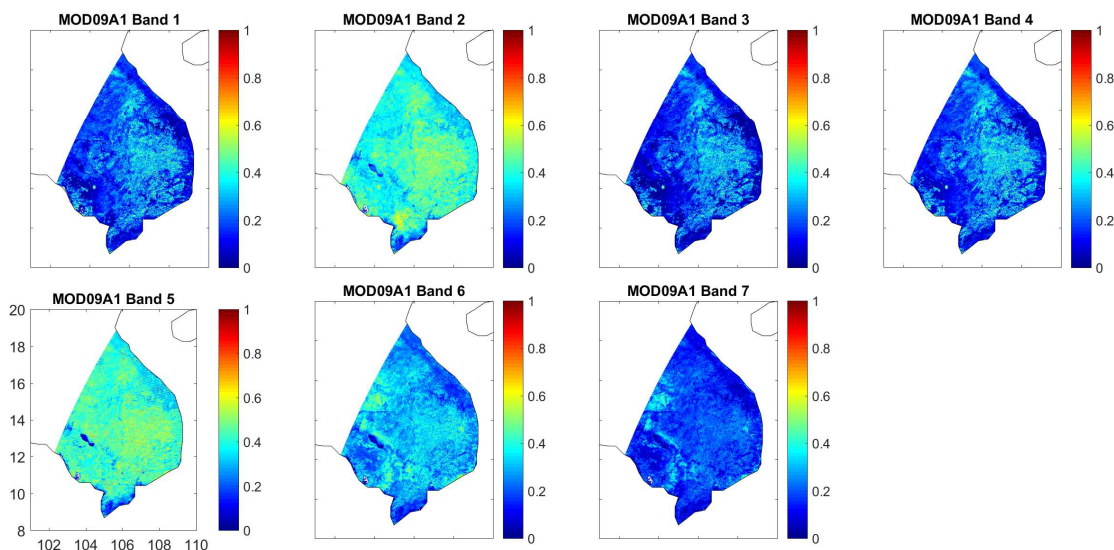


FIGURE 2.4: MOD09A1 land surface reflectance maps from the seven MODIS bands, over Cambodia and Vietnam in January, 2007.

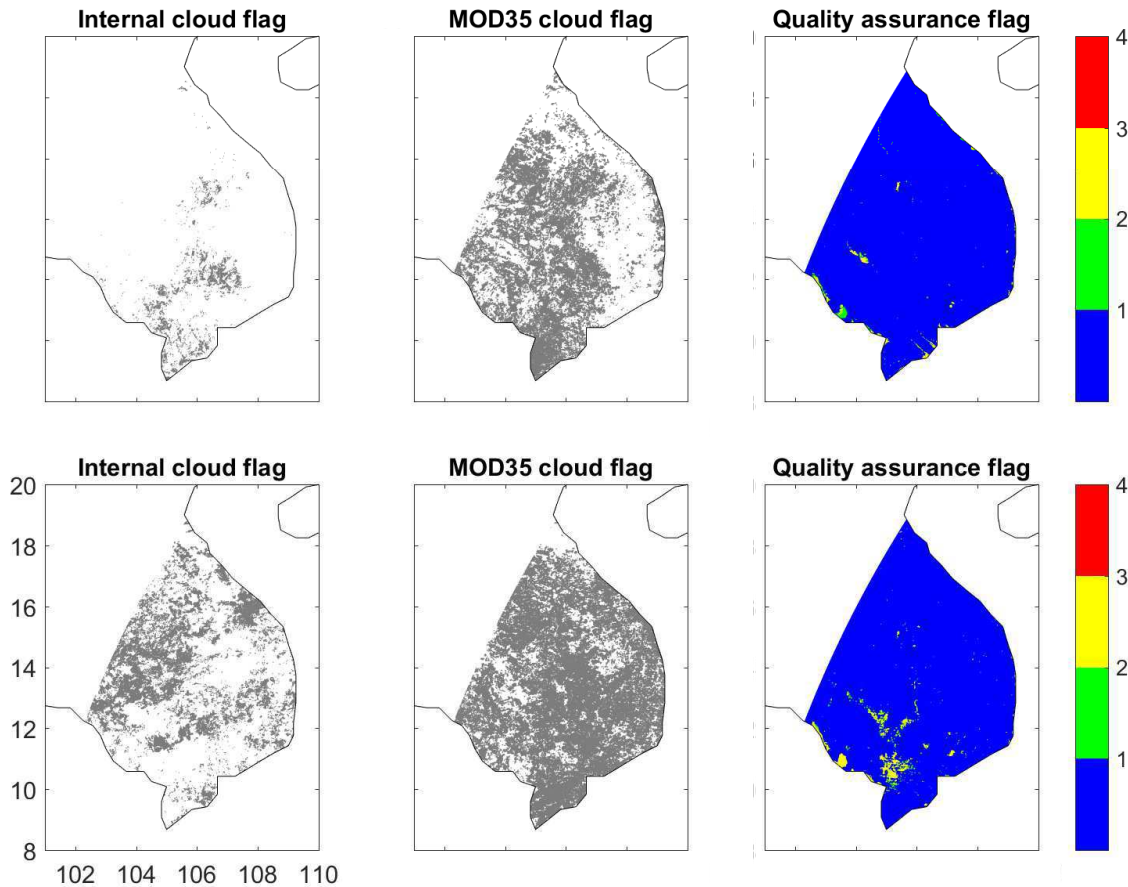


FIGURE 2.5: MOD09A1 cloud covered maps derived from the two cloud flags (internal cloud flag and MOD35 cloud flag), and quality assurance maps over Cambodia and Vietnam in January (top) and in October (bottom), 2007.

maps derived from the two cloud flags, and quality assurance maps over Cambodia and Vietnam in January (top) and in October (bottom), 2007. It is clear that the internal cloud algorithm map is totally different, and covers much less than the one derived from the MOD35 cloud flag for both seasons. The quality assurance flags are supposed to be very good (even during the wet season in October) since they are mostly marked as 1, meaning that it is corrected at ideal quality for all the seven MODIS bands. Relationships between the seven MOD09A1 surface reflectance bands, the two cloud flags, and the quality assurance flags are assessed by analyzing time series of these variables for four different pixels during the year 2007 (see Figure 2.6). It confirms that information from the two cloud flags are totally different for all the four pixels since the same pixels are cloud covered at different times in the year. Quality assurance flags seem to be good for

most of the time during the year. However, the times when the quality assurance flags are not good do not occur at the same time when these pixels are marked as cloud covered. Clearly, there is no strong connection between the quality assurance flags and these two cloud flags. The same conclusion can be made for the seven surface reflectance bands and the two cloud flags since the peaks of surface reflectance do not occur at the same time when these pixels are marked as cloud covered. For example, for the second pixel that is located in the middle of the ocean, time series of these surface reflectance bands show two straight peaks in July and August. It is believed that these peaks are due to the presence of clouds, but it is not confirmed by any cloud flag.

To conclude, there is no clear relationship between the seven MOD09A1 surface reflectance bands, the two cloud flags, and the quality assurance flag over the Mekong River Delta and Cambodia. This hypothesis is confirmed by Crétaux and Soudani (personal communication). In addition, surface reflectance values of all the seven bands are not very smooth and they must be filtered before being used in the next steps. The blue band surface reflectance values (band 3) will be used to detect cloud covered pixels using the Sakamoto et al. [2007] criterion (greater than 0.2).

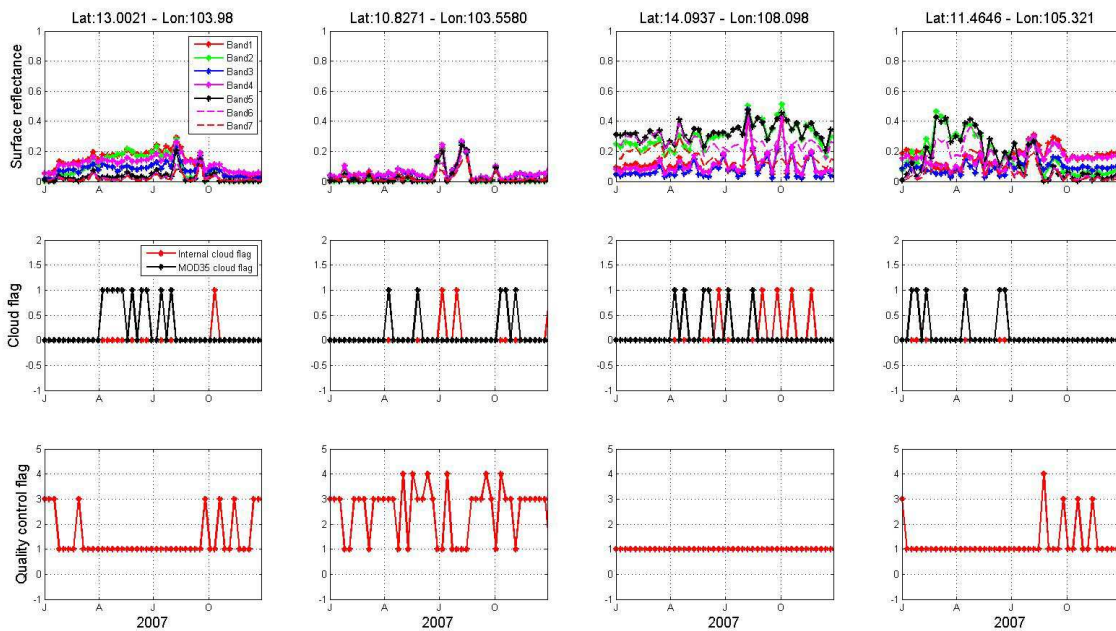


FIGURE 2.6: Time series of the seven surface reflectance bands, the cloud flags and the quality assurance flags for four selected pixels.

### 2.3.2 Smoothing method selection

To reproduce surface water maps based on the Sakamoto et al. [2007] methodology, it is required to use smoothed indices to reduce noise on the classification methodology. Three smoothing methods (simple moving average, weighted moving average, and Gaussian smoothing), with the same window size (5), are tested for the area shown in Figure 2.4. Surface water time series (from the original Sakamoto work) are used as the reference. Comparisons between the Sakamoto reference and results derived from the three smoothing methods are shown in Figure 2.7 and Table 2.5. In general, results derived from all the three smoothing methods are close and similar to the reference data. The weighted moving average method always gives the lowest temporal correlation to the Sakamoto reference data. The Gaussian and simple moving average smoothing give higher temporal correlations with the Sakamoto reference (greater than 94%), but the simple moving average smoothing provides higher correlations to the reference than the Gaussian smoothing. As the simple moving average smoothing fits better to the reference, this method is selected to smooth surface reflectance data and indices (EVI, LSWI, and DVEL, for example) before moving to the classifying step.

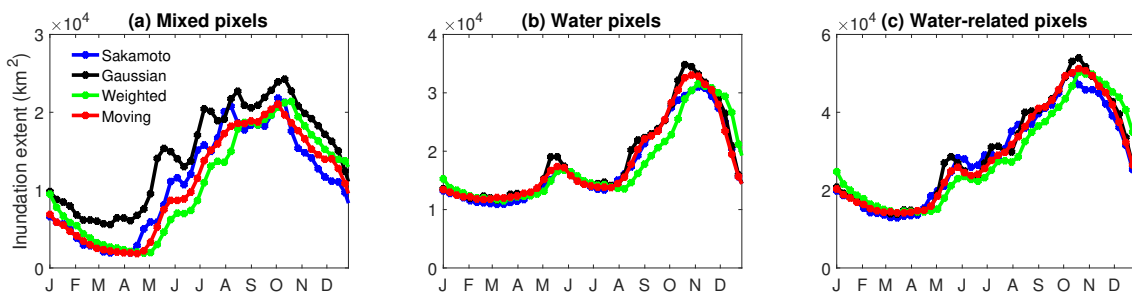


FIGURE 2.7: Surface water time series (2007) derived from three smoothing methods for the area shown in Figure 2.4, and comparisons with the Sakamoto results.

TABLE 2.5: Linear temporal correlations between surface water time series (2007) derived from the three smoothing methods and the Sakamoto reference data.

	Weighted	Gaussian	Moving
<b>Mixed pixels</b>	88%	94%	96%
<b>Water pixels</b>	93%	97%	99%
<b>Water-related pixels</b>	92%	96%	98%

## 2.4 Results

### 2.4.1 Comparing Sakamoto et al. [2007] and Bergé-Nguyen and Crétaux [2015] methodologies

In this section, comparisons between surface water maps and surface water time series for the year 2007 over Cambodia the Mekong River Delta, derived from [Bergé-Nguyen and Crétaux \[2015\]](#) and [Sakamoto et al. \[2007\]](#) methodologies, are discussed. A few changes are applied for the two classification methods. For [Bergé-Nguyen and Crétaux \[2015\]](#) method, only three classes (open water pixels, mixed pixels, and non-water related pixels) are classified instead of five classes as in the original work. It is assumed that vegetation pixels and dry land pixels belong to the non-water related class, and the aquatic vegetation pixels belong to the mixed-pixel class. Band 5 surface reflectance and NDVI data are also smoothed using the simple moving method. For the [Sakamoto et al. \[2007\]](#) method, the wavelet transformation for noise reduction is not applied. The original results from [Sakamoto et al. \[2007\]](#) for the year 2007 are used as the reference, but it has a spatial resolution of 230 m. Therefore, the original images have been re-gridded to a common spatial resolution of 500 m by averaging every 230 m pixels within each 500 m pixel, to match the other surface water maps.

Figure 2.8 shows the surface water maps over the Vietnam Mekong Delta and Cambodia, derived from the three methods in January (top) and October (bottom) 2007, respectively. During the dry season, [Bergé-Nguyen and Crétaux \[2015\]](#) method detects more mixed pixels than the [Sakamoto et al. \[2007\]](#) one, and most of the mixed pixels are not located around the Tonle Sap Lake or the Mekong river. During the rainy season, [Bergé-Nguyen and Crétaux \[2015\]](#) method marks most of pixels located around the Tonle Sap Lake and the lower Vietnam Mekong delta as fully inundated pixels. Surface water maps derived from the modification of the [Sakamoto et al. \[2007\]](#) method and from their original work have a strong agreement, especially for the water pixels. Linear spatial correlations of the surface water maps based on these classification methods are shown in Table 2.6. As expected, the spatial correlations between surface water maps derived from the original and modified [Sakamoto et al. \[2007\]](#) methods are very high for both the

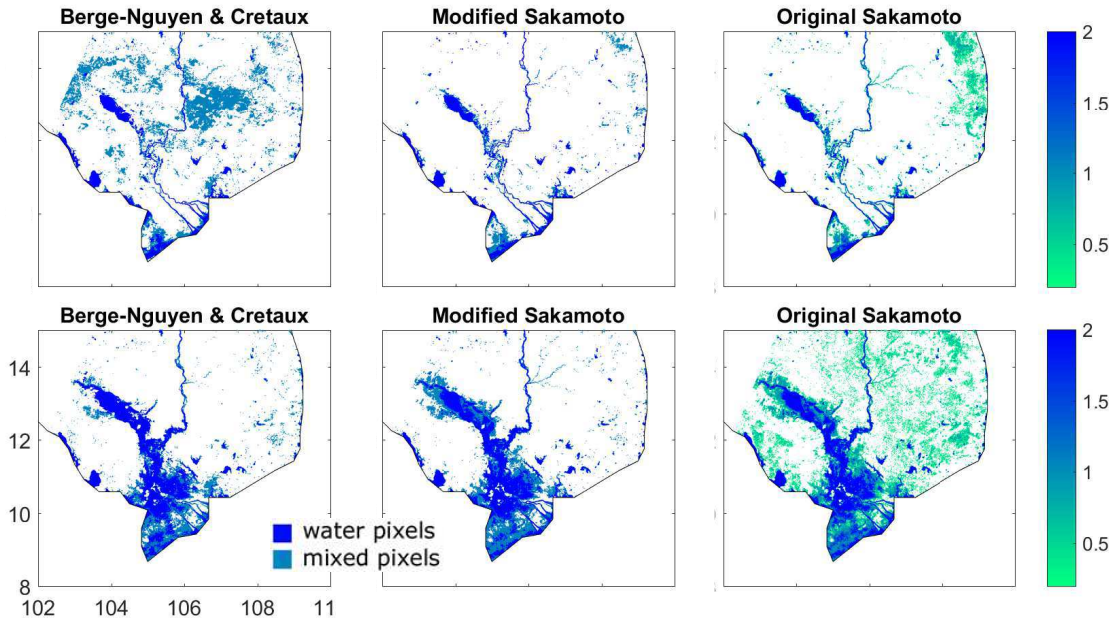


FIGURE 2.8: Surface water maps over the Vietnam Mekong Delta and Cambodia derived from [Bergé-Nguyen and Crétaux \[2015\]](#) (left), modified [Sakamoto et al. \[2007\]](#) (middle) and original [Sakamoto et al. \[2007\]](#) (right) methods in January (top) and October (bottom), 2007. Original Sakamoto surface water maps are re-gridded from 230 m to 500 m spatial resolution.

TABLE 2.6: Spatial correlations between surface water maps shown in Figure 2.8.

	Dry season	Rainy season
<b>Bergé-Nguyen &amp; Crétaux/Modified Sakamoto</b>	76%	94%
<b>Bergé-Nguyen &amp; Crétaux/Original Sakamoto</b>	78%	95%
<b>Modified Sakamoto/Original Sakamoto</b>	95%	96%

dry and wet seasons (always greater than 95%), while the spatial correlations of surface water maps derived from the [Bergé-Nguyen and Crétaux \[2015\]](#) method and the two other surface water maps are lower in the dry seasons (around 77%) and it increases during the rainy seasons (to 94% and 95%, respectively).

Figure 2.9 shows surface water time series for the area shown in Figure 2.8, derived from the three classification methods, for the year 2007. The [Bergé-Nguyen and Crétaux \[2015\]](#) method fails to capture the dynamics of mixed pixels during the dry season when it shows an increase and a high peak of surface water in February while it is decreasing with the reference data (Figure 2.9a). This is mainly because of the wrong detection of mixed pixels over local areas (mostly located in the northwest of the Tonle Sap Lake, and the east of the Mekong River).

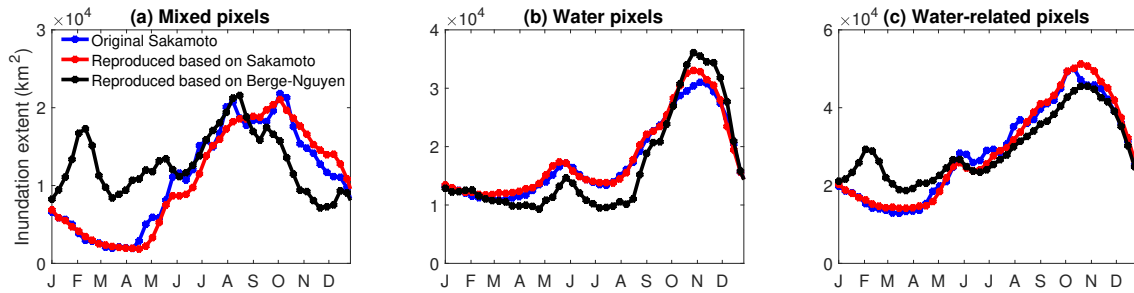


FIGURE 2.9: Surface water extent time series over the Vietnam Mekong Delta and Cambodia, derived from three classification methods for 2007.

For the water pixels, all classification methods have a similar seasonal cycles with a small peak in June and a larger one in November, but the [Bergé-Nguyen and Crétaux \[2015\]](#) method detects less water pixels than the [Sakamoto et al. \[2007\]](#) method most of the time, except in the middle of the wet season (Figure 2.9b). To conclude, classification results derived from the [Bergé-Nguyen and Crétaux \[2015\]](#) and the [Sakamoto et al. \[2007\]](#) classifications methods show similar seasonal cycle with higher spatial and temporal correlations for water pixels and lower for mixed pixels, as expected. The [Sakamoto et al. \[2007\]](#) algorithm seems to be better and more suitable than the [Bergé-Nguyen and Crétaux \[2015\]](#) one, for surface water and flood detection over the Vietnamese Mekong delta and Cambodia. The [Bergé-Nguyen and Crétaux \[2015\]](#) method does not work very well in this region because its simple algorithm was originally developed to detect water in arid and semi-arid regions.

## 2.4.2 Comparing unsupervised and supervised methodologies

In this section, surface water maps derived from supervised and unsupervised classification techniques (section 2.2) are compared to results obtained when applying the modified Sakamoto method. Figure 2.10 shows surface water maps in January (top) and October (bottom) 2007, while Figure 2.11 presents surface water extent time series derived from unsupervised, supervised and modified Sakamoto methods, respectively. Both unsupervised and supervised classifications detect more water-related pixels than Sakamoto method, but the surface water dynamic is very similar. Supervised classification uses the training dataset based on the Sakamoto criterion, therefore, surface maps and water extent time

series derived from these two methodologies are very close, with high spatial and temporal correlations. Table 2.7 shows confusion matrix between surface water maps derived from supervised classification and Sakamoto method for the high-water period (Figure 2.10 - bottom). The supervised classification detects correctly  $\sim 95\%$  water pixels and  $\sim 98\%$  non-water pixels compared to Sakamoto results.

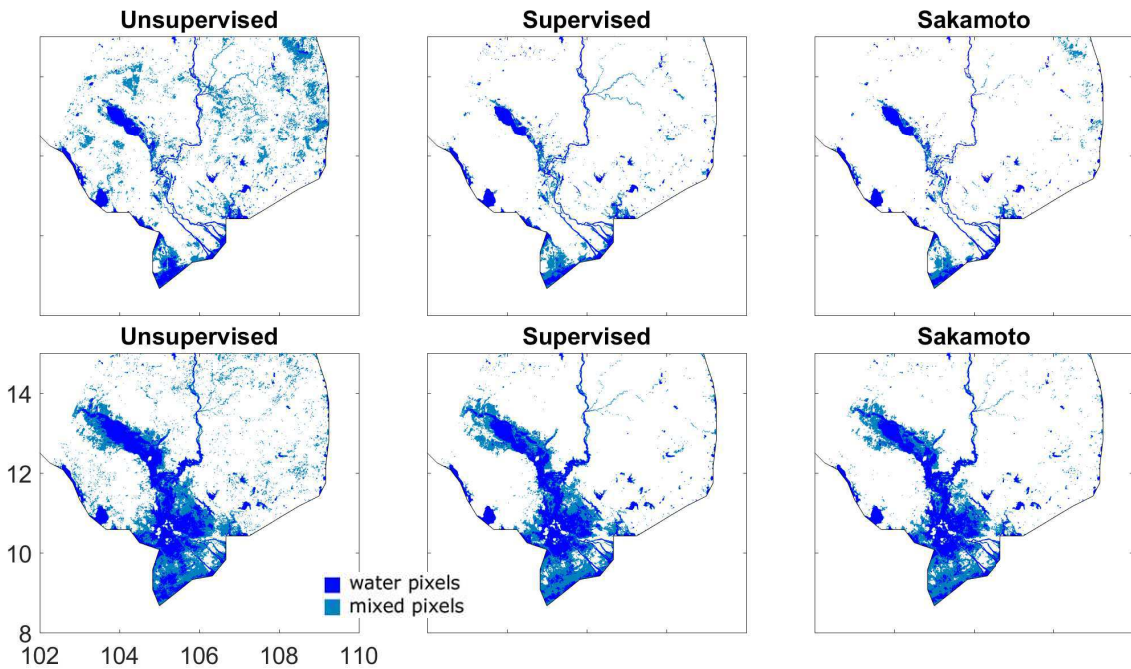


FIGURE 2.10: Surface water maps derived from (left) unsupervised classification, (middle) supervised classification and (right) modified Sakamoto et al. [2007] methods for (top) the dry season and (bottom) the rainy season of the year 2007, over the Mekong River Delta and Cambodia.

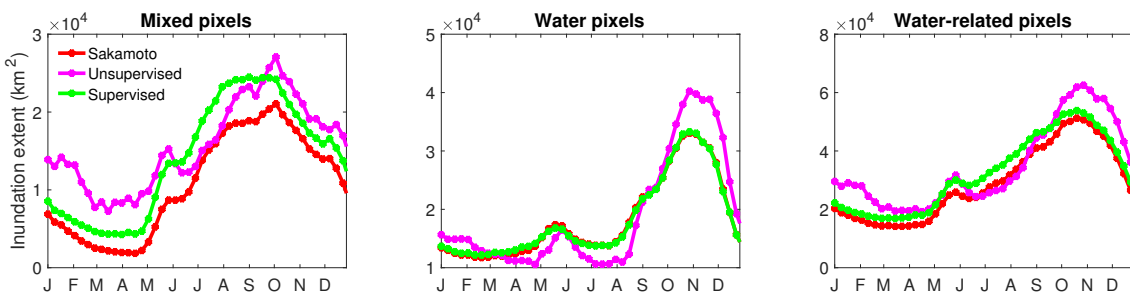


FIGURE 2.11: Surface water extent time series for the year 2007 over the area in Figure 2.8 derived from (left) unsupervised, (middle) supervised, and (right) modified Sakamoto classification methods.



TABLE 2.7: Confusion matrix of inundated maps derived from supervised classification and Sakamoto method for the high-water period.

		Supervised		
		Mixed	Non-water	Water
Sakamoto	Mixed	77.2%	0.4%	1.1%
	Non-water	1.7%	97.8%	0%
	Water	0.99%	0%	94.6%

A surface reflectance curve of a mixed environment (with water, submerged and emergent vegetation) derived from the SAILHFlood radiative transfer model [Beget et al., 2013], is plotted in Figure 2.12, along with the bary-centers of mixed-pixel clusters derived from the unsupervised and supervised techniques. The mixed-pixel bary-centers are mostly close to that from the SAILHFlood model, except for band 1 (the red band, at  $\sim 0.6 \mu\text{m}$ ). Differences may come from different conditions between a real mixing complex environment and a simpler mixing environment created by the SAILHFlood model.

Although results from the unsupervised and supervised classification methods are close and share similar dynamics, and in good agreement with the SAILHFlood model, there are still differences in term of the total surface water extent detected by each method, especially for the mixed pixels. Dominated bands for

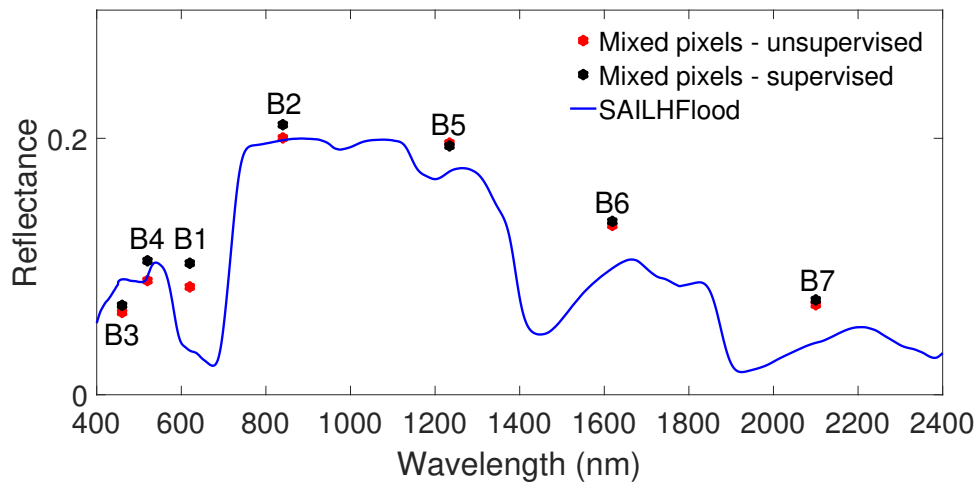


FIGURE 2.12: Seven MODIS band bary-centers of mixed-pixel cluster derived from unsupervised and supervised classifications, along with surface reflectance curve for a mixing environment derived from the SAILHFlood radiative transfer model [Beget et al., 2013].

the two classification methods are bands 1, 2, 6 and 7 because bands 1, 3 and 4 are highly correlated, as well as between bands 2 and 5. In this section, using the unsupervised and supervised classification methods for surface water detection is like a training exercise, and results are used to confirm performance of other classification methods.

### 2.4.3 Expanding Sakamoto methodology

After comparing surface water maps and surface water extent time series derived from two different classification methods (Sakamoto et al. [2007] and Bergé-Nguyen and Crétaux [2015]), along with tests with an unsupervised and a supervised classifications, it is concluded that the Sakamoto et al. [2007] algorithm is suitable for flood and surface water detection over the Vietnamese Mekong Delta and Cambodia using visible/infrared MODIS/Terra MOD09A1 products. Sakamoto et al. [2007] method is expanded to build surface water maps for the 2001-2016 period. Figure 2.13 shows new surface water extent time series (2001-2016), in comparison with the original data from Sakamoto (2001-2007) for the area shown in Figure 2.10. The seasonal dynamics and variability of the new surface water estimate is very close to the original one for the common period (2001-2007), with very high temporal correlations (0.95, 0.99 and 0.98 for mixed pixels, water pixels and water-related pixels, respectively). In addition, the new mixed-pixel time series is smoother than the original data. This 16-year surface water

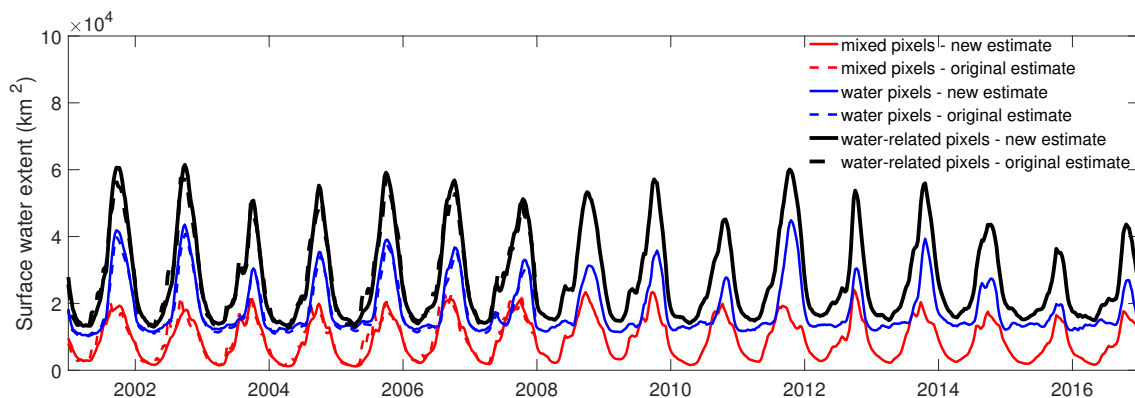


FIGURE 2.13: Surface water extent time series for the 2001-2016 period over the Mekong River Delta shown in Figure 2.8.

extent record allows us to identify recent drought years (2010, 2014 and 2015, for example) when the total surface water extents were low.

The Sakamoto et al. [2007] method is continuously expanded to the north-west to make surface water maps for a larger area. Tests have been made, but this classification method does not work well over the Tibetan Plateau with the presence of the Himalaya mountains because this method cannot distinguish between water pixels and snow/ice pixels. Open water bodies and snow/ice have similar signatures to the EVI that is the main indicator for this method. Finally, this method is applied for the area lower than latitude 20°N. An example of surface water maps for that area is shown in Figure 2.14-left, and the corresponding surface water extent time series (2001-2016) is presented in Figure 2.14-right.

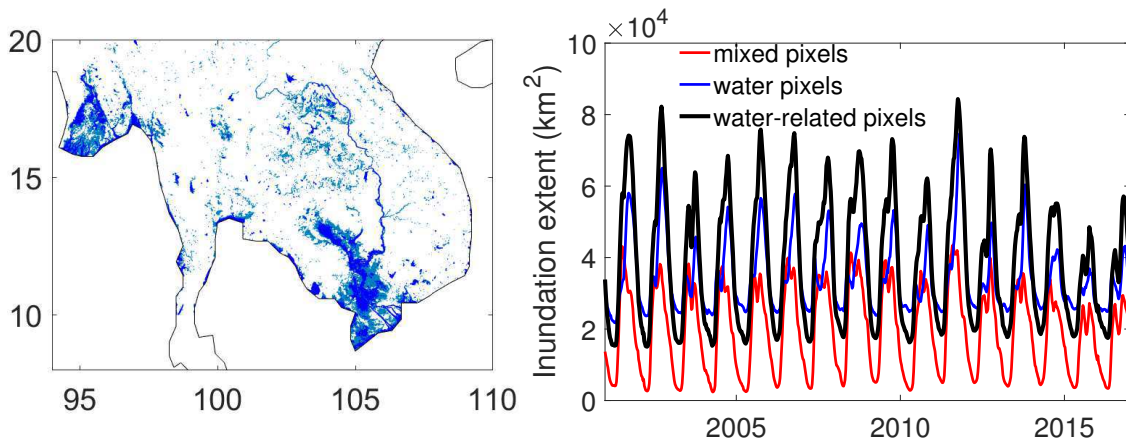


FIGURE 2.14: Left: An example of the extended surface water map. Right: Surface water time series (2001-2016) for area shown on the left.

## 2.5 Discussions and conclusions

As all visible/infrared satellite measurements, MODIS/Terra surface reflectance observations (MOD09A1) have limitations due to their inability to penetrate clouds and vegetation. Over the Vietnamese Mekong River Delta and Cambodia, information provided from the two MODIS cloud flags are not very useful. Surface reflectance data of band 3 (blue band) are used to identify cloud-covered pixels (values greater than 0.2), and a linear interpolation is applied to fill the missing

data. A simple moving smoothing method is used to reduce noise before the classifying step. Four classification methods have been applied to MOD09A1 data to create and compare surface water maps over the Vietnamese Mekong Delta and Cambodia, for 2007. The surface water maps and time series derived from all the four methods are comparable, showing similar seasonality and dynamics. Output surface water maps show significant differences in detecting mixed pixels. Sakamoto et al. [2007] classification method is chosen to be applied to a longer period (2001-present) because this method has been specifically designed to detect flood and to monitor surface water extent over tropical regions like the Lower Mekong River Delta. As already observed in Figure 2.13, the new surface water extent is in a strong agreement with the original Sakamoto et al. [2007] estimate, for the 2001-2007 period.

For the larger area shown in Figure 2.14-left, the result surface water map looks reasonable. The Sakamoto et al. [2007] classification algorithm can detect the Mekong River, smaller lakes, reservoirs and a big inundated area over the Gulf of Martaban in the south of Myanmar. This region has a strong wetland seasonal dynamics as shown in Figure 2.14-right. Maximum states of mixed pixels (red curve) and fully inundated pixels (blue curve) do not always occur at the same time. At the time of this study, there is no other available inundated product with similar spatial resolution for this large area over the same period for comparison.

MODIS surface water maps in a regular basis (every 8 days) at 500 m spatial resolution over the Vietnamese Mekong Delta and Cambodia can be a helpful product for scientific community. It could be an input for regional hydrological models to study the impacts of the climate change to the Lower Mekong basin. Regular MODIS surface water maps (from 2001 to present time) over the studied region are stored on the University of Science and Technology of Hanoi (USTH) Space and Aeronautics Department website, and they are provided freely to final users. Data are updated regularly.

English version: <http://space.usth.edu.vn/en/news/projects/>

Vietnamese version: <http://space.usth.edu.vn/vi/news/du-an/>



## **Chapter 3**

# **Surface Water Monitoring within the Mekong Delta and Cambodia using SAR Sentinel-1 Satellite Observations**



## 3.1 Introduction

In this chapter, the use of Synthetic Aperture Radar (SAR) satellite observations for surface water monitoring will be studied. The use of SAR satellites started in 1978 with the launch of the Seasat, the first civil SAR satellite. Since then the SAR instruments have been available on many sensors and platforms with different frequencies, polarizations, swaths and resolutions (see Table 3.1 for a list of some recent SAR satellite instruments). SAR images have become an important source of data for a wide range of applications in remote sensing to map the Earth surfaces (for topography, oceanography, interferometry, glaciology, geology, forestry, volcano, earthquake, flood and surface water monitoring, for example). SAR remote sensing has some basic advantages compared to optical remote sensing. SAR sensors can provide observations in all weather conditions regardless of cloud presence, with spatial resolution comparable to optical satellite images. They are also independent from the sunlight, providing day and night operation that will be very useful in some urgent situations. In addition, SAR observations are not affected by atmospheric constituents. These advantages make SAR observations more suitable than visible/infrared observations for many applications, especially in tropical regions.

Flood detection using different SAR observations has been studied by many authors, showcasing the advantages of SAR instruments compared to optical instruments. Wang [2004] used a single decision tree classifier on two sets of JERS-1 SAR data to classify surface water within the states of North Carolina and South Carolina into five land cover types (water, marsh, flooded forest, field, and non-flooded forest). Although the classifier was simple, they reported an overall classification accuracy of nearly 90%. Pierdicca et al. [2013] showed the potential of the COSMO-SkyMed data for flood detection by showing case studies in several locations all over the globe (e.g., Tarano River overflow, Italy, April 2009; Pakistan inundation, July–September 2010; Thailand flood, October 2010; and Australia flood, January 2011). COSMO-SkyMed instruments provided very high resolution X-band SAR images, but covered limited areas (the highest spatial resolution is  $\sim 1$  m for an observation area of  $10 \text{ km} \times 10 \text{ km}$ ). X-band data from



TerraSAR-X instrument were also reported suitable for flood mapping under forest canopy in a temperate forest zone in Estonia [Voormansik et al., 2014]. Martinis et al. [2015] compared four flood detection approaches over several areas (the Netherlands, Mali, Germany, and China) using SAR data from the TanDEM-X mission. Although these four approaches were designed according to different requirements, their performances were satisfactory over the studied areas (17 out of 20 water masks reaching an overall accuracy larger than 90%). Other studies using SAR data for water monitoring locally and regionally under different environments can be listed (e.g., Bartsch et al. [2008]; Brisco et al. [2009]; Reschke et al. [2012]). There are some studies that used SAR observations to monitor surface water, focusing on the Vietnam Mekong Delta. Nguyen and Bui [2001] mapped flood occurrence for the year 1996 over the Delta using five ERS-2 observations. Kuenzer et al. [2013] used 60 Envisat ASAR observations during the years 2007–2011 to study the flood regime in the Delta. Mapping water bodies at global scale using SAR data was limited due to the lack of global observations, and the fact that SAR data were not easy to access freely. Santoro et al. [2015a] used multi-year (2005–2012) Envisat ASAR observations to create, for the first time, a global potential water body map at a spatial resolution of 150 m from SAR data. Errors concentrated along shorelines and coastline, but this global water map has an accuracy of ~80% compared to reference data.

Since 2014, the free and open data access policy of the European Space Agency (ESA) within the Copernicus space program makes SAR Sentinel-1 observation now regularly and freely accessible for scientific and educational purposes, over large parts of the globe. Similar to previous SAR instruments, Sentinel-1 instruments show strong potential for detecting open water bodies at high spatial resolution [Amitrano et al., 2014; Santoro et al., 2015b]. With the advantage of higher temporal resolution than previous SAR instruments, Sentinel-1 has the ability to monitor the seasonal cycle of water extent every six days over Europe and the boreal region, and with slightly reduced temporal sampling elsewhere. In this study, I propose a methodology using Sentinel-1A SAR observation for monitoring water surface extent within the Mekong Delta and Cambodia since the beginning of 2015. It is based on a Neural Network (NN) algorithm, trained on visible Landsat-8 images (30 m spatial resolution). The Sentinel-1 SAR data and the ancillary observations are described in Section 3.2, including the pre-processing steps.

TABLE 3.1: List of recent SAR instruments on board satellites. Sources: NASA, ESA & JAXA.

Satellites	Agency	Freq.-Pol.	Resolution-Swath
ERS-1 (1991-2000)	ESA	C-VV	25 m - 100 km
JERS (1992-1998)	JAXA	L-HH	20 m - 100 km
ERS-2 (1995-2011)	ESA	C-VV	25 m - 100 km
RADARSAT-1 (1995-2013)	CSA	C-HH	10-100 m - 45-500 km
ENVISAT (2002-2012)	ESA	C-HH/VV/HV	25-1000 m - 50-500 km
ALOS-PALSAR (2006-2011)	JAXA	L Polarimetric	10-100 m - 100-350 km
Terra-X (2007-present)	DLR	X Polarimetric	HR SpotLight: 1 m-10×5 km SpotLight: 2 m-10×10 km StripMap: 3 m-30×50 km ScanSAR: 18 m-100×150 km
RADARSAT-2 (2007-present)	CSA	C Polarimetric	UltraFine: 3 m - 20 km Standard: 25 m - 100 km ScanSAR: 100 m - 500 km
ALOS-PALSAR-2 (2013-present)	JAXA	L Polarimetric	SpotLight: 3×1 m - 25×25 km Stripmap: 3-10 m - 30-70 km ScanSAR: 60,100 m - 350,490 km
Sentinel-1 (2013-present)	ESA	C-HH/VV/VH/HV	StripMode: 5 m - 80 km Interfero WS: 5×20 m - 250 km Extra WS: 20×40 m - 400 km

Section 3.3 presents the NN methodology, along with sensitivity tests. Results and comparisons with other products are provided and discussed in Section 3.4. An improvement of the NN methodology is presented in Section 3.5. Section 3.6 concludes this study.

## 3.2 Sentinel-1 SAR data and the ancillary datasets

### 3.2.1 Sentinel-1 SAR data

Sentinel-1 is a satellite project funded by the European Union and carried out by the European Space Agency. It is a two satellite constellation working at C-band (5.405 GHz). The major objective of the satellites is the observation and monitoring of land and ocean surfaces day and night, under all weather conditions [ESA, 2015]. The satellite operates in four exclusive imaging modes with different spatial resolutions (the highest being 5 m) and swaths (up to 400 km). The first Sentinel-1A satellite of the pair was launched on 3 April 2014, and the second Sentinel-1B satellite was launched on 22 April 2016. The Sentinel-1 satellites fly along a sun-synchronous, near-polar circular orbit at an altitude of  $\sim 693$  km. The incidence angle varies between  $29^\circ$  and  $46^\circ$ . The two satellites provide a revisiting time of 6 days (it was 12 days before the launch of the Sentinel-1B satellite). Sentinel-1 satellites have dual polarization capabilities (HH, VV, HH + HV and VV + VH), giving final users the ability to access a large variety of applications, including the monitoring of surface water. SAR images from Sentinel-1 satellites are freely downloaded from the Sentinel scientific data hub (<https://scihub.copernicus.eu/>).

20 m resolution (10 m pixel spacing) Level-1 Ground Range Detected (GRD) Sentinel-1 images from the Interferometric WideSwath (IW) mode are used in this study. These images have been detected and projected to ground range using an Earth ellipsoid model provided by ESA. Over the Mekong Delta and Cambodia, there are two polarizations available: the VH and VV polarizations. Some pre-processing steps have to be carried out using the free Sentinel Application Platform (SNAP) software developed by ESA, before the analysis steps (Figure 3.1). These pre-processing steps are described in the “SAR Basics with the Sentinel-1 Toolbox in SNAP tutorial” (<http://step.esa.int/main/doc/tutorials/>).

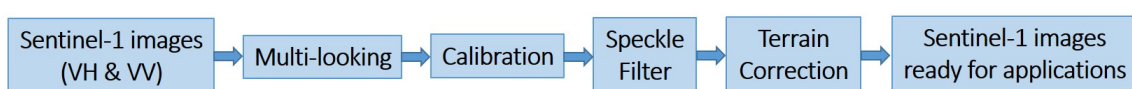


FIGURE 3.1: Synthetic Aperture Radar (SAR) Sentinel-1 pre-processing steps.

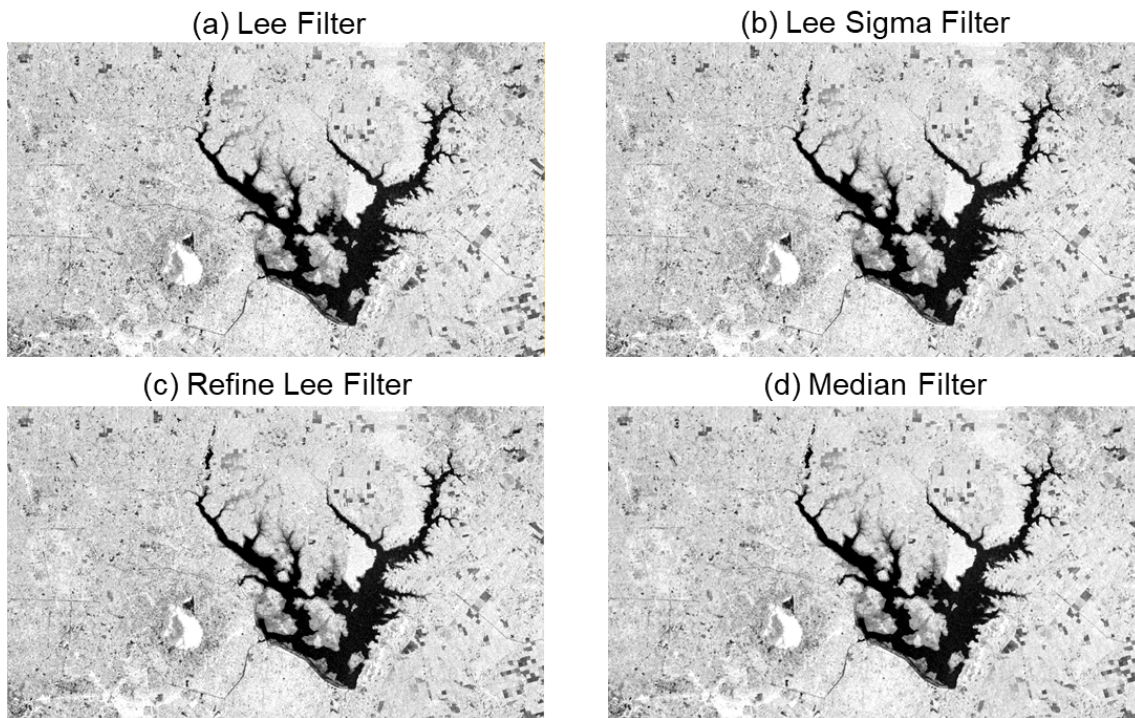


FIGURE 3.2: Sentinel-1 backscatter coefficient VH polarization over a lake derived from four different filters. The image was acquired on 10 May 2015.

First, multi-looking processing is applied to each single Sentinel-1 image (both polarizations) to convert to 30 m spatial resolution (to match with Landsat-8 images). Applying multi-looking at the beginning of the chain reduces the processing time for the next steps since the size of the image is several times smaller than the original one. Second, the image is calibrated to convert values of the raw image from digital number to radar backscatter coefficient ( $\sigma_0$ ). Third, the Refined Lee filter is applied to reduce the speckle noise and to smooth the radar backscatter coefficient data because this filter maintains details of the standing water boundary [Liu, 2016]. Other filters (Lee, Lee Sigma or Median, for example) were tested, and results showed little differences in terms of water detection (see Figure 3.2). Next, the "terrain correction" tool is used to compensate for distortions in the SAR images, so that the geometric presentation of the image will be as close as possible to the real world. At the end of this step, the image is also re-projected from the satellite projection to the Earth geographic projection, and is ready for applications. To fully cover the Vietnamese Mekong Delta and Cambodia, at least five Sentinel-1 SAR images are needed. Figure 3.3 (top) provides

examples of the SAR backscatter coefficients for VH (a) and VV (b) polarizations, along with the incidence angle (c), over the Tonle Sap Lake, on 17 December 2015.

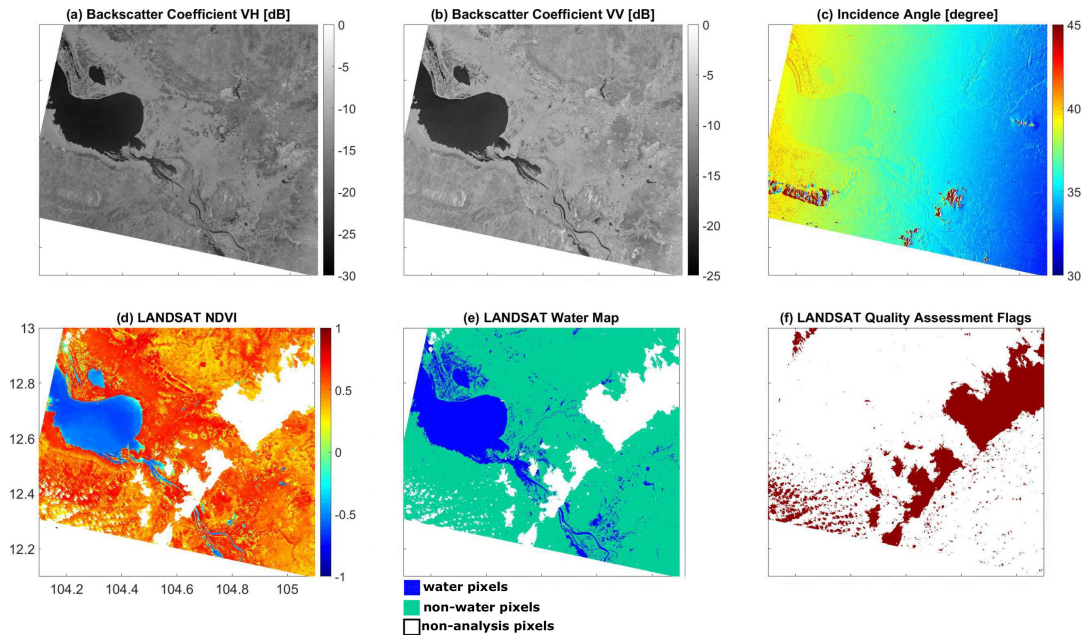


FIGURE 3.3: Examples of satellite observations from Sentinel-1 (top) and from Landsat-8 (bottom), over the southeast of the Tonle Sap Lake (Cambodia) after the pre-processing steps: (a) SAR backscatter coefficient at VH polarization; (b) SAR backscatter coefficient at VV polarization; (c) SAR incidence angle; (d) The Normalized Difference Vegetation Index (NDVI) from Landsat-8; (e) Surface water estimated from Landsat-8; and (f) Landsat-8 quality flags. The white areas are cloud-covered pixels detected by the Landsat quality flags, and have been removed. Both Sentinel-1 and Landsat-8 images were acquired on 17 December 2015.

## 3.2.2 Ancillary datasets

### 3.2.2.1 Inundation maps derived from Landsat-8 data

Landsat-8 satellite collects visible and shortwave images (30 m spatial resolution). NIR wavelength reflects less solar radiation than the red wavelength over water bodies [McFeeters, 1996; Xu, 2006], and surface water maps can be derived from the NDVI maps (water pixels and non-water pixels correspond to negative and positive values of NDVI, respectively) [Rouse Jr. et al., 1974; Rokni et al., 2014]. Other indices have been used to detect water, but the NDVI is effective

when properly corrected from the atmospheric contamination. In this study, official and reliable atmospherically corrected Landsat-8 NDVI images are ordered directly from the U.S. Geological Survey (USGS) website (<https://espa.cr.usgs.gov/index/>). To limit cloud effects, only images with less than 10% of cloud contamination are used. The selected images are further filtered using the Landsat-8 quality assessment to remove pixels that might be affected by instrument artifacts or subject to cloud contamination. Figure 3.3 (bottom) shows the NDVI from Landsat-8 (d), the resulting surface water map based on negative NDVI values (e), and the quality flag (f), for the same regions and the same day (17 December 2015) as previously presented. Over the Lower Mekong basin (lower than latitude number 15), there are ~250 Landsat-8 images available between January 2015 and January 2016. However, there is only ~10% (27 images) with less than 10% cloud contamination. Among the remaining images, only 1/3 was selected for this study since they were observed with a time difference of less than 3 days from a Sentinel-1 image.

### 3.2.2.2 Inundation maps derived from MODIS/Terra data

In this study, the MODIS/Terra surface reflectance products (MOD09A1) described in Chapter 2 are used to create surface water maps based on the Sakamoto methodology [Sakamoto et al., 2007]. MODIS surface water maps (500 m spatial resolution) over the Vietnam Mekong Delta and Cambodia will be compared to the corresponding surface water maps derived from SAR Sentinel-1 observations for 2015.

All Sentinel-1, Landsat-8 and MODIS/Terra observations used in this study are listed in Tables 3.2 and 3.3. Sentinel-1 and Landsat-8 training observations are used to train the NN (Section 3.2). Sentinel-1 and Landsat-8 test observations are used to test, optimize, and evaluate the performance of the NN (Sections 3.3.3 and 3.4.1). NN evaluation is also based on comparisons with MODIS surface water estimates (Section 3.4.3).

TABLE 3.2: List of 9 Sentinel-1 and corresponding Landsat-8 training (top) and test (bottom) observations used in this study over the Vietnamese Mekong Delta and Cambodia. Maximum gap between Sentinel-1 and Landsat-8 observations is only 3 days. The cloud cover percentage is indicated for each Landsat-8 observation.

<b>Sentinel-1 and Landsat-8 Training Observations</b>			
<b>Image No</b>	<b>Sentinel-1</b>	<b>Landsat-8</b>	<b>Clouds</b>
1	16 April 2015	14 April 2015	6.29%
2	21 April 2015	21 April 2015	0.05%
3	19 August 2015	18 August 2015	7.94%
4	17 December 2015	17 December 2015	4.84%
5	29 March 2016	31 March 2016	6.22%
6	9 June 2016	10 June 2016	3.94%
<b>Sentinel-1 and Landsat-8 Test Observations</b>			
<b>Image No</b>	<b>Sentinel-1</b>	<b>Landsat-8</b>	<b>Clouds</b>
1	5 January 2016	2 January 2016	0.16%
2	3 February 2016	3 February 2016	7.5%
3	22 February 2016	19 February 2016	0.29%

TABLE 3.3: List of 20 Sentinel-1 and corresponding MODIS/Terra observations used in this study over the Vietnamese Mekong Delta and Cambodia.

<b>Sentinel-1 and MODIS/Terra Observations</b>			
<b>Image No</b>	<b>Date</b>	<b>Image No</b>	<b>Date</b>
1	10 January 2015	11	14 August 2015
2	3 February 2015	12	26 August 2015
3	15 February 2015	13	7 September 2015
4	11 March 2015	14	19 September 2015
5	4 April 2015	15	1 October 2015
6	28 April 2015	16	13 October 2015
7	15 June 2015	17	25 October 2015
8	27 June 2015	18	6 November 2015
9	9 July 2015	19	30 November 2015
10	21 July 2015	20	24 December 2015

## 3.3 Methodology

### 3.3.1 Surface water information from the Sentinel-1 SAR images

Flat water surfaces act like mirrors and reflect almost all incoming energy in the specular direction, thus providing very low backscatter for large observation incidence angles. With this physical principle, detection of surface water is often based, at least partly, on the application of a threshold on the SAR backscatter coefficient, with the low backscatter values attributed to water bodies [Nguyen and Bui, 2001; Wang, 2004; Pierdicca et al., 2013; Kuenzer et al., 2013]. However, SAR backscatter coefficients over water surfaces are also affected by several mechanisms related to the interaction of the signal with vegetation or with possible surface roughness. The backscattered signals over flooded vegetation in wetlands can be enhanced due to the double-bounce scattering mechanism [Hess et al., 1990; Kasischke and Bourgeau-Chavez, 1997; Pope et al., 1997]. On the other side, the backscatter coefficients can be affected by vegetation canopy (e.g., rice) above the water surfaces due to volume scattering from the plant components (stems or leaves) [Liu et al., 2016b]. The backscatter coefficients (especially the VV polarization) can also be influenced by the wind-induced surface roughness over open water [Gstaiger et al., 2012; Kuenzer et al., 2013]. Finally, there might be ambiguities between surface water and other very flat surfaces (such as arid regions), that could provide very similar backscatter signatures [Prigent et al., 2015].

Based on a reference water mask derived from Landsat-8 NDVI, Figure 3.4 presents the histograms of the backscatter coefficients for VH and VV polarizations, separately for water and non-water pixels over the incidence angle range of 30°–45° for the area shown in Figure 3.3. For both polarizations, the water and non-water histograms are rather well separated, with thresholds of –22 dB and –15 dB for the VH and VV polarizations, respectively. Using these thresholds, the surface water has been classified separately for each polarization. The classification derived from the VH polarized image had a stronger spatial linear correlation with the reference water mask than the one derived from the VV polarized



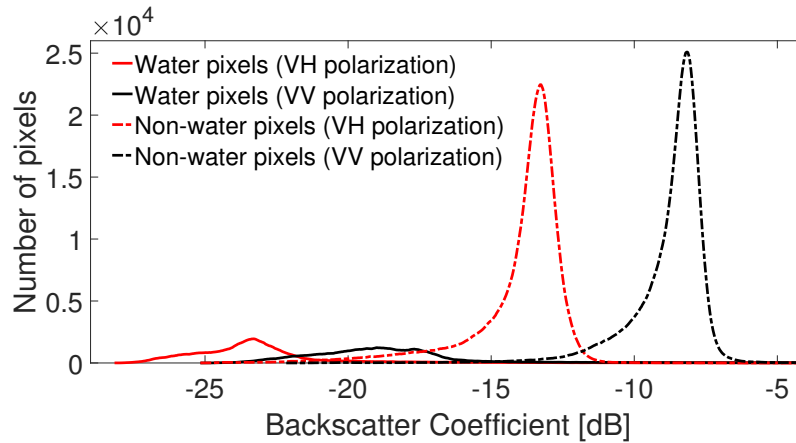


FIGURE 3.4: For surface water delineated with Landsat-8, histograms of the water and non-water pixels for the SAR backscatter coefficients in VH and VV polarizations for the area shown in Figure 3.3 (over the incidence angle range of 30° to 45°).

image (72% compared to 62%), confirming a higher sensitivity of the VH polarization to the presence of surface water [Santoro et al., 2015b]. Using both polarizations for the classification increased the correlation (76%), confirming that the two polarizations carry different information and that using both of them increases the retrieval accuracy. These findings confirmed the study by Henry et al. [2006] where water detection with VV polarization was further refined using multiple-polarizations.

The effect of the backscatter incidence angle is also tested here. For a collection of pixels located over water (rivers, reservoirs, or lakes), the backscatter coefficient is plotted as a function of the incidence angle between 30° and 45° (Figure 3.5). Similar negative correlations between incidence angle and backscatter coefficients can also be found in Santoro et al. [2015a] with ASAR data over water bodies (from  $\sim -5$  dB at 20° to  $\sim -20$  dB at 45° of incidence angle).

As a conclusion, the SAR backscatter coefficients (VH and VV polarizations) are both sensitive to the presence of water, but with slightly different sensitivities. The effect of the incidence angle, although rather limited within the 29°–46° range of Sentinel-1 SAR, has to be accounted for if a high detection accuracy is required. Simple tests on thresholding techniques illustrated the limitations of these approaches and here I suggest developing a new scheme to delineate the surface

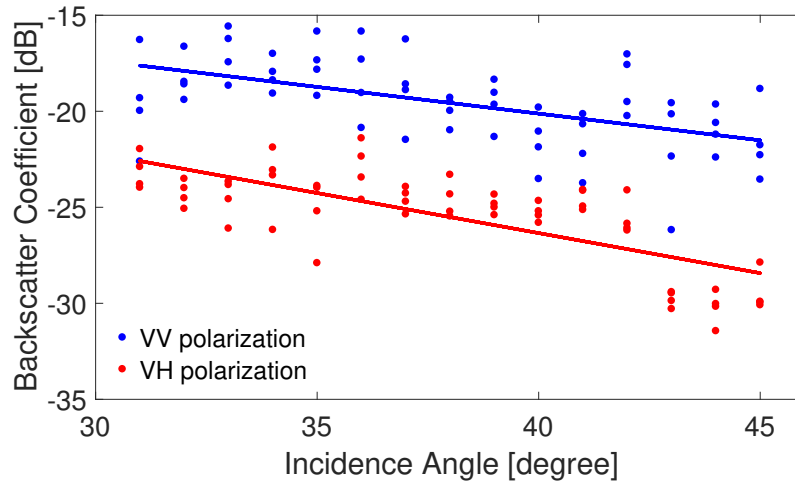


FIGURE 3.5: The SAR backscatter coefficients (VH and VV polarizations) from the Sentinel-1 as a function of the incidence angle over water bodies. The linear regression lines are also plotted.

water based on Neural Networks (NN). The temporal dynamics of the backscatter coefficients can also be a source of information and can help disentangle the influence of the other surface parameters [Santoro et al., 2015a]. This temporal information will be investigated and discussed in Section 3.5.

### 3.3.2 A Neural Network-based classification

Here, I propose training a NN to produce surface water maps from SAR Sentinel-1 observations, over the Vietnam Mekong Delta and Cambodia. In remote sensing, NNs are often used as a regression tool to estimate a quantity. For each pixel, NN input satellite observations are represented by a vector  $x$ , and the network outputs (i.e., the retrieval) is represented by a vector  $y$ . However, NNs can also be used as classifiers. In this case, when trained with binary output values ( $y = 0$  for non-water, 1 for water surfaces), the NN becomes a statistical model for the conditional probability  $y = P(\text{surface} = \text{water}/x)$ , i.e., the probability of the surface being covered by water knowing the satellite observations  $x$ . The NN output can then directly be used as an index for water presence probability, but a threshold can also be applied to classify the state as being covered by water or not. The threshold needs to be optimized in order to satisfy some quality criteria, such as overall accuracy or false alarm rates.

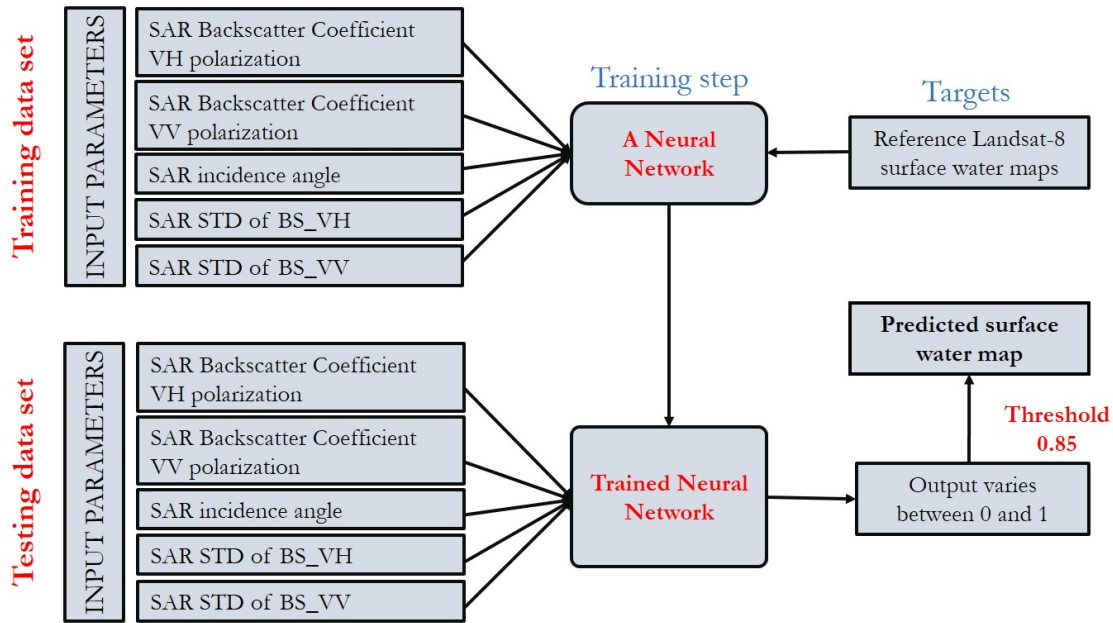


FIGURE 3.6: The block diagram of the proposed Neural Network (NN) algorithm.

The NN classifier needs to be trained in order to perform an optimal discrimination between water and non-water states. A supervised learning is chosen: the NN will be designed to reproduce an already existing classification. A dataset including a collection of SAR information  $x$  and associated surface water state  $y$  is first built. Part of it is then used during the training stage in order to determine the optimal parameters of the NN model. The reference dataset in the selected area is provided here by a Landsat-8 surface water map (NN outputs), in spatial and temporal coincidence with the SAR Sentinel-1 data (NN inputs). A maximum time difference of 3 days is tolerated, as the two satellites do not fly in phase. Six Landsat-8 surface water maps are selected, along with the corresponding Sentinel-1 observations (see Table 3.2 for details on the training dataset). The selection process for the Landsat-8 images has been described in Section 3.2.2.1. For each image in the training dataset, the number of non-water pixels is much higher than the number of water pixels. To avoid giving too much weight to the non-water pixels, an equalization of the training dataset is performed: an equal number of non-water and water pixels is selected in the training dataset. For this purpose, non-water pixels are selected randomly in the images, to match the

number of water pixels. The total number of training samples is  $\sim 10$  million pixels, half water pixels, half non-water pixels. It takes  $\sim 5$  h to train the NN (with the use of a personal computer), but when the training is completed, a surface water map can be produced quickly (after  $\sim 3$ – $4$  minutes) from any new set of satellite inputs  $x$ . A test dataset is chosen to measure the performance of the NN retrieval scheme with data not used in the training process. The NN methodology is summarized in Figure 3.6.

Several tests are necessary to determine the optimum inputs to the NN, in addition to the obvious ones, i.e., the backscatter coefficients for both polarizations. To limit ambiguities between flat arid surfaces and surface water, and to better capture small rivers, the spatial homogeneity of the backscatter coefficients appeared to be a relevant parameter. The standard deviations of the backscatter

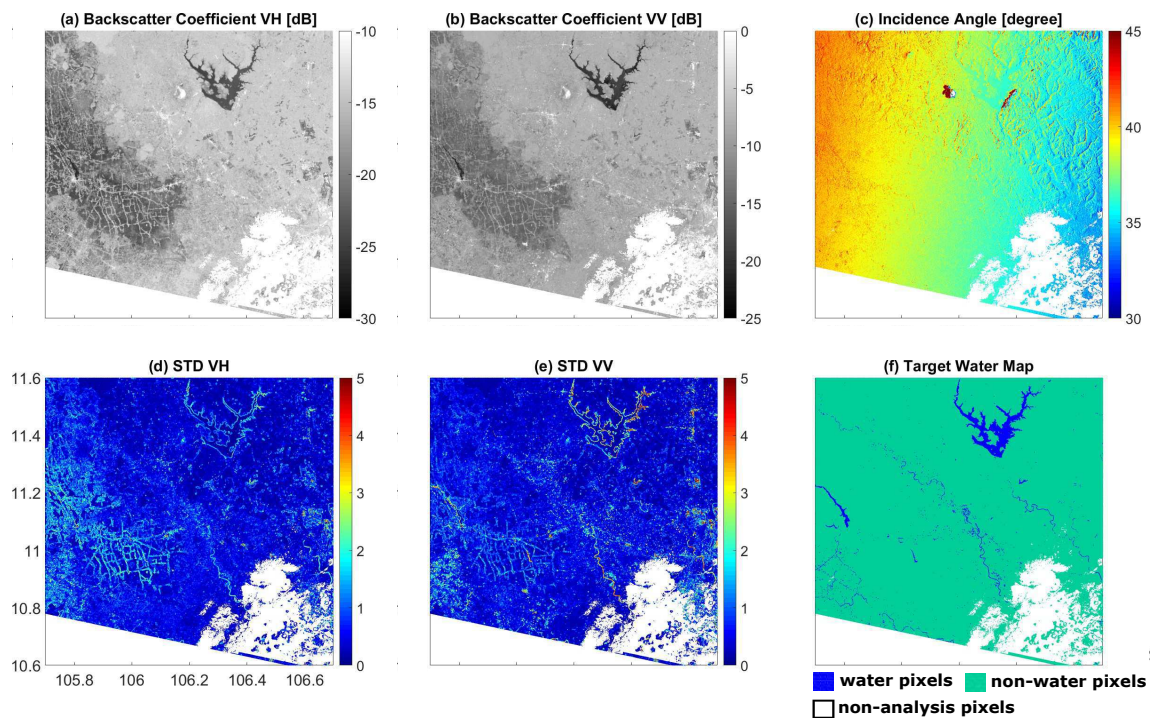


FIGURE 3.7: Example of the five inputs and the target for the NN. (a) SAR backscatter coefficient VH polarization; (b) SAR backscatter coefficient VV polarization; (c) SAR incidence angle; (d) SAR standard deviation of backscatter coefficient VH polarization; (e) SAR standard deviation of backscatter coefficient VV polarization; and (f) Target surface water map based on NDVI from Landsat-8. The white areas are cloud-covered pixels detected by the Landsat quality flags, and they have been removed. Sentinel-1 and Landsat-8 images were acquired on 16 and 14 April 2015, respectively.

coefficients are computed locally over  $150 \text{ m} \times 150 \text{ m}$  boxes. As a result, the NN uses five different inputs  $x$ :

- SAR backscatter coefficient VH polarization (BS\_VH);
- SAR backscatter coefficient VV polarization (BS\_VV);
- SAR standard deviation of backscatter coefficient VH over  $150 \text{ m} \times 150 \text{ m}$  (STD\_VH);
- SAR standard deviation of backscatter coefficient VV over  $150 \text{ m} \times 150 \text{ m}$  (STD\_VV);
- and SAR incidence angle.

Figure 3.7 presents an example of the set of five inputs and the target surface water map used to train the NN. Missing areas in the maps correspond to Landsat-8 low quality pixels and are excluded from the training. The NN model is asked to find a relationship between these five input parameters and the corresponding water and non-water state.

### 3.3.3 NN sensitivity tests

In this section, a test dataset of three SAR Sentinel-1 images and three corresponding Landsat-8 reference surface water maps is used to make several sensitivity tests in order to optimize the performance of the NN classification (see details of the test data sets in Table 3.2). Three different sensitivity tests are carried out: (1) selecting the best threshold of the NN output to classify land/water surface; (2) understanding the effect of the equalization of the water and non-water pixels in the NN training dataset; and (3) finding the most important satellite NN inputs. The NN performances have been evaluated based on: spatial correlation between the SAR and Landsat-8 surface water maps, overall accuracy of the NN, as well as higher values of true positive (TP) and true negative (TN) percentages. True positive value indicates the NN ability to correctly detect water pixels, while true negative value illustrates its ability to correctly detect non-water pixels (compared to the Landsat-8 surface water maps).

### 3.3.3.1 Selection of an optimized threshold for the NN output

The first test is conducted to optimize the output threshold to distinguish water from non-water pixels. Figure 3.8 shows the histogram of the output of the NN, separating the water and non-water pixels according to the related Landsat-8 surface water map. The histograms of the water and non-water clusters intersect around 0.9, meaning that the optimal threshold to separate water from non-water pixels is close to this number. Different thresholds on the NN output values are tested (0.80, 0.85, and 0.90): for each one, the confusion matrix and the overall accuracy are calculated, with the corresponding Landsat-8 images as references. The overall accuracy and the spatial correlation increase from 98% to 99% when the threshold increases from 0.80 to 0.90 (Table 3.4), but the true positive pixel detection decreases from 92% (with threshold 0.80) to 89% (with threshold 0.90) and the false negative pixel detection increases from 8% to 11%. A threshold of 0.85 is selected here because of its good water detection performance and because it results in the predicted water surface closest to the reference map: 4430 km<sup>2</sup> from the Landsat-8 versus 4420 km<sup>2</sup> from the SAR results, i.e., a limited overestimation of 0.4% as compared to the reference map.

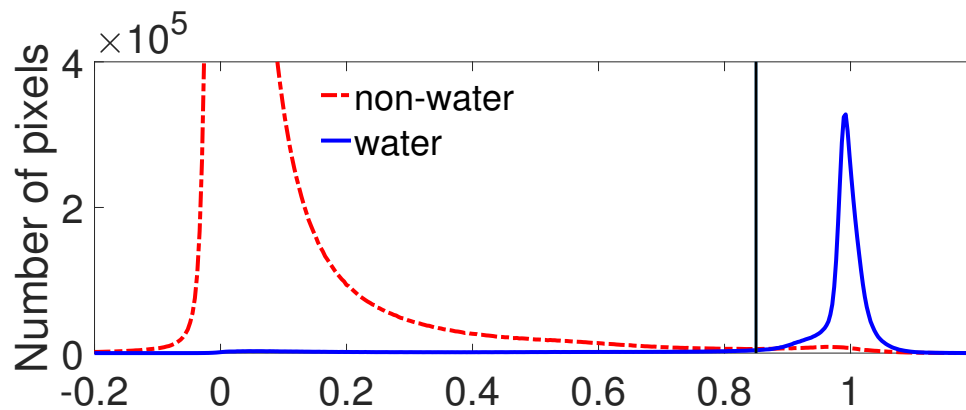


FIGURE 3.8: Histograms of the NN outputs, for water (blue) and non-water (dashed red) pixels separately, according to the corresponding Landsat-8 surface water maps. The NN uses the five initial inputs and the training dataset is equalized. The y axis range is selected to illustrate the peak of the water histogram. The black vertical line shows the chosen threshold (0.85).

TABLE 3.4: Confusion matrix of the NN classification for different thresholds.  
The NN uses the five initial inputs and the training dataset is equalized.

<b>Output Threshold: 0.80</b>				
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Overall Accuracy</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99.3%	0.7%	98%	91%
<b>Water(1) (Actual)</b>	8%	92%		
<b>Output Threshold: 0.85</b>				
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Overall Accuracy</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99.5%	0.5%	99%	92%
<b>Water(1) (Actual)</b>	9%	91%		
<b>Output Threshold: 0.90</b>				
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Overall Accuracy</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99.6%	0.4%	99%	91%
<b>Water(1) (Actual)</b>	11%	89%		

### 3.3.3.2 Equalization of water and non-water pixel number

For this test, instead of using an equal number of water and non-water pixels in the training dataset, 10% of each Sentinel-1 image is selected randomly to train the neural network, meaning that the number of non-water pixels is several times higher (10–15 times depending on each image in the training dataset) than the number of water pixels (as seen in Figure 3.8). The intersection between histograms of the NN outputs for water pixels (blue) and non-water pixels (red) moves to 0.5 (see the histogram in Figure 3.9), meaning that the value 0.5 should be selected to separate water and non-water clusters. As shown in Table 3.5, the resulting NN is very efficient at detecting non-water pixels with a true negative

detection of 99.7%, but it misses 14% of the actual water pixels (86% of true positive detection only, compared to 91% with the equalized training dataset—Table 3.4). The true positive detection of water pixels decreases because in the training database the non-water pixels are more numerous and as such have more weight in the retrieval than the water pixels. As a consequence, the NN is more effective at detecting non-water pixels, and less effective at detecting water pixels. It is concluded that the use of an equalized training data set is very important in this classification framework.

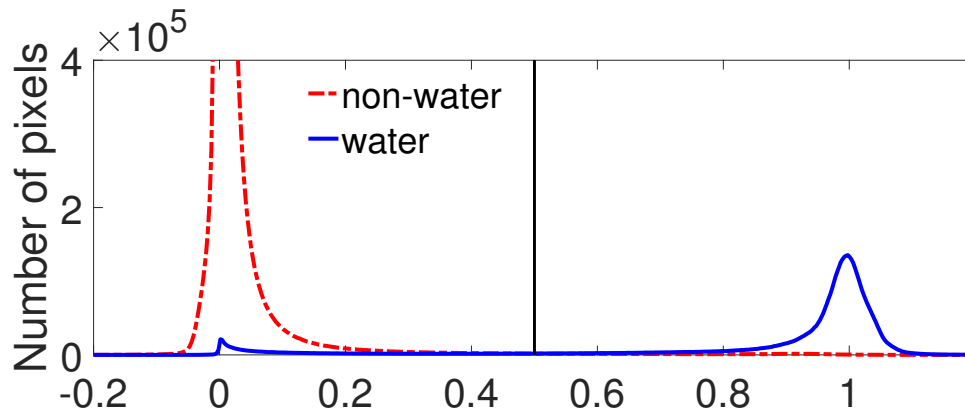


FIGURE 3.9: Histograms of the NN outputs, for water (blue) and non-water (dashed red) pixels separately, according to the corresponding Landsat-8 surface water maps. The NN uses the five initial inputs but the training dataset is not equalized. The y axis range is selected to illustrate the peak of the water histogram. The black vertical line shows the chosen threshold (0.5).

TABLE 3.5: Confusion matrix of the NN without equalization of the training dataset, for an optimum threshold of 0.5 on the NN outputs.

	<b>Non-water (0) (Predicted)</b>	<b>Water (1) (Predicted)</b>
<b>Non-water(0) (Actual)</b>	99.7%	0.3%
<b>Water(1) (Actual)</b>	14%	86%

### 3.3.3.3 Analyzing the weight of each NN satellite input

To identify the most relevant inputs for the NN classification of the water surface, 15 NNs are trained based on all 15 different combinations from the five input parameters, and their performances are evaluated following various criteria. Table



3.6 presents the best results with one to five inputs and illustrates how the overall accuracy of the NN classification increases when the number of satellite inputs increases, as compared to the reference Landsat-8 dataset. The NN trained with only the VH backscatter coefficient has a spatial correlation of 78% and a true positive accuracy (correctly detecting water pixels) of 77% compared to the reference data. The spatial correlation increases to 79%, and the true positive accuracy rises to 85% when the standard deviation of the VV backscatter coefficient is added as an input to the NN. The VV backscatter coefficient helps to increase the performance of the NN since both spatial correlation and true positive accuracy increase to 87% and 90%, respectively. The standard deviation of the VH backscatter coefficient does not significantly improve the accuracy of the NN classification. This is due to the strong linear correlation (88%) between the spatial standard deviations of the VH and the VV backscatter coefficients (the other linear correlations among the five input parameters of the NN are provided in Table 3.7). Similar to the standard deviation of the VH backscatter coefficient, the incidence angle does not have a strong impact on the performance of the NN since its accuracy remains nearly the same after adding the incidence angle as a new input. The input parameters of the NN classification are listed below, from the most important to the least important one in the NN processing:

- Backscatter coefficient VH polarization (BS\_VH)
- Standard deviation of backscatter coefficient VV polarization (STD\_VV)
- Backscatter coefficient VV polarization (BS\_VV)
- Incidence angle
- Standard deviation of backscatter coefficient VH polarization (STD\_VH)

To conclude, the water detection ability of the proposed NN increased when the input parameters are carefully selected and when an optimal output threshold is selected. An equal number of water and non-water pixels should be used in the training dataset to ensure that the NN performs equally well in classifying water and non-water clusters. The STD\_VH provides limited additional information to the NN due to its strong linear correlations with the other NN inputs. The incidence angle also plays a limited role in the NN performance. This is partly due to the rather narrow range of incidence angles, from 29° to 46°.

TABLE 3.6: The NN classification performances when adding input parameters, one at a time.

<b>One Input: BS_VH</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	98%	2%	<b>78%</b>
<b>Water(1) (Actual)</b>	23%	77%	
<b>Two Inputs: BS_VH + STD_VV</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	98%	2%	<b>79%</b>
<b>Water(1) (Actual)</b>	15%	85%	
<b>Three Inputs: BS_VH + STD_VV + BS_VV</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99%	1%	<b>87%</b>
<b>Water(1) (Actual)</b>	10%	90%	
<b>Four Inputs: BS_VH + STD_VV + BS_VV + Angle</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99.5%	0.5%	<b>91%</b>
<b>Water(1) (Actual)</b>	10%	90%	
<b>Five Inputs: BS_VH + STD_VV + BS_VV + Angle + STD_VH</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-Water(0) (Actual)</b>	99.5%	0.5%	<b>92%</b>
<b>Water(1) (Actual)</b>	9%	91%	

TABLE 3.7: Linear correlations among the five potential NN inputs.

	BS_VH	BS_VV	STD_VH	STD_VV	ANGLE
BS_VH	100%				
BS_VV	84%	100%			
STD_VH	24%	20%	100%		
STD_VV	21%	21%	88%	100%	
ANGLE	25%	22%	11%	6%	100%

### 3.4 Results and comparisons with other surface water products

The following results and comparisons involve the optimized version of the NN classification with five input parameters (an equalization of water and non-water pixels, and an output threshold at 0.85). In Section 3.4.1, the SAR-predicted surface water maps are calculated for two test areas, then compared to Landsat-8 surface water maps over the Tonle Sap Lake in Cambodia and over the Mekong river in Vietnam (see test dataset in Table 3.2). Other regions are tested but the results are not shown here. Due to the lack of in situ local surface water maps at the time of this study, a reference dataset is not available to confirm the accuracy of the Landsat-8 based maps. Therefore, an inter-comparison between Sentinel-1 estimate and other existing estimates is the only way to evaluate the new wetland product based on SAR Sentinel-1 data. First, the results are evaluated with respect to the floodability map derived mainly from the HydroSHEDS topography dataset [Lehner et al., 2006], developed by Aires et al. [2017] (Section 3.4.2). Second, time series of the SAR-derived surface water over the Mekong Delta is compared to the MODIS/Terra-derived inundation maps based on the methodology described by Sakamoto et al. [2007], for 2015 (Section 3.4.3)

### 3.4.1 Evaluation of the SAR NN classification method with Landsat-8 images

Figure 3.10 shows results of the NN classification applied over the Tonle Sap Lake in Cambodia (top) and over the Mekong river in Vietnam (bottom), in February 2016. Figure 3.10a,d shows the SAR-predicted surface water maps, Figure 3.10b,e presents the reference Landsat-8 surface water maps, whereas the differences between these two surface water maps are shown in Figure 3.10c,f.

Over the Tonle Sap Lake, both Sentinel and Landsat images were acquired on the same day (3 February 2016). The spatial correlation between the two surface water maps is 94%. The confusion matrix for this area is given in Table 3.8 (top).

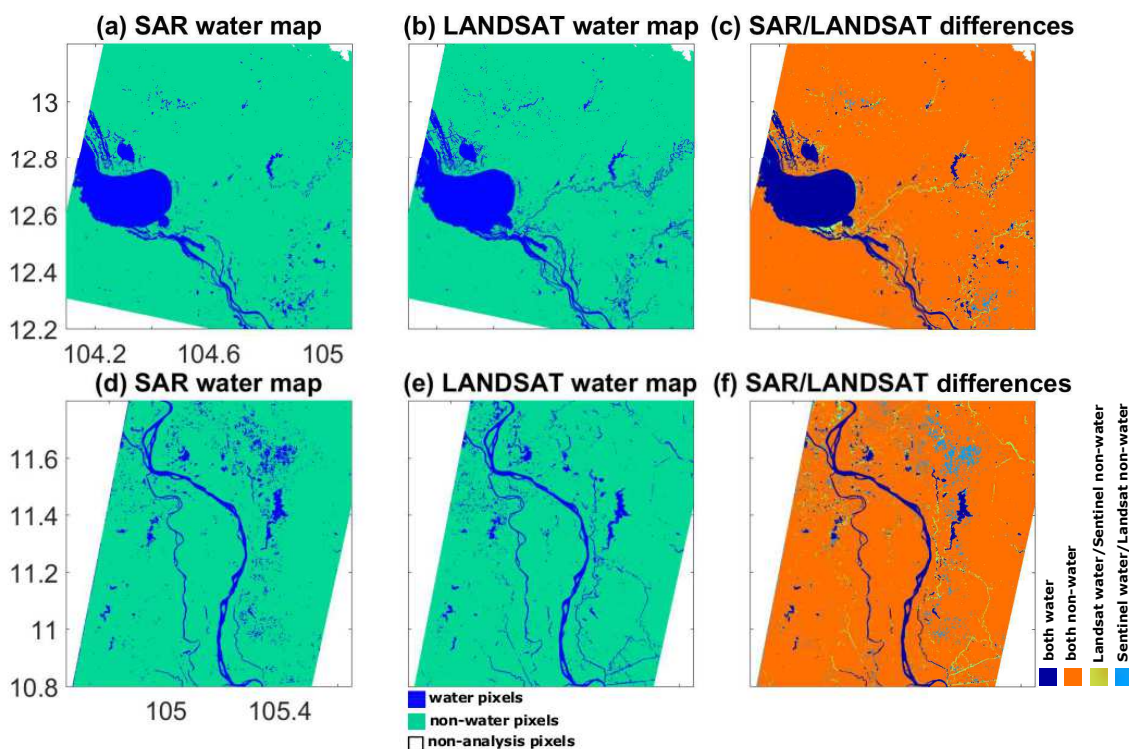


FIGURE 3.10: (a,d) SAR surface water maps; (b,e) Landsat-8 surface water maps; and (c,f) their differences; over the Tonle Sap Lake (left), and over the Mekong river (right), for February 2016. Blue color presents water pixels while orange color presents non-water pixels detected by both Sentinel and Landsat, green color is Landsat water/Sentinel non-water pixels, and light blue color is Sentinel water/Landsat non-water pixels.

TABLE 3.8: Confusion matrices (in numeric and percentage forms) of the SAR-predicted surface water maps and the Landsat-8 reference surface water maps, over the Tonle Sap Lake (top) and over the Mekong River (bottom).

<b>Tonle Sap Lake</b>		
	<b>Non-water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>
<b>Non-water(0) (Actual)</b>	11,641,078 (99.6%)	44,493 (0.4%)
<b>Water(1) (Actual)</b>	71,884 (6.5%)	1,023,457 (93.5%)
<b>Mekong River</b>		
	<b>Non-water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>
<b>Non-water(0) (Actual)</b>	10,983,583 (99.2%)	85,096 (0.8%)
<b>Water(1) (Actual)</b>	51,611 (14.3%)	309,982 (85.7%)

Overall accuracy of the classification is 99%, with a true positive water detection of 93.5%, and a false negative percentage of 6.5%. The classification correctly detects more than 99.6% of non-water pixels compared to the reference map. The classification slightly underestimates the surface water coverage by  $\sim 2.5\%$ . This is 961 km<sup>2</sup> compared to the reference surface water map derived from the Landsat-8 images of 986 km<sup>2</sup>.

The second case study is carried out over the Mekong River and its surrounding areas (latitude range [10.8°N–11.8°N] and longitude range [104.6°E–105.6°E]). The optical Landsat-8 images were taken on 19 February 2016 and the SAR Sentinel-1 images were taken 3 days later, on 22 February 2016. These Sentinel and Landsat images were not acquired on the same day, but within 3 days in the middle of the dry season when land surfaces in this area are not expected to change much. Similar to the first case study, the classification works well, even though the environment here is rather complex, with rivers and vegetated wetlands. The overall accuracy is 98.8%, with a spatial correlation of nearly 82% with the Landsat-8 reference surface water map. Confusion matrix for this

area is shown in Table 3.8 (bottom) where the true positive percentage is 85.7%, the false negative percentage 14.3%, and 99.2% of non-water pixels are classified correctly. The total surface water area derived from Landsat data is 325 km<sup>2</sup>, and it is 355 km<sup>2</sup> predicted from the NN.

Similar results are found when applying the NN classification to other areas. To conclude this comparison, the proposed NN methodology correctly detected ~90% of the water pixels observed by Landsat-8, with a spatial correlation of ~90%. The NN works better over open water bodies than over other heterogeneous environments. For instance, the NN has difficulties detecting small river branches (Southeast of the Tonle Sap Lake in Figure 3.10—top panel) although they are clearly detected with Landsat-8 images. The NN can provide water maps with high accuracy compared to the reference Landsat-8 water maps. Errors could come from the following factors:

- The SAR responses can be affected by complex interactions with the terrain and the vegetation, especially along small river banks. It can be difficult to account for this local complexity in the methodology.
- In the SAR water detection method, as in any other classifications method scheme, different parameters are selected to optimize the overall performance of the method, but local ambiguities can still exist.
- Sentinel-1 and Landsat-8 data are not always acquired on the same day.
- Using Landsat-8 quality flags, cloud-covered pixels are expected to be removed, but some cloud and cloud-shadow pixels can still be present. This can cause some ambiguities in the NN training dataset.
- Reference surface water maps derived from negative NDVI values on the Landsat-8 images are not always perfect. Water under vegetation can be difficult to detect with Landsat-8 observations. The NDVI values can also be impacted for highly turbid waters where the NIR reflectance can be higher than the red reflectance.

### 3.4.2 Evaluation using a topography-based floodability index

A global floodability index based on topography has been developed by [Aires et al. \[2017\]](#). It uses mainly the hydrological data and maps based on the Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS) dataset [[Lehner et al., 2006](#)] that has been derived from elevation measured by the Shuttle Radar Topography Mission (SRTM) satellite. This floodability index provides a static map of an estimate of the probability for a pixel to be inundated (between 0% and 100%) at the spatial resolution of 90 m, based only on topography information (such as slope in the pixel, distance to the closest river, difference of elevation with the closest river). Figure 3.11a presents this floodability index map over the Mekong Delta and Cambodia. As expected, all rivers and lakes in this area have a very high probability of being inundated (over 80%). Since this index is based only on topography, its reliability is higher for natural environments and it can be less precise over regions with strong anthropic impact such as irrigated areas. The floodability data are down-scaled from 90 m to 30 m spatial resolution to compare with predicted SAR surface water maps over the Tonle Sap Lake and the Vietnamese Mekong Delta. Each floodability pixel is divided into a  $3 \times 3$  matrix with the same value, and projected onto the Sentinel-1 grid. By comparing these two products, it is possible to see where and how Sentinel-1 water pixels are located with respect to the floodability index, and test the consistency between these two independent products. Figure 3.11b–e shows floodability maps at 30 m spatial resolution and predicted Sentinel-1 water maps, over four different areas in the Mekong Delta. SAR surface water areas are generally located in areas with high predicted inundation probabilities, as expected (see Table 3.9). A total of 98% of the SAR surface water pixels are located in areas where the floodability index is greater than 60%, while only 2% of the SAR surface water pixels are located in areas with a lower floodability index ( $\leq 60\%$ ). As mentioned earlier, the floodability index only relies upon topography information, and it can be less precise over regions with strong anthropic activities, such as irrigation. There are many rice paddies in the Lower Mekong Delta, and these irrigated fields can be missed by the floodability index, contributing to the 15% errors of SAR water pixels located in areas with a floodability index less than 80%. In the future, in complex-topography environments where SAR only data could not provide the

required accuracy for the water classification (the Red River Delta in the North of Vietnam, for example), the floodability index information could be added as another input to the NN to improve the classification performance.

TABLE 3.9: Performance of the SAR surface water classification for different ranges of floodability index.

Floodability Index	$\leq 40$	40–60	60–80	$\geq 80$
Percentage of surface water pixels detected by the NN classification	1%	1%	13%	85%

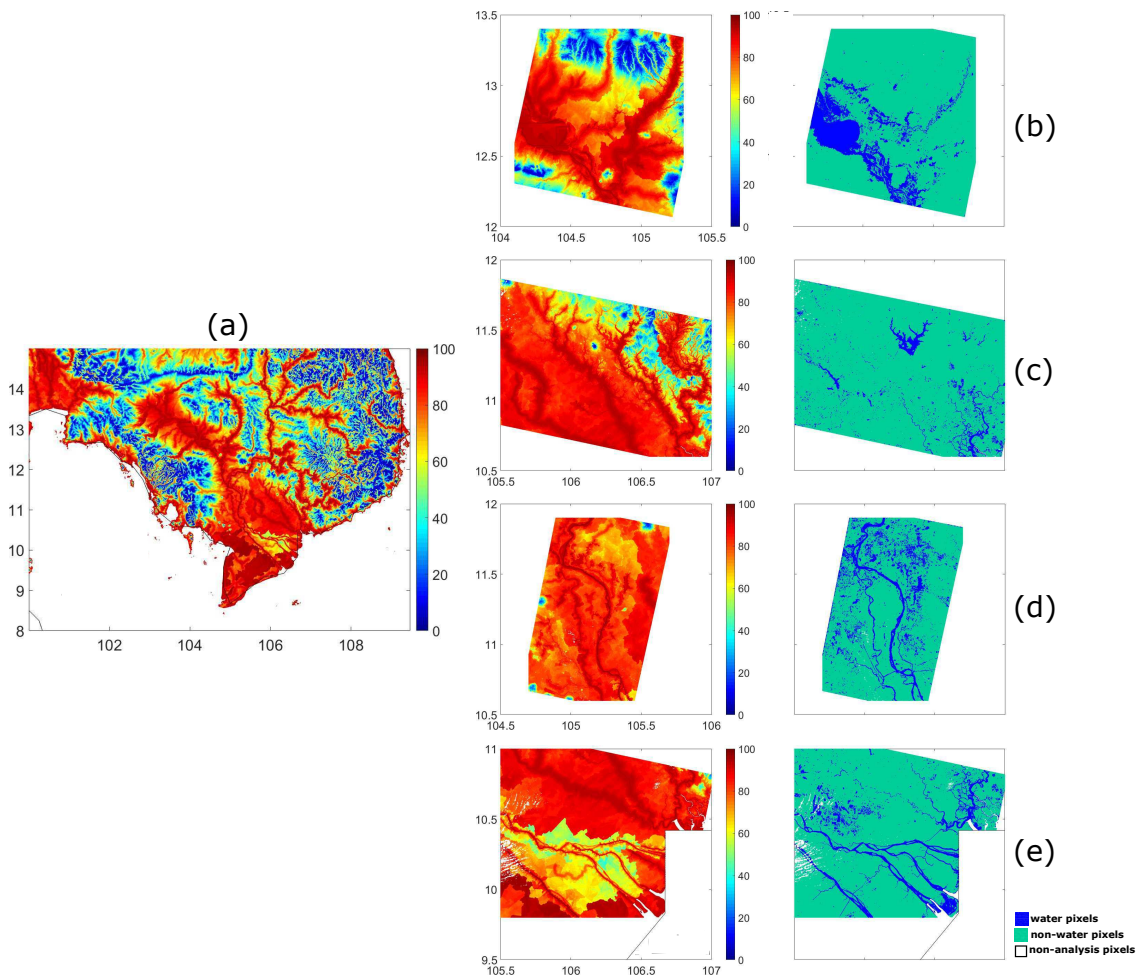


FIGURE 3.11: (a) Topography-based floodability index map over the Mekong Delta from Aires et al. [2017]. (b–e) Comparisons of floodability index maps and SAR-predicted surface water maps for four areas over the Vietnamese Mekong Delta and Cambodia.



### 3.4.3 Comparisons with MODIS/Terra-derived inundation maps

In this section, the 30 m SAR surface water maps are compared to the 500 m MODIS/Terra-derived surface water maps, for a region in the Mekong Delta. One year (2015) of SAR Sentinel-1 and MODIS/Terra data are extracted, over the same region (latitude [9.8°N–11.3°N]; longitude [104.75°E–107°E]). The MODIS surface water maps are derived from the method described by Sakamoto et al. [2007]. The methodology is re-produced to calculate surface water maps with three different states of non-water, water, and mixed pixels, respectively (as already described in Chapter 2). The total MODIS surface water is the sum of the water pixels (100% area is inundated) and mixed pixels (part of these pixels is inundated). For a mixed pixel, two hypothesis are tested: 25% or 50% of the pixel is inundated.

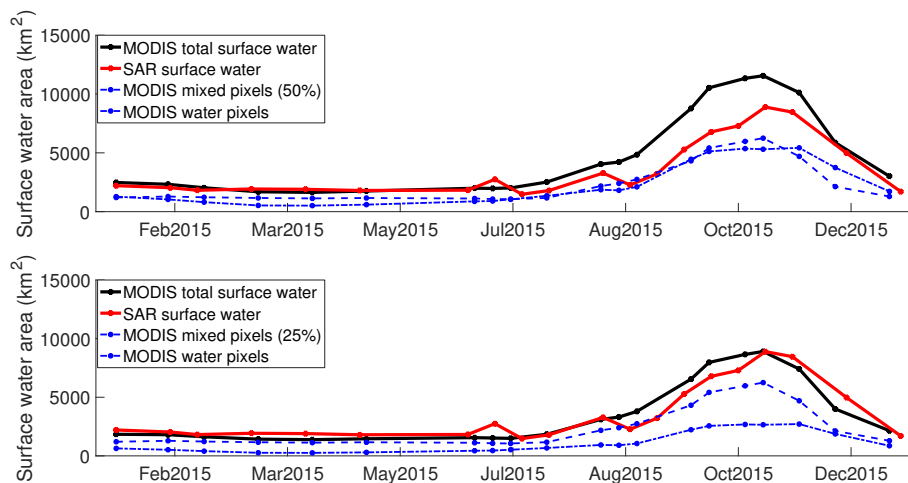


FIGURE 3.12: Time series of the surface water detected by SAR (red) and MODIS data (black), over the Mekong Delta (Latitude [9.8°N–11.3°N]; Longitude [104.75°E–107°E]), for 2015. Two hypotheses are tested for the MODIS mixed pixels: 50% inundated (top Panel), and 25% inundated (bottom Panel).

Twenty SAR Sentinel-1 observations are available over the selected region for the year 2015 (less than two images per month—see Table 3.3). The surface water extent calculated from the SAR and MODIS data are presented in Figure 3.12. With the first assumption (25% of a mixed MODIS pixel is covered by water), the two surface water extents have very similar seasonal cycles and amplitudes, with a correlation of  $\sim 99\%$  (Figure 3.12-bottom). For the second assumption (the surface water extent of a mixed pixel increases to 50%), the difference in surface water areas increases, but without significant changes in the seasonal cycle and still

with high correlation to the SAR surface water time series. For both hypotheses, the SAR and MODIS surface water extents reach their maximum at the same time (around 20 October 2015). Total inundated areas derived from SAR and MODIS are very close during the dry season (January to July). The cloud contamination of the MODIS estimate is low during that season. During the rainy season, more cloud contamination is expected in the MODIS estimates, and the discrepancies between the two surface water extents increase. The SAR-derived surface water estimate is expected to be more reliable due to its insensitivity to the cloud cover, but at this stage there is no convincing dataset at this spatial resolution to confirm it, as mentioned before.

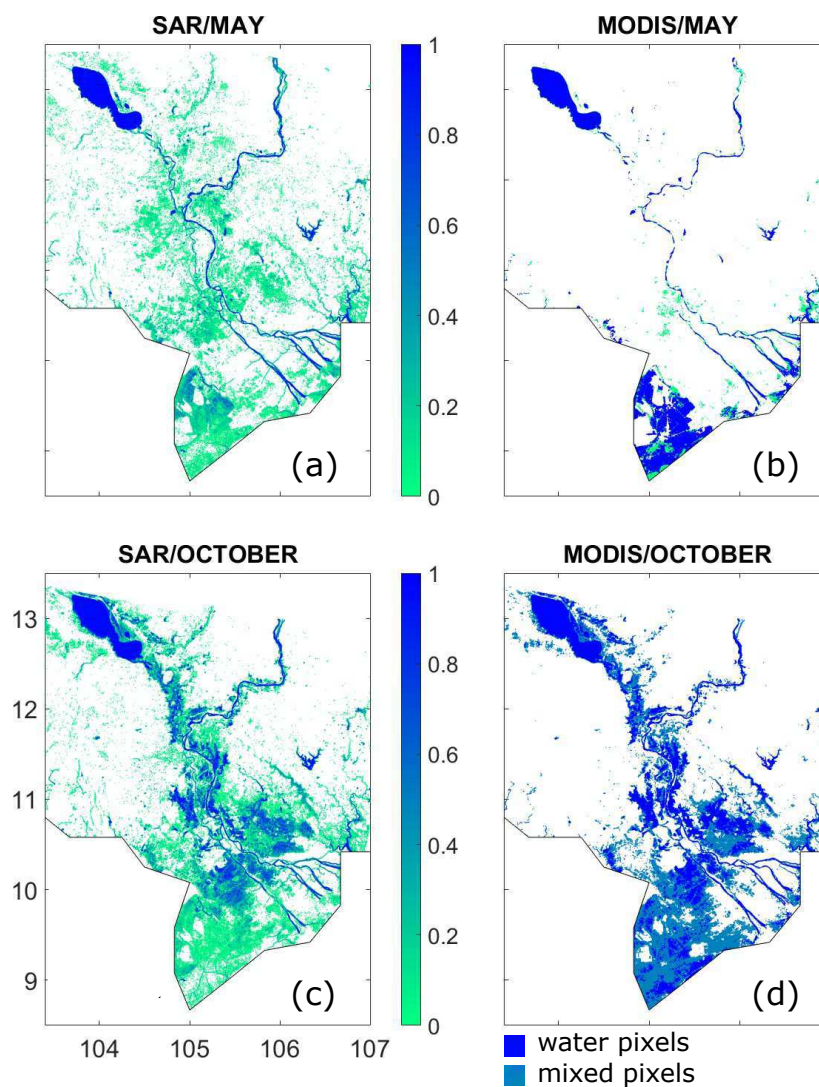


FIGURE 3.13: (a,c) SAR and (b,d) MODIS surface water maps at 500 m resolution over the Mekong Delta in May (a,b) and October (c,d) 2015.

To evaluate the consistency of the spatial structure between the SAR-derived and the MODIS-derived surface water maps, 10 SAR Sentinel-1 images were downloaded to cover the Mekong Delta and the Tonle Sap Lake (five images in May and five images in October 2015). For comparison purposes and to calculate the spatial correlation, the SAR surface water maps are aggregated from the 30 m resolution to the 500 m resolution of the MODIS-derived inundation maps (see Figure 3.13a,c). As a consequence, Sentinel-1-derived inundation maps are not binary (0 for non-water pixels or 1 for water pixels), but they are converted into a percentage of surface water at 500 m spatial resolution. For the dry season (Figure 3.13a,b—May 2015), the spatial correlation between the two surface water maps is 68%. A total of 4% of the area is inundated for the SAR estimation, while it is 5% for the MODIS estimates. For the rainy season (October 2015), the spatial correlation of the two maps increases to 78%, with 8% inundated area with the SAR and 11% with MODIS. For these calculations, the hypothesis of 25% inundation of the MODIS mixed pixels is used. Although SAR-derived and MODIS-derived water maps have a very similar seasonal cycle and similar spatial distribution of the water bodies, confirming the wetland seasonal cycle over this region, there are differences in the total surface of inundated areas. It comes mainly from the difference of spatial resolution between the two satellites. First, MODIS sensors cannot detect very small surface water fractions due to their spatial resolution. Second, the MODIS mixed pixels include water surfaces, vegetation surface and bare soil, and the percentage of each surface type within the pixel is not quantified.

### 3.5 Improvement of the Neural Network

One of the limitations of the proposed NN is its inability to classify perfectly between water surfaces and very dry, flat surfaces because of their similar backscatter signatures [Prigent et al., 2001, 2007, 2015]. Both surface types reflect almost all incoming signals to other directions, making their backscatter coefficient very low, and they all appear very dark in the images for both VH and VV polarizations. In this section, temporal information from a set of several SAR Sentinel-1

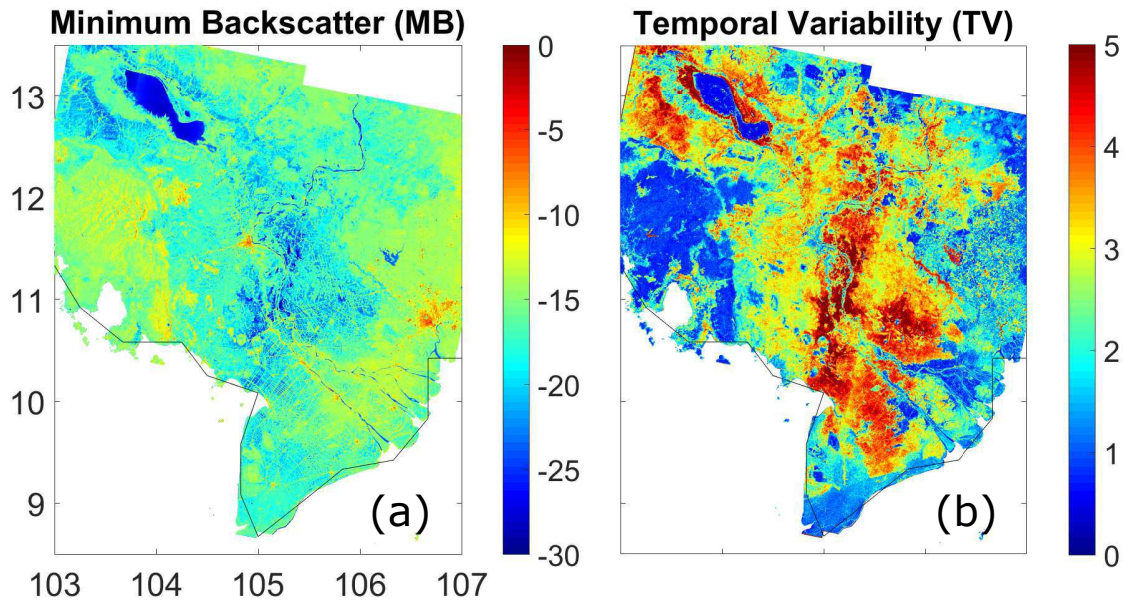


FIGURE 3.14: (a) Minimum backscatter and (b) Temporal variability of the VH polarization over the Vietnam Mekong Delta and Cambodia, calculated from 12 other SAR Sentinel-1 observations in 2016.

observations are added to the NN to examine if they can improve the performance of the NN over these problematic areas.

Two new parameters derived from the VH polarization are added to the input of the NN: the temporal variability (TV\_VH) and the minimum backscatter (MB\_VH). The temporal variability is defined as the standard deviation of the multi-temporal of the backscatter dataset, and the minimum backscatter is defined as the 5<sup>th</sup> percentile of the backscatter dataset [Santoro et al., 2015b]. These two new parameters are calculated from 12 other Sentinel-1 observations. The TV and MB from the VV polarization are not studied because of its high linear correlations with that derived from the VH polarization (more than 85% as seen in Table 3.7 between BS\_VH and BS\_VV, or between spatial STD\_VH and STD\_VV). Examples of TV\_VH and MB\_VH maps, calculated from 12 other Sentinel-1 observations between January and December 2016, are shown in Figure 3.14. As a consequence, a new NN with 7 input parameters is trained. Performance comparisons between the 5-input NN (NN5) and the 7-input NN (NN7) are discussed in the following section.

### 3.5.1 Performance comparisons between NN5 and NN7

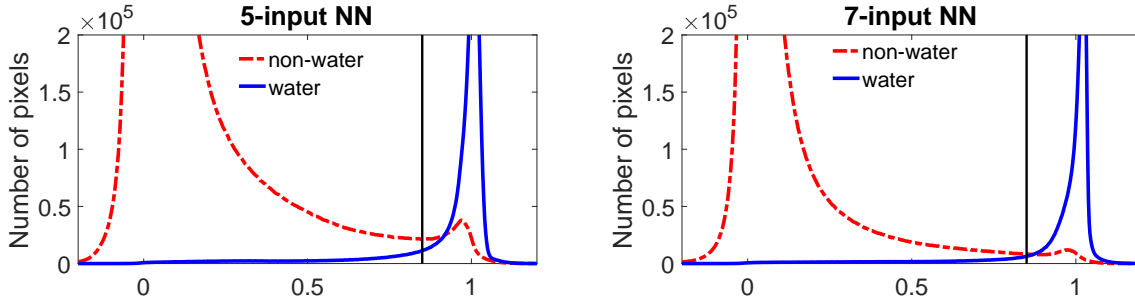


FIGURE 3.15: Histograms of the 5-input (left) and 7-input (right) NNs, for water (blue) and non-water (dashed-red) pixels separately, according to the corresponding Landsat-8 surface water maps. The training datasets is equalized. The y axis range is selected to illustrate the water histogram. The black vertical lines show the chosen threshold (0.85).

TABLE 3.10: Confusion matrices of the SAR-predicted surface water maps derived from the NN5 (top) and from the NN7 (bottom), and the Landsat-8 surface water reference maps. The output threshold 0.85 is used.

5-Input NN				
	Non-Water(0) (Predicted)	Water(1) (Predicted)	Overall Accuracy	Spatial Correlation
Non-water(0) (Actual)	98.7%	1.3%	97.6%	85%
Water(1) (Actual)	13.4%	86.6%		
7-Input NN				
	Non-Water(0) (Predicted)	Water(1) (Predicted)	Overall Accuracy	Spatial Correlation
Non-water(0) (Actual)	99%	1%	98.2%	88.3%
Water(1) (Actual)	11.2%	88.8%		

The NN7 is trained the same way as the NN5, except that it has two more input parameters as already introduced. A test dataset is used to compare performances of the two NNs. Histograms of the outputs of the two NNs are shown in Figure 3.15, along with the corresponding Landsat-8 surface water reference

maps. It is clear that the small peak of the non-water curve (dashed red) from the NN7 (right) is lower than that from the NN5 (left). This means that the performance of the NN7 is better than the NN5, because the NN7 shows less ambiguities than the NN5. The chosen output threshold (0.85) is applied to outputs of the two NNs to make SAR-predicted surface water maps, then compared to the Landsat-8 surface water reference maps. Detailed comparisons are shown by the confusion matrices in Table 3.10. All indices (true positive and true negative detection, overall accuracy and spatial correlation) from the NN7 are slightly higher than that from the NN5. The spatial correlation increase from 85% with the NN5 to 88.3% with the NN7, while the overall accuracy increases by 0.6% (from 97.6% to 98.2%). The NN7 is better than the NN5 in detecting water pixels since the true positive detection increases by more than 2%, from 86.6% to 88.8%.

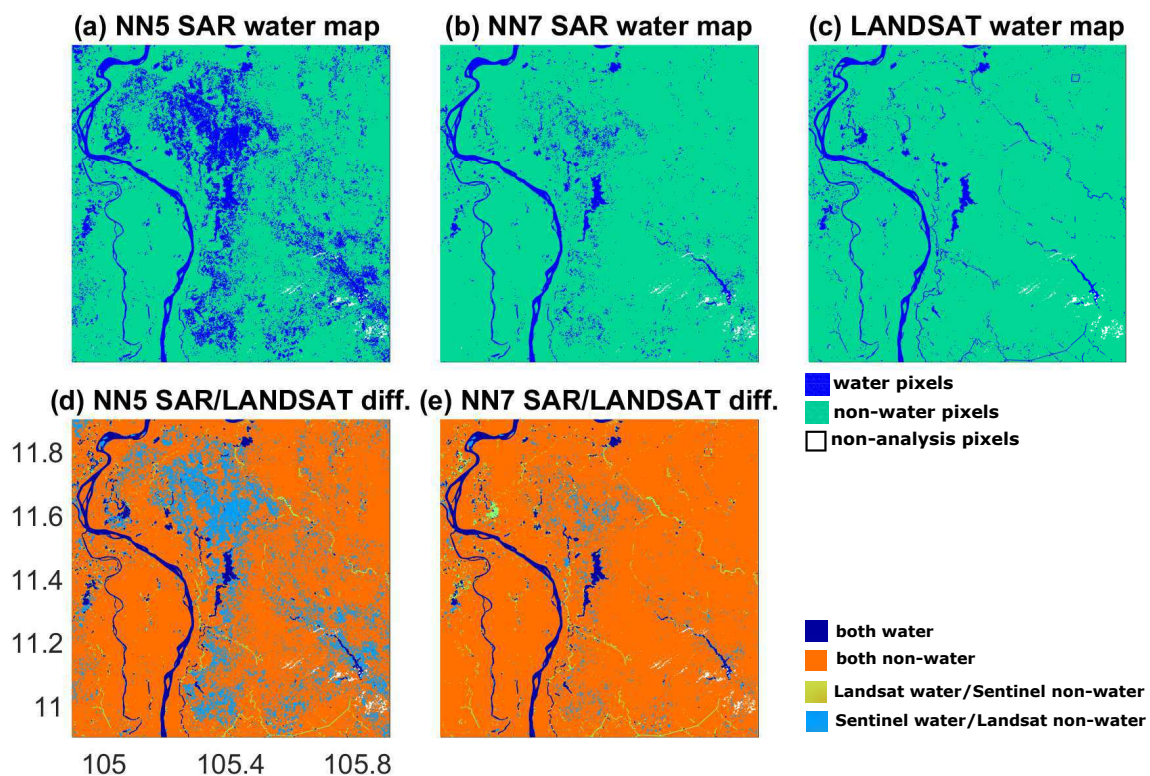


FIGURE 3.16: (a,b) SAR surface water maps derived from the NN5 and NN7; (c) Landsat-8 surface water map; and (d,e) their differences. Sentinel-1 images were acquired on 10 April 2016, while Landsat-8 images were acquired 4 days earlier on 6 April 2016. Landsat-8 cloud-covered pixels is less than 1%, and have been removed.

TABLE 3.11: Confusion matrices of the SAR-predicted surface water maps derived from the NN5 (top) and from the NN7 (bottom), and the Landsat-8 surface water reference maps, for an  $1^\circ \times 1^\circ$  area shown in Figure 3.16. The output threshold 0.85 is used.

<b>5-Input NN</b>				
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Overall Accuracy</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	12,531,691 (96.8%)	410,031 (3.2%)	<b>96.4%</b>	<b>60%</b>
<b>Water(1) (Actual)</b>	58,440 (15.6%)	315,208 (84.4%)		
<b>7-Input NN</b>				
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Overall Accuracy</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	12,860,918 (99.4%)	80,804 (0.6%)	<b>98.9%</b>	<b>80%</b>
<b>Water(1) (Actual)</b>	62,834 (16.8%)	310,814 (83.2%)		

Surface water maps derived from the NN5 and NN7, focusing on a problematic area (latitude [10.9°N–11.9°N]; longitude [104.9°E–105.9°E]), are shown in Figure 3.16a,b, along with the reference Landsat-8 surface water map (Figure 3.16c), as well as the differences between SAR-derived and Landsat-derived water maps (Figure 3.16d,e). For this  $1^\circ \times 1^\circ$  area, the NN5 detects much more water pixels than the NN7. Total inundated surface derived from the NN5 is 652 km<sup>2</sup>, nearly double that derived from the NN7 and from Landsat-8 data (352 km<sup>2</sup> and 336 km<sup>2</sup>, respectively). Compared to the Landsat-8 reference water map, most of the wrong classified water pixels from the NN5 are corrected as non-water pixels by the NN7. Confusion matrices of the SAR and Landsat derived surface water maps for this area are shown in Table 3.11. Although the true positive detection slightly decreases (84.4% to 83.2%), but all other indices increase when using the NN7. The spatial correlation with the Landsat-8 reference surface water map rise by 20% (from 60% to 80%). True negative detection increases from 96.8% to 99.4%, making the overall accuracy going from 96.4% to 98.9%.

By comparing the NN5 and NN7 over a problematic area, it is proved that

the NN performance can be slightly improved by adding temporal information derived from the temporal variability and the minimum backscatter of the VH polarization. Adding two more parameters to the NN input partly solves problems related to the similar backscatter signatures between water surfaces and dry, flat surfaces.

The new trained NN7 is used to produce regular surface water maps for the area shown in Figure 3.14, for the period from January 2015 to June 2017. Its surface water time series is shown in Figure 3.17. It shows clearly the effect of the super drought in the Mekong basin in 2015 that made the maximum surface water just a half of that in 2016. An inundated frequency map (for the 2015-2017 period) is created based on all available SAR-derived surface water maps (Figure 3.18). Comparisons with the same inundated frequency maps derived from optical MODIS, and Landsat data will be shown in Chapter 4. Monthly temporal surface water changes over the Mekong Delta and Cambodia are illustrated in Figure 3.19, showing when and where the surface water has changed in 2016.

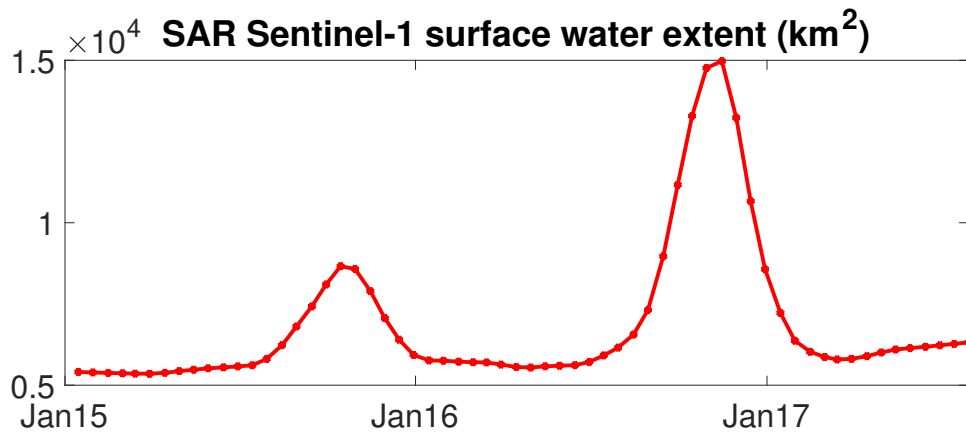


FIGURE 3.17: SAR Sentinel-1 surface water time series from January 2015 to May 2017 for the area shown in Figure 3.14.

### 3.6 Conclusions and perspectives

This study presents a methodology to monitor and quantify surface water under all weather conditions within the Vietnam Mekong Delta and Cambodia, using



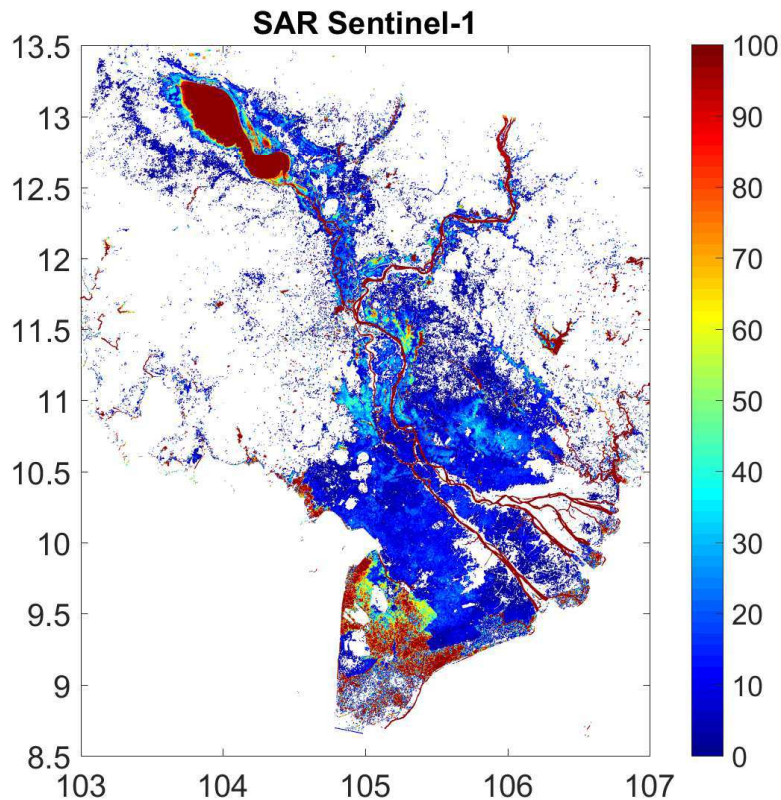


FIGURE 3.18: Inundation frequency between January 2015 and June 2017 derived from SAR-predicted surface water maps over the Vietnam Mekong Delta and Cambodia.

high quality Sentinel-1 SAR observations, freely available online. The methodology is based on a neural network classification trained with optical Landsat-8 images at 30 m spatial resolution. The information content of each satellite input is analyzed and the inputs are selected to optimize the performance of the classification. This method allows for the detection of surface water with good accuracy when compared to results derived from visible and NIR data under clear sky conditions, as well as when compared to a floodability map derived from topography data. Surface water maps derived from the proposed NN show a spatial correlation of  $\sim 90\%$  when compared to Landsat-8 water maps, with a true positive water detection of  $\sim 90\%$ . Compared to MODIS/Terra water maps over the Delta in 2015, this product shares the same wetland seasonal cycle and dynamics, with a temporal correlation of  $\sim 99\%$ . Temporal information in the SAR backscatter (i.e., minimum and standard deviation of the VH polarization time series)

are also useful for the NN to improve its ability in detecting water pixels. More detailed comparisons with MODIS surface water maps, as well as other hydrological parameters will be shown and discussed in Chapter 4.

In the future, the NN methodology will be applied to other areas under similar environments in Southeast Asia and in other parts of the globe, then in more vegetated environments. The final goal is to develop a general method capable of performing at the global scale to exploit the full spatial coverage of the Sentinel-1 mission. For this purpose, several approaches will be tested to improve the retrieval scheme. First, the introduction of a priori information from a topography-based floodability index will increase information on flooding and reduce ambiguities in the SAR signal with other surface parameters. Second, with the launch of the optical Sentinel-2 satellite, Sentinel-2 observations could be used to replace Landsat-8 data, and to train the SAR surface water classification under clear sky conditions. The classification could then be extended to the cloudy areas using the SAR data. Third, the high-resolution inundation extent retrieval maps could be post-processed in order to reduce the inherent noise in such high-spatial retrievals. Random walk techniques are planned to test for that purpose.

Part of the work in this chapter was selected to write a paper named "**Surface Water Monitoring within Cambodia and the Vietnamese Mekong Delta over a Year, with Sentinel-1 SAR Observations**", published in *Water* in May 2017.

Regular (twice a month) Sentinel-1 surface water maps for this area from January 2015 to present time are stored on the University of Science and Technology of Hanoi (USTH) Space and Aeronautics Department website, and they are provided freely to final users. The surface water maps are updated monthly.

English version: <http://space.usth.edu.vn/en/news/projects/>

Vietnamese version: <http://space.usth.edu.vn/vi/news/du-an/>

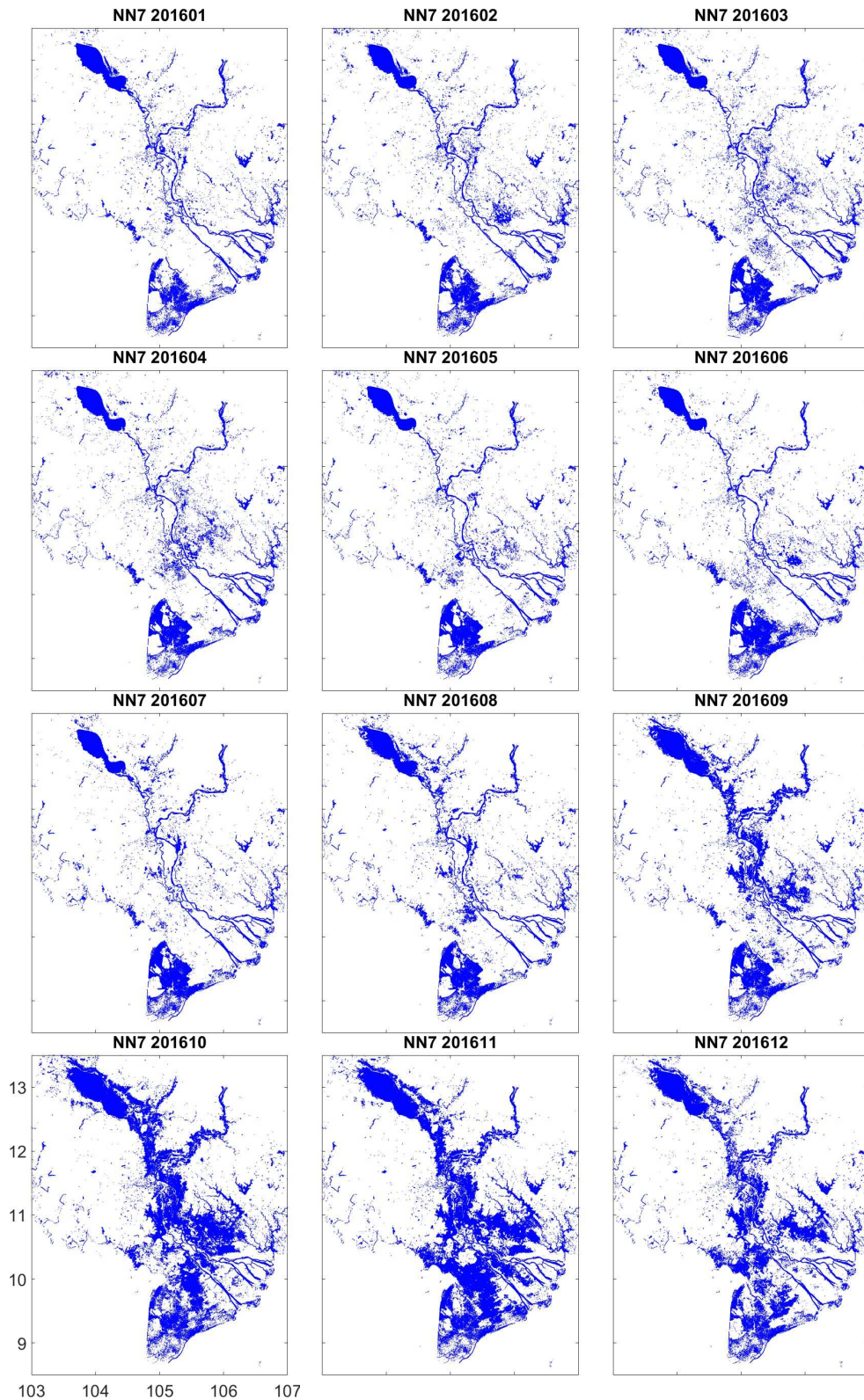


FIGURE 3.19: Monthly temporal changes of surface water extent in the Vietnam Mekong Delta and Cambodia, for 2016.

## **Chapter 4**

# **Toward the analyses of the change in surface water volume within the lower Mekong Delta**



## 4.1 Introduction

Understanding the variation of surface water storage is now recognized as a major task in climate research and in water resource management because it is a key component to study the terrestrial branch of the global water cycle [Bullock and Acreman, 2003]. Quantifying the temporal variation of the surface water volume has many applications, and our current knowledge of the interannual variability of the land surface water storage is still rather limited [Alsdorf et al., 2007]. Estimations of surface water volume variations within large river basins mostly relied on in situ observations and hydrological models [Coe et al., 2002; Goteti et al., 2008]. Recently, variations of the surface water volumes at different river basins can be estimated using satellite remote sensing techniques [Frappart et al., 2006a,b, 2011; Papa et al., 2013, 2015].

This chapter is developed to take advantages of satellite observations to estimate the variation of the surface water volume in the lower Mekong basin over the last decades. By definition, the surface water volume is the product of the surface water extent and the surface water height. The MODIS-derived surface water extent dataset is used as the first component. The methodology is well described in Chapter 2. Here, MODIS-derived surface water extent is evaluated with other datasets, including SAR-derived surface water extent, precipitation and in situ water level data. It is also compared to the Multivariate ENSO index to understand the effects of the ENSO index to variation of surface water extent. The surface water height at satellite virtual stations are extracted from satellite altimetry data, then are also evaluated by comparing with results from previous studies and with in situ water level data at gauge stations when available. The consistency between the two components, surface water extent and height, is checked first by comparing their monthly variations for the common period. Then the surface water volume variation is calculated from the two components, based on a methodology described in Frappart et al. [2008, 2011]. Results are evaluated by comparing with the total land surface water volume variation from GRACE data, and with variation of the discharge of the Mekong River from in situ measurements.

Section 4.2 introduces all datasets used in this chapter. Evaluations of the MODIS-derived surface water extent are presented in section 4.3. The analyses of the satellite altimeter data to produce surface water level maps are described in section 4.4. Section 4.5 presents the methodology to calculate the water volume variation and evaluates its results. Section 4.6 concludes this chapter.

## **4.2 Datasets**

### **4.2.1 Satellite altimetry data**

Radar satellite altimetry data from three satellites (ENVISAT, JASON-2 and SARAL) are used to estimate time series of surface water level over the Tonle Sap Lake, and along the main stream of the Mekong River. Altimetry data used in this study were developed, validated, and distributed by the Centre of Topography of the Oceans and the Hydrosphere (CTOH) in the Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS), France (<http://ctoh.legos.obs-mop.fr/>). The Environmental Satellite (ENVISAT), the successor of the ESA European Remote Sensing satellites (ERS-1 and ERS-2), is the largest civilian Earth observation mission with ten instruments onboard, including an altimeter. The Joint Altimetry Satellite Oceanography Network satellite series (JASON-1,2,3) are the successors of the TOPEX/Poseidon mission. It is a joint project between NASA and the Centre National d'Études Spatiales (CNES) with the main objective focusing on oceanography (measurement of ocean surface topography, surface wind speed and mean wave height). The Satellite with ARgos and ALtiKa (SARAL), is a joint project between the Indian Space Research Organization (ISRO) and CNES. SARAL/ALTIKA instrument was designed to fill the gap between ENVISAT and Sentinel-3, and it provides a better spatial resolution and a higher vertical resolution for oceanographic applications (sea surface height, wave height, wind speed, ice, coastal areas, and continental water bodies, for example). A summary of the satellites with altimeters is presented in Table 4.1. ENVISAT/SARAL (left) and JASON-2 (right) ground tracks in the Mekong basin are displayed in Figure 4.1. Note that ENVISAT and SARAL satellites are designed

to have the same ground tracks until the ENVISAT had to change to a lower orbit in 22 October 2010.

TABLE 4.1: Several commonly used satellites providing altimetry data.

Satellites	Agency	Re-visiting (days)	Altitude (km)	Instrument	Frequency (GHz)
ENVISAT (2002-2012)	ESA	35	800	Radar Altimeter 2	13.575 3.2
JASON-1 (2001-2013)	NASA/ CNES	10	1336	Poseidon 2	13.575 5.3
JASON-2 (2008-present)	NASA/ CNES	10	1336	Poseidon 3	13.575 5.3
JASON-3 (2016-present)	NASA/ CNES	10	1336	Poseidon 3b	13.575 5.3
SARAL (2013-present)	CNES/ ISRO	35	800	ALTIKA	35.75

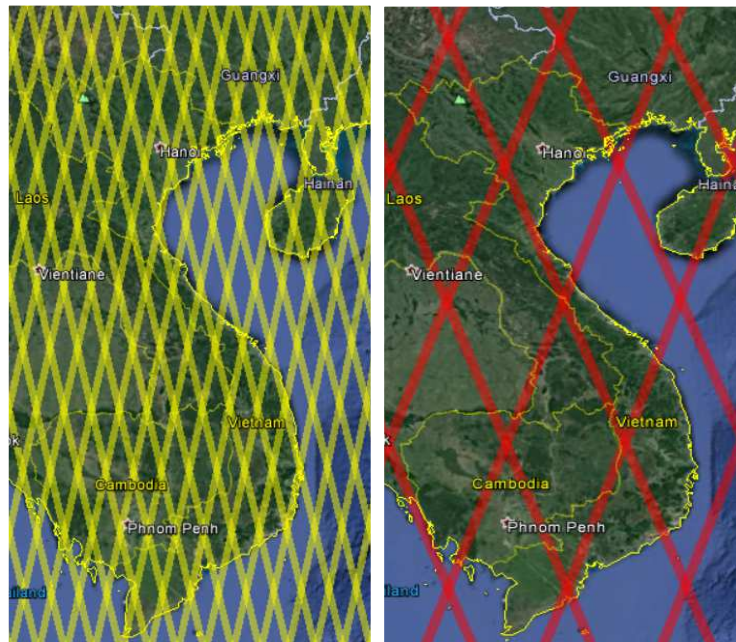


FIGURE 4.1: ENVISAT/SARAL (left) and JASON-2 (right) ground tracks over the Mekong basin (from China to Vietnam).

#### 4.2.2 GRACE data

The Gravity Recovery And Climate Experiment (GRACE) data can be used to derive the monthly variation of the total land water storage in the lower Mekong



basin with an accuracy of  $\sim 1.5$  cm of equivalent water thickness [Papa et al., 2013]. It is based on measurements of the spatiotemporal changes in the gravity field of the Earth. Since its launch in March 2002, monthly GRACE gravity solutions data are provided by three different processing centers: the GeoforschungsZentrum Potsdam (GFZ), the Center for Space Research at University of Texas, Austin (CSR), and the Jet Propulsion Laboratory (JPL). To reduce noise in the gravity field solutions, the average of the three products is used in this study as suggested in Sakumura et al. [2014].

### 4.2.3 In situ water level and river discharge data

In situ water level provided by the Mekong River Commission (MRC) are used as the reference for comparisons with results derived from satellite altimetry data. Daily in situ water level data at 8 gauge stations along the main stream of the Mekong River in Laos PDR and Cambodia, over the 2008-2016 period, can be found at the MRC's webpage: [http://ffw.mrcmekong.org/historical\\_rec.htm](http://ffw.mrcmekong.org/historical_rec.htm). Water level data were available only during rainy seasons between 2008 and 2012, before the full records for each year were provided from 2013 until present time. Table 4.2 shows locations of the 8 MRC gauge stations and the distances to its closest ENVISAT/SARAL satellite virtual stations (VS). A "virtual station" can be defined as intersections between the satellite ground tracks and a water surface [Roux et al., 2010].

TABLE 4.2: Locations of 8 in situ gauge stations along the Mekong River main stream in Laos PDR and Cambodia, and the distance to its closest ENVISAT/SARAL satellite virtual station.

Stations	Latitude (°N)	Longitude (°E)	Country	Distance to the closest VS (km)
Tonle Sap	11.48	104.48	Cambodia	$\sim 80$
Kompong Cham	11.59	105.29	Cambodia	$\sim 25$
Kratie	12.28	106.00	Cambodia	$\sim 30$
Pakse	15.06	105.48	Laos PDR	$\sim 35$
Mukdahan	16.35	104.44	Laos PDR	$\sim 65$
Nakhon Phanom	17.25	104.46	Laos PDR	$\sim 12$
Pak Sane	18.22	103.39	Laos PRD	$\sim 20$
Nong Khai	17.52	102.43	Laos PRD	$\sim 20$

Monthly in situ discharge data at Tan Chau (latitude: 10.48°N, longitude: 105.13°E) and Chau Doc (latitude: 10.42°N, longitude: 105.06°E) stations in Vietnam are also collected for comparisons, but data are available only for a limited period (2002-2010). The sum of discharge at the two stations is almost equal to the total discharge of the Mekong river when it flows from Cambodia to Vietnam. In situ discharge data are provided by the Vietnam Southern Regional Hydrometeorological Center (<http://www.kttv-nb.org.vn/>).

#### **4.2.4 Precipitation data**

Precipitation rate data derived from the gauge-calibrated Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis (TRMM/TMPA, 3B42 V7) over the Mekong basin are also used for comparisons. Precipitation rate data are available from 1998 to 2012, but only data for the common period (2001-2012) are used. To understand the contribution of local and upstream rainfall of the Mekong basin to the variation of surface water over the Mekong Delta and Cambodia areas, precipitation data are divided into two main parts. The first part is located lower than latitude 15°N to the Mekong River mouths, and it represents the local rainfall. The second part is located between latitude 15° until the source of the basin in China, and it represents the upstream rainfall.

The Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) dataset [Yatagai et al., 2012] can be used for comparisons, but this dataset is more suitable for historical studies because it only provides precipitation data for the 1951-2007 period.

#### **4.2.5 The El Niño-Southern Oscillation index**

The El Niño-Southern Oscillation (ENSO) is a naturally occurring phenomenon that involves fluctuating of winds and sea surface temperature over the tropical eastern Pacific Ocean. ENSO is one of the most important climate phenomena on Earth because it can influence temperature and precipitation in many regions across the globe. ENSO has three phases: the two opposite phases are known as

"El Niño" and "La Niña", and a "Neutral" phase is in the middle of the continuum. During the El Niño phase (or the warm phase), the sea surface temperature in the Pacific Ocean is usually warmer than the normal state. In contrast, the sea surface temperature in the Pacific Ocean is cooler than the normal state during the La Niña phase (or the cool phase). El Niño and La Niña events occur on average every two to seven years, and typically, El Niño occurs more frequently than La Niña.

In this chapter, the effects of the ENSO on the variations of surface water extent and precipitation in the lower Mekong basin are examined by comparing their time series variations with an ENSO index. The Multivariate ENSO Index (MEI) is chosen because this index integrates multiple climate variables that make it suitable for climate-land-atmosphere interaction studies [Wolter and Timlin, 1998]. MEI has a long monthly record back to 1950, and it is updated monthly. MEI data are developed by Wolter and Timlin [1993, 1998], and are provided on a National Oceanic & Atmospheric Administration's (NOAA) webpage: <https://www.esrl.noaa.gov/psd/enso/mei/>.

### **4.3 Evaluations of the MODIS-derived surface water extent**

As already mentioned previously, variation of the surface water volume is calculated as the product of the surface water extent and height. The accuracy of the two components need to be checked to guarantee good quality of the surface water volume variation. For validation of the first component, the MODIS-derived surface water extent is evaluated with the SAR-derived surface water extent (the output of Chapter 3), precipitation data, and in situ water level data before being used to calculate surface water volume variation in the lower Mekong basin. The MODIS-derived surface water extent is also compared to the ENSO index to understand the relationship between them.

### 4.3.1 Comparison with SAR-derived surface water extent

Time series of surface water extent derived from optical MODIS observations and SAR Sentinel-1 observations over a common area in the Mekong basin (latitude: 8.5°- 13.5°N; longitude: 103°- 107°E) for two years (2015-2016) are shown in Figure 4.2. Similar to findings in Chapter 3 (see Figure 3.12), there is an amplitude gap between MODIS- and SAR-derived surface water extent, but the two products share the same dynamics with a high linear temporal correlation (94%) over the common period. The two surface water extent time series confirm the wetland seasonality and its variation in the lower Mekong basin. However, it is impossible to conclude which dataset is better because no reference dataset for comparison is available. SAR-derived products are expected to be more reliable and closer to the reality than MODIS-derived ones because of its cloud penetration ability, as well as its higher spatial resolution (30 m compared to 500 m of MODIS imagery).

Inundation frequency maps in the studied area, calculated from different optical Landsat, MODIS and SAR Sentinel-1 satellite observations are shown in Figure 4.3. Landsat inundation frequency map is calculated from the global water occurrence product by Pekel et al. [2016], based on ~3 million optical Landsat satellite images during 32 years (from March 1984 to October 2015). This product gives the frequency of global water occurrence at 30 m spatial resolution. MODIS inundation frequency map is calculated based on 16-year optical MODIS observations (2001-2016, as presented in Chapter 2), and Sentinel-1 inundation frequency map is calculated based on 2-year SAR observations (2015-2016, as presented in

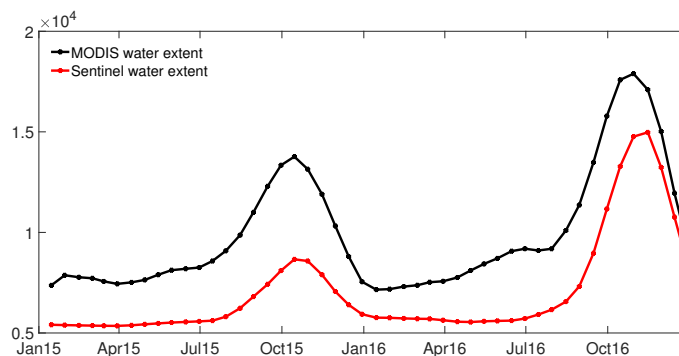


FIGURE 4.2: Surface water extent time series derived from optical MODIS and SAR Sentinel-1 observations in the Mekong Delta and Cambodia, over the common period (2015-2016).

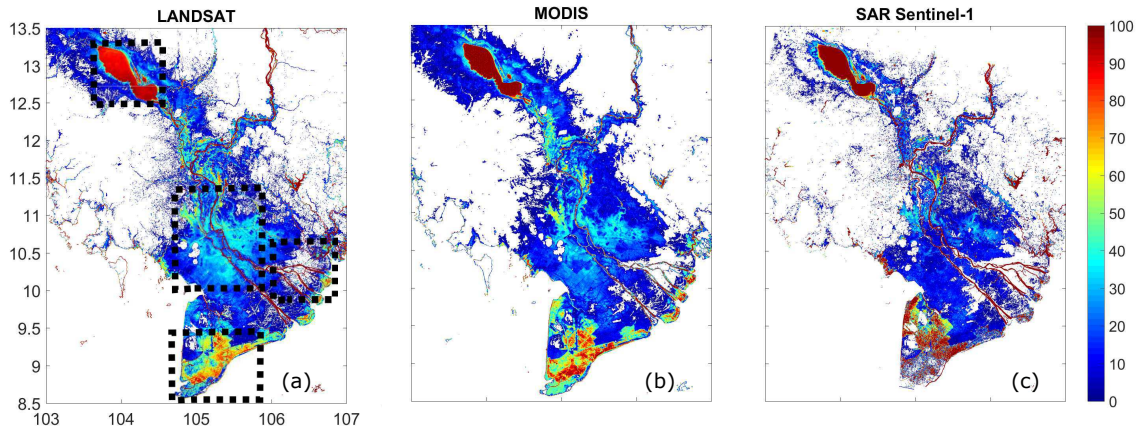


FIGURE 4.3: Inundation frequency derived from 32-year optical Landsat observations (a) [Pekel et al., 2016], 16-year optical MODIS observations (b), and 2-year SAR Sentinel-1 observations (c), over the Mekong Delta and Cambodia.

Chapter 3). Although the total number of Landsat, MODIS, and Sentinel-1 observations used to produce the three inundation frequency maps are different, spatial distributions of wetlands, rivers, lakes and water bodies are very similar in these maps. Lakes and rivers appear clearly in all the three maps, with high inundation frequency. The wetland distribution for the area over the Cambodia-Vietnam border (latitude:  $10^{\circ}$ – $11.5^{\circ}$ N, longitude:  $104.5^{\circ}$ – $106^{\circ}$ E) is similar in the three maps, especially between the two maps derived from optical data. It is less extended in the SAR map because the map was created based on limited SAR observations in just 2 years. In addition, the southern part of the Mekong Delta (latitude:  $8.5^{\circ}$ – $9.5^{\circ}$ N, longitude:  $105^{\circ}$ – $106^{\circ}$ E) appears similarly in the three maps, with high inundation frequency. There are water fields in this area where local farmers raise aquatic animals all along the year. Despite the similarity, there are differences among the three maps. Inundation frequency of the Tonle Sap Lake derived from the Landsat map (Figure 4.3a) is lower than that from the MODIS and SAR maps (Figure 4.3b,c). It may be due to the lack of free-cloud images observed by the optical Landsat sensors. For a region located to the east of the Mekong delta (latitude:  $10^{\circ}$ – $11^{\circ}$ N, longitude:  $106^{\circ}$ – $107^{\circ}$ E), it is dryer in MODIS and SAR inundation frequency maps than in Landsat map. The land use change due to urban development in this area could be the reason to explain the difference in the three maps.

### 4.3.2 Comparison with precipitation data

Time series of the MODIS-derived surface water extent are compared to satellite-based TMPA precipitation data. Relationships between surface water extent and local precipitation (over the Mekong basin lower than latitude 15°N) are presented in Figure 4.4, while Figure 4.5 shows comparisons with precipitation over the higher Mekong basin (higher than latitude 15°N). Local precipitation partly contributes to the variation of surface water extent. For example, less local rainfall than mean values during the wet season in 2010 contributed to a lower peak of surface water extent at the same time. The year after, local rainfall increased making an increase to the maximum of surface water extent. Peaks of local precipitation normally occur in Septembers, 2-3 weeks before the peaks of surface water extent. Their anomalies show some good agreements during wet years (2002, 2003, 2005, 2006 and 2011), as well as during dry years (2002, 2004, 2007, 2010, and 2012). Rainfall over the higher Mekong basin affects less the variation of surface water extent over the Mekong Delta than the local rainfall. Over the higher Mekong basin, precipitation reaches their maximum 1-month earlier, normally in Augusts, compared to the maximum of surface water extent in the Mekong Delta.

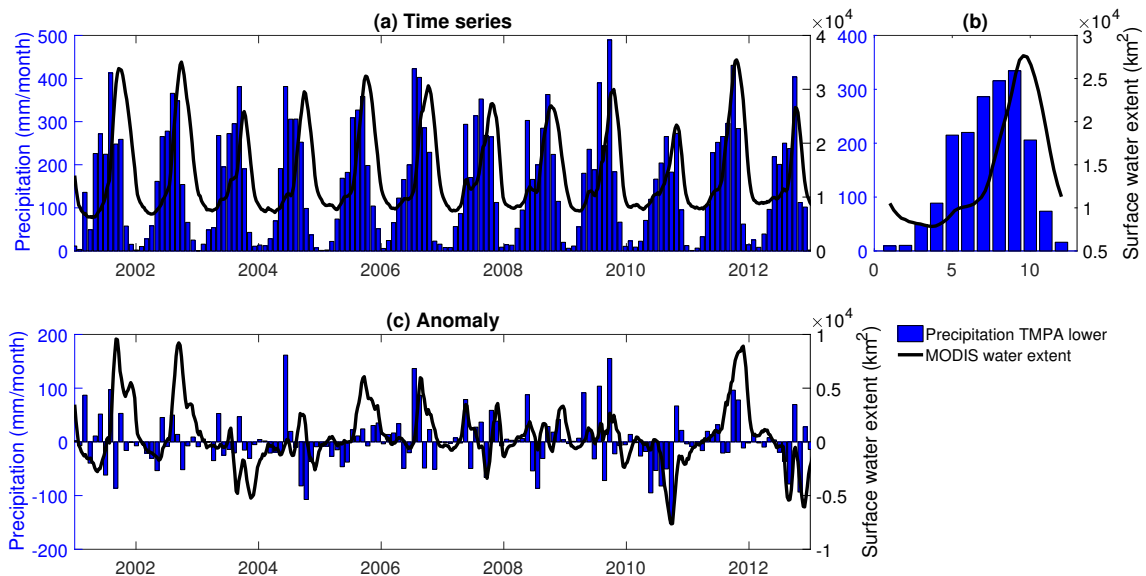


FIGURE 4.4: (a) Time series, (b) Monthly-mean annual cycle, and (c) Anomaly of MODIS-derived surface water extent over the Mekong Delta and Cambodia, and satellite-based precipitation over the Mekong basin located lower than latitude 15°N, over the 2001-2012 period.

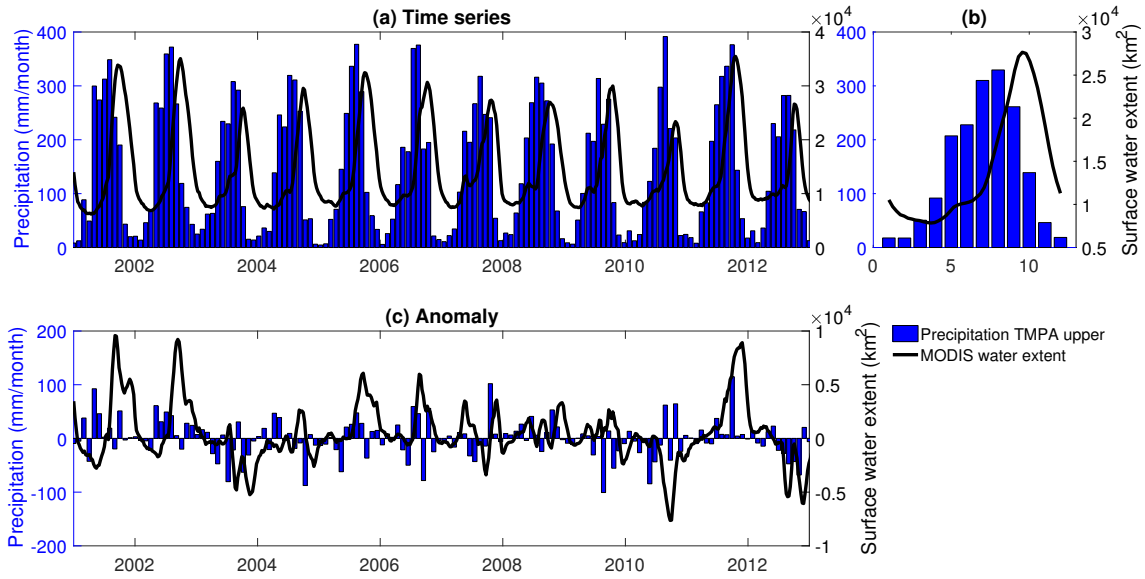


FIGURE 4.5: Similar to Figure 4.4, but satellite-based precipitation data cover the Mekong basin located higher than latitude 15°N.

There is no strong relationships between their anomalies, except for some strong events like the drought in 2012.

### 4.3.3 Comparison with in situ water level data

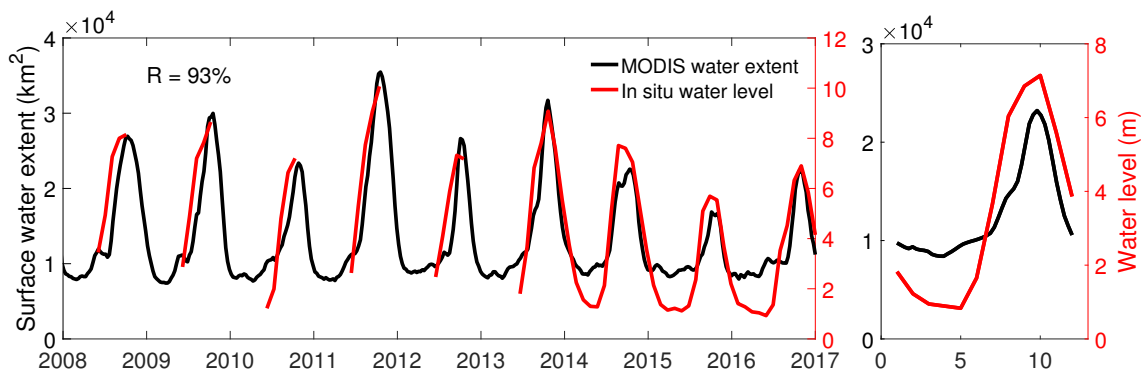


FIGURE 4.6: Left: Time series of MODIS-derived surface water extent over the Mekong Delta and Cambodia, and in situ water level at the Tonle Sap Lake gauge station, over the common period (2008–2016). Right: Monthly-mean annual cycle from the two time series.

Time series of the MODIS-derived surface water are compared to in situ water level data at the Tonle Sap gauge station, over the 2008–2016 period (Figure 4.6). Similar dynamics between surface water extent and in situ water level can be

observed for the common period, with a high linear temporal correlation (93%). Monthly-mean annual cycles extracted from the two time series are shown in Figure 4.6-right.

### 4.3.4 Comparison with the Multivariate ENSO Index

Figure 4.7 shows comparisons between time series and anomalies of the MODIS-derived surface water extent, TMPA precipitation and GIEMS-derived surface water extent in the lower Mekong basin, and the Multivariate ENSO index (MEI) over their common periods. Details of the Global Inundation Extent from Multi-Satellites (GIEMS) will be presented in Chapter 5. Relationships between the variation of MODIS-derived surface water extent in the lower Mekong basin and the MEI over the 2001-2016 period, are shown in Figure 4.7a,b. A linear correlation of 62% is found between MODIS-derived surface water extent anomaly and the MEI during a 5-year period (from May 2001 to June 2016). The effects of the strong 2014-2016 El Niño event on the lack of surface water extent in the lower Mekong basin is clearly observed. The link between surface water extent and the MEI

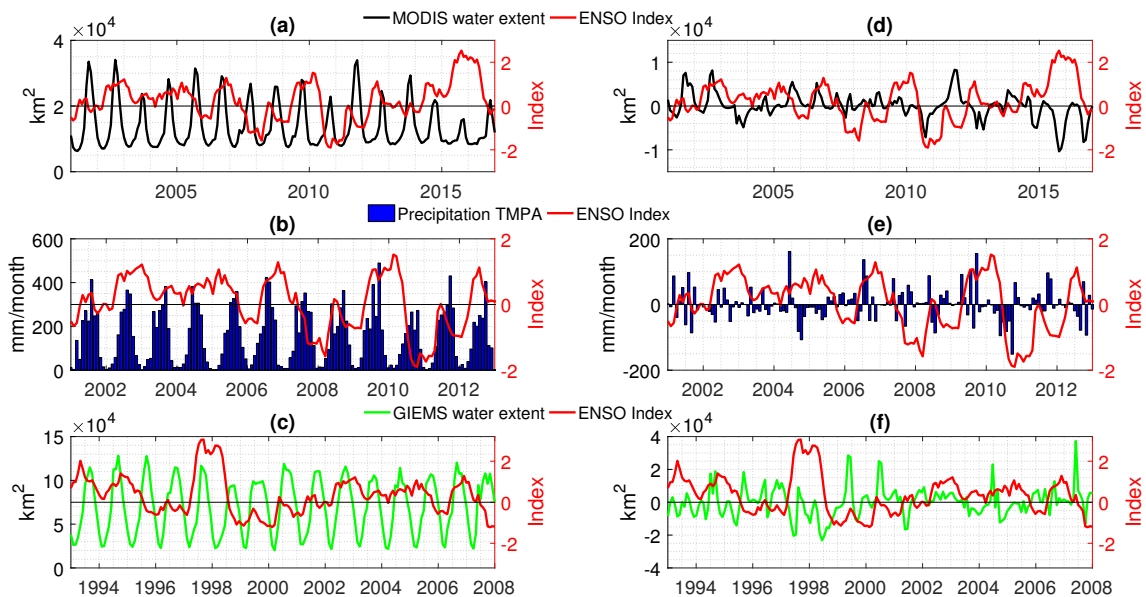


FIGURE 4.7: Comparisons between time series (left) and anomalies (right) of MODIS-derived surface water extent (a & d), precipitation (b & e) and GIEMS-derived surface water extent (c & f) over the lower Mekong basin, and the Multivariate ENSO Index, over their common periods.



is not obvious in other years when the ENSO is in its "Neutral" phase, as expected. Similar conclusion can be made between variation of local precipitation in the Mekong Delta and the MEI over the 2002-2012 period (Figure 4.7b,e). Updated TMPA precipitation data for recent years are not available to evaluate the effects of the 2014-2016 El Niño event on the variation of local precipitation in the Mekong Delta. Surface water extent in the Mekong basin, derived from the GIEMS dataset over the 1993-2007 period [Prigent et al., 2007, 2012], is also accessed to study the relationships between surface water extent and the ENSO in previous years. Similar to what have been observed during the 2014-2016 El Niño event, there was a decrease of GIEMS-derived surface water extent in the lower Mekong basin during the 1997-1998 El Niño event (Figure 4.7c,f).

## **4.4 Surface water height from altimetry data**

In this section, the principle of radar satellite altimeter is described first, then is applied to extract surface water height at all possible satellite virtual stations along the Mekong River and in the Mekong Delta. Similar to the surface water extent, the surface water height need to be evaluated by comparing to results from previous studies, and to variation of the in situ water level data at the closest gauge stations when available. Then, surface water height at virtual stations are linearly interpolated to build monthly maps of surface water height at 500 m spatial resolution to prepare for the calculation of the surface water volume variation in the next section.

### **4.4.1 The principle of radar satellite altimeter**

The principle of using radar satellite altimetry data to measure surface water height is well documented in the literature, and can be found elsewhere [Birkett, 1995, 1998; Fu and Cazenave, 2001; Crétaux and Birkett, 2006; Frappart et al., 2006a,b, 2017; Papa et al., 2015]. A summary is provided here (see Figure 4.8): the satellite altimeter transmits a microwave pulse in the nadir direction to the Earth surface, and receives the echo reflected back by the observed surface. The

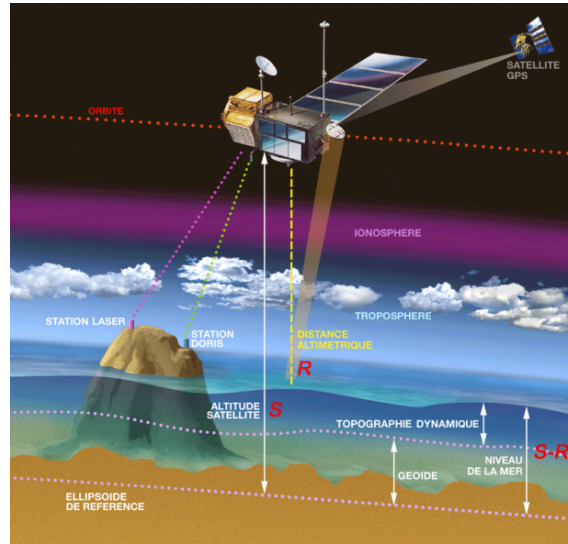


FIGURE 4.8: The principle of radar satellite altimetry to measure surface water height. Source: CNES.

round-trip time between the satellite and the Earth surface ( $t$ ) can be measured precisely by analyzing the returned signal. Assuming that the microwave pulse is propagating at the speed of light ( $c$ ), the distance between the satellite and the Earth surface ( $R$ ) can be calculated using equation 4.1

$$R = \frac{ct}{2} \quad (4.1)$$

Several systems (Laser Retro-Reflector, GPS, or DORIS) allow for an accurate measurement of the satellite altitude ( $S$ ), with respect to a reference ellipsoid. The surface water height ( $H$ ) with respect to the reference ellipsoid is determined by the difference between the satellite altitude ( $S$ ) and the altimeter range ( $R$ ), taking into account various corrections (see equation 4.2).

$$H = S - R - \Sigma Corrections \quad (4.2)$$

There are many factors needed to be corrected to take into account various physical phenomena in order to increase the accuracy of the satellite measurements. Instrument corrections have to be applied because the altimeter instruments are not perfect, and the satellite is moving when making measurements. Propagation corrections have to be applied because the electromagnetic waves are perturbed when they propagate and interact with the atmosphere. Atmospheric corrections

are also important to correct for the atmospheric dynamics. Surface state correction is another factor needed to take into account since the water surface is not always flat. Last but not least, geophysical corrections for the tides (solid earth, or polar tides) need to be applied to increase the accuracy of measurements.

#### 4.4.2 Satellite-based surface water height estimations along the Mekong River

Along the Mekong River main stream (from China to Vietnam), there is a total of 38 satellite virtual stations between the satellite ground tracks (from ENVISAT and SARAL) and the Mekong River (Figure 4.9). Using the Multi-mission Altimetry Processing Software (MAPS) provided by CTOH/LEGOS, ENVISAT/SARAL



FIGURE 4.9: Locations of 38 virtual stations between ENVISAT/SARAL ground tracks and the Mekong River. Blue presents VSs where altimetry data are good enough to extract surface water height. Figure is created on Google Earth.

satellite altimetry data over all the 38 VSs are processed using the Ice-1 retracker algorithm [Frappart et al., 2006a]. Among all 38 VSs, there are only 14 VSs where altimetry data are good enough to extract the surface water height of the water bodies (blue VSs numbered: 1-4, 7-11, 14-15, and 31-33). The accuracy of surface water level estimations over rivers strongly depends on the width and the morphology of the rivers and their banks [Papa et al., 2015]. Over the 24 other VSs, the river banks are too steep and the river widths are too narrow (smaller than 500 m) limiting the number of interactions between the satellite ground tracks and the Mekong River. As a consequence, altimetry data are too noisy to extract the surface water level information at these 24 VSs. For the 14 working VSs, most of them (11 VSs) are located under latitude 19°N where the river widths are large enough (larger than 1 km).

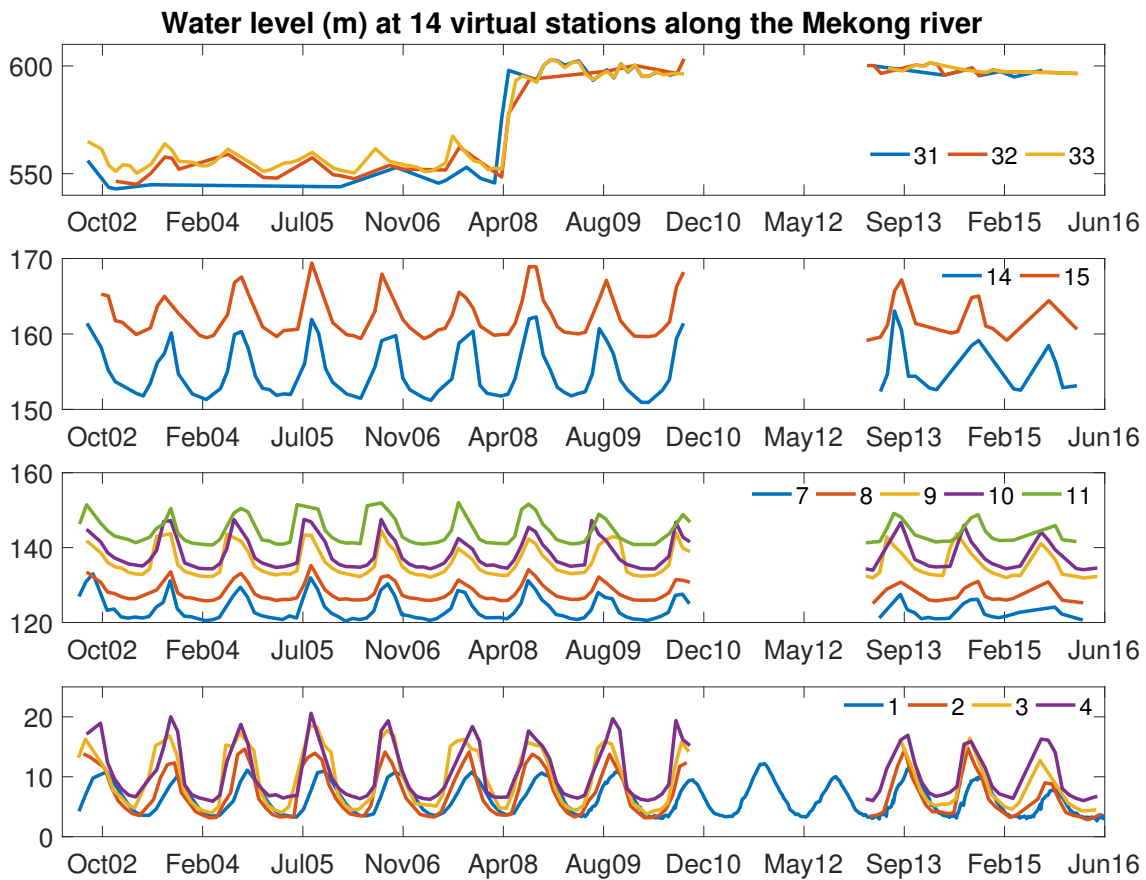


FIGURE 4.10: Water level time series derived from ENVISAT, SARAL and JASON-2 satellite altimetry at 14 VSs along the Mekong River, over the 2002-2016 period.

Surface water level time series at the 14 working VSs, over the 2002-2016 period, are shown in Figure 4.10. For most VSs (except the station numbered 1), there was a period without any data (from November 2010 to August 2013) as neither ENVISAT nor SARAL altimetry data were available. The VS numbered 1 (located over the Tonle Sap Lake) is the only VS where ENVISAT/SARAL and JASON-2 ground tracks overlap. By removing the bias in the reference ellipsoid between the three satellites, all satellite altimetry data can be jointly used to extract a continuous surface water level time series at the Tonle Sap VS, over the 2002-2016 period.

### **4.4.3 Evaluations of satellite-based surface water height data**

In this section, satellite-based surface water height time series at 14 working VSs (see section 4.4.2) are evaluated by comparing to results from other studies, and to in situ surface water level data at gauge stations when available.

#### **4.4.3.1 Comparison with other studies**

From Figure 4.10, reasonable surface water height variations and dynamics can be captured at the 14 working VSs. VSs 1-4 are located over the Tonle Sap Lake and near the intersections between the Tonle Sap Lake and the Tonle Sap river, have mean annual amplitudes range between 8-12 m. It is similar to findings in other altimetry studies over the Mekong basin [Frappart et al., 2006b]. VSs 7-11, and 14-15 are located in the center of Laos PRD. Annual surface water level amplitudes in these VSs vary between 10-12 m, and their water level peaks occur at the same time in a year (around September - October). VSs 31-33 are located 10-25 km above the Jinghong Dam in China (latitude: 22.05°N, longitude: 100.77°E), where the river width is rather small (~ 0.5 km). This small river width limits the number of intersection between the satellite tracks and the river. However, the most interesting finding is that the satellite altimetry data can capture an increase of water height (~ 50 m) in these three VSs after April 2008 when the Jinghong Dam started its operation. This finding is similar to results from Liu et al. [2016a]

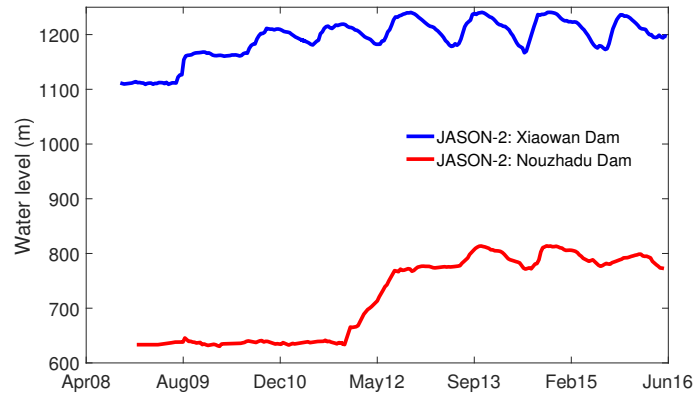


FIGURE 4.11: Water level time series for the 2008-2016 period, derived from JASON-2 satellite altimetry data at 2 VSs located  $\sim 20$  km and  $\sim 40$  km above the Manwan and Nouzhadu Dams, respectively.

where they used ENVISAT altimetry data and found the same impressive trend of water height above the Jinghong Dam after the first half of 2008.

The impact of dam and reservoir construction to the water height is also analyzed for two other Chinese dams: the Xiaowan Dam [latitude:  $24.70^{\circ}\text{N}$ , longitude:  $100.09^{\circ}\text{E}$ ] and the Nouzhadu Dam [latitude:  $22.40^{\circ}\text{N}$ , longitude:  $100.25^{\circ}\text{E}$ ]. JASON-2 altimetry data at 2 VSs located  $\sim 20$  km and  $\sim 40$  km, respectively, above these two dams are collected and processed. Figure 4.11 shows surface water height time series at these two VSs for the 2008-2016 period. For the first JASON-2 VS located  $\sim 20$  km above the Xiaowan Dam, a  $\sim 60$ -m increase (blue line) can be observed at the beginning of September 2009, just after the first generator was commissioned. Same conclusion can be found in Liu et al. [2016a]. For the second JASON-2 VS located  $\sim 40$  km above the Nouzhadu Dam, an increase of  $\sim 150$  m can be observed (red line) since the dam started operating its first generator at the beginning of 2012.

To conclude, my findings on the variation of the surface water height in the lower Mekong basin are consistent with previous results. Satellite altimetry data are not good enough when the river widths are too narrow (like rivers in China). However, satellite altimeters can capture the variations of the surface water height at VSs located 40-50 km above electrical dams.

#### 4.4.3.2 Comparison with in situ water level data

Daily in situ water level data at 8 gauge stations along the Mekong River are used for comparisons to satellite-derived surface water height data at the corresponding VSs (see details of these gauge stations in Table 4.2).

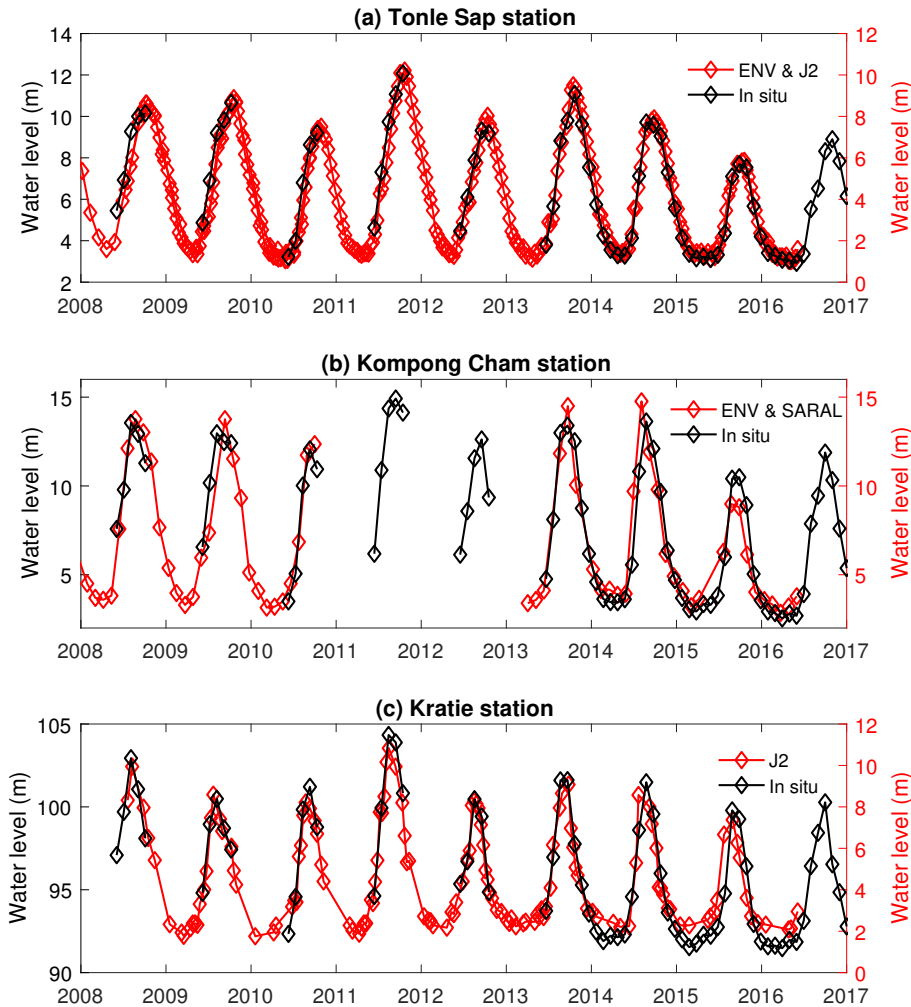


FIGURE 4.12: Water level time series derived from satellite altimetry, and in-situ measurements at three stations within Cambodia, for the 2008-2016 period.

Figure 4.12 compares surface water level time series derived from satellite altimetry data to that from in situ measurements, at three Cambodia stations (Tonle Sap, Kompong Cham, and Kratie), for the common 2008-2016 period. Note that altimetry-derived surface water level data are calculated with respect to a reference ellipsoid, therefore, there is no common height when comparing to in situ water level data, but very high temporal correlation and same amplitude of the

annual signal. The Tonle Sap station (Figure 4.12a) is the only location where ENVISAT/SARAL and JASON-2 VSs are close enough, and their data can be used jointly. The satellite VSs are located ~80 km upstream compared to the Tonle Sap gauge station. Satellite-derived surface water levels over the Tonle Sap VS show a very good agreement with in situ data, with similar dynamics in the peak-to-peak height variations over the common period. Similar to the Tonle Sap station, satellite-derived surface water level data at the Kompong Cham (Figure 4.12b) and the Kratie (Figure 4.12c) VSs agree very well with in situ data.

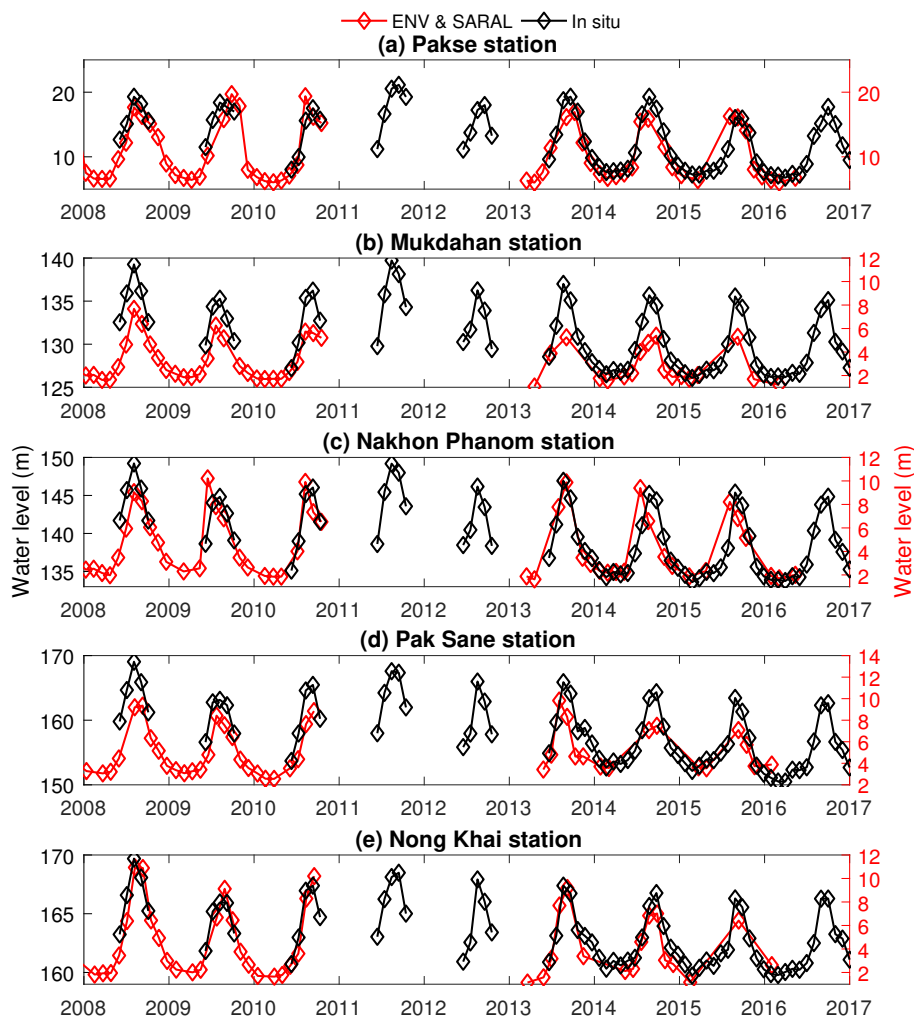


FIGURE 4.13: Water level time series derived from satellite altimetry, and in-situ measurements at five stations within Laos PRD, for the 2008-2016 period.

Figure 4.13 compares satellite-derived surface water height time series to in situ data at five locations located in Laos PRD (Pakse, Mukdahan, Nakhon, Pak Sane and Nong Khai), over the same period as already mentioned. JASON-2 data



are not available in this area due to the limitation of intersections between the satellite ground track and the Mekong River (see Figure 4.1-right). Because of the small river width of the Mekong River in Laos PDR, the number of intersections between the SARAL satellite ground track and the Mekong River is limited. As a consequence, surface water height information derived from the SARAL satellite are not regular compared to that derived from the ENVISAT satellite over the 5 VSs in Laos PDR.

Compared to in situ water level data, satellite-derived surface water height data are rather close, with the same seasonality and amplitude. ENVISAT-derived surface water height data fit better to the in situ data than SARAL-derived data. It is because the number of intersections between ENVISAT and the Mekong River at satellite VSs is much higher than that between SARAL and the river.

By comparing satellite-derived surface water height data to the results from previous studies by other authors, and to in situ water level observation at gauge stations, it is concluded that the processes I applied to extract river height information from satellite observations are correct. More comparisons between surface water extent over the Mekong Delta, and satellite-based surface water height over the Tonle Sap Lake will be shown in the next sections.

#### **4.4.4 Interpolated surface water height maps**

Monthly maps of surface water height over the Mekong delta and Cambodia can be estimated based on satellite altimetry surface water level data at VSs. Monthly time series of surface water level at 45 ENVISAT VSs are collected for the 2003-2009 period. Figure 4.14 shows locations of the 45 ENVISAT VSs in the Mekong delta and Cambodia. Due to the 35-day revisit time of the ENVISAT, some data are lacking during 1 month. Water level data between two consecutive months are interpolated to complete the water level dataset at all 45 VSs. At a given month, satellite altimetry surface water level data at the VSs are linearly interpolated over the inundated areas (from the corresponding MODIS-derived surface water extent map in Chapter 2) to create a surface water level map. Figure 4.15 shows examples of surface water level maps at 500 m spatial resolution in January, April,

July, and October 2003, respectively. Over the Tonle Sap Lake, water levels decrease from January to April, and then increase from July to October. Water levels over the southern part of the Mekong delta remain stable all along the year. Note that surface water level data are respected to the reference ellipsoid.

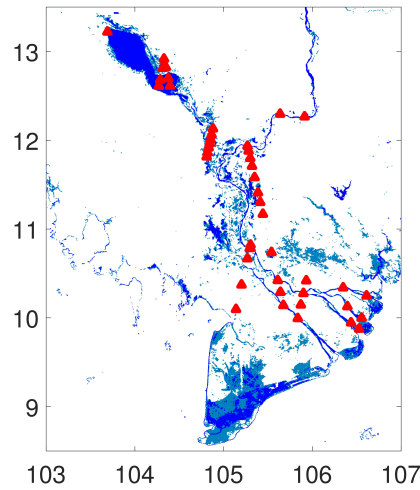


FIGURE 4.14: Locations of 45 ENVISAT VSs in the lower Mekong basin.

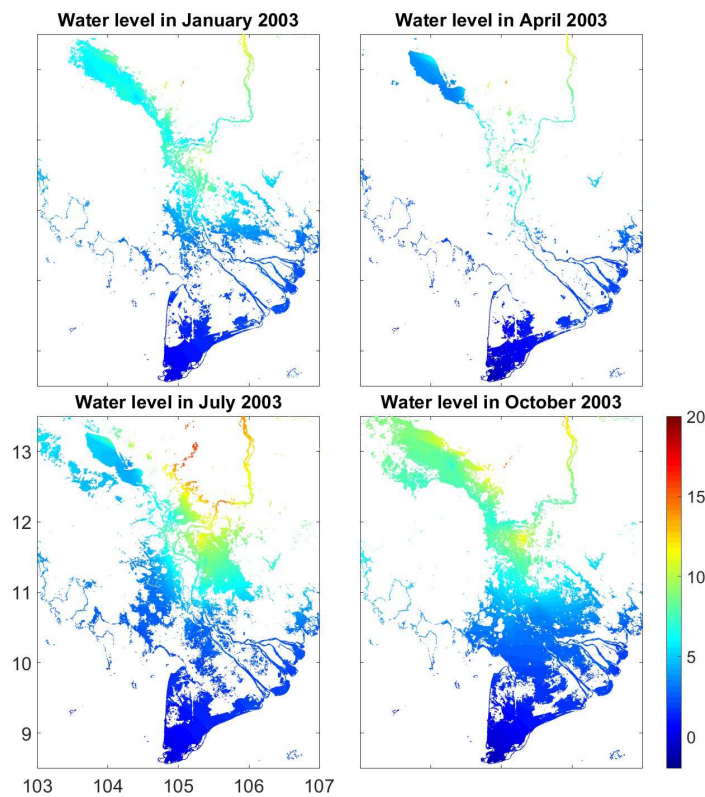


FIGURE 4.15: Surface water level maps (500 m spatial resolution) in the Mekong delta and Cambodia in January, April, July and October 2003, respectively.

## 4.5 Surface water volume variations

### 4.5.1 Comparison between MODIS-derived surface water extent and satellite-based surface water height

Time series of the MODIS-derived surface water extent and time series of surface water height at a satellite VS are compared to check the consistency between the two components. The Tonle Sap Lake VS (latitude: 12.40°N, longitude: 104.20°E) is chosen because the Lake is the largest water body in the lower Mekong basin. Figure 4.16 shows the two time series over the common period (2002-2015). MODIS-derived surface water extent and satellite-derived surface water height data at the Tonle Sap Lake show similar seasonal dynamics, with a linear temporal correlation of 90%. During rainy seasons, both surface water extent and water level data show low peaks in 2003, 2010, 2014 and 2015, as well as higher peaks in 2009, 2011 and 2013. MODIS-derived surface water extent normally reaches its maximum levels at the same time as the maximum states of the surface water height at the Tonle Sap Lake, but its minimum levels occur before that of the surface water height at the Lake. This is consistent with reality because during rainy seasons (June-October), the floodwater of the Lower Mekong River Delta flows directly into the Tonle Sap Lake through the Tonle Sap River, then the Tonle Sap Lake stores and slowly releases water to the Mekong Delta during dry seasons (November-May) [Cochrane et al., 2014].

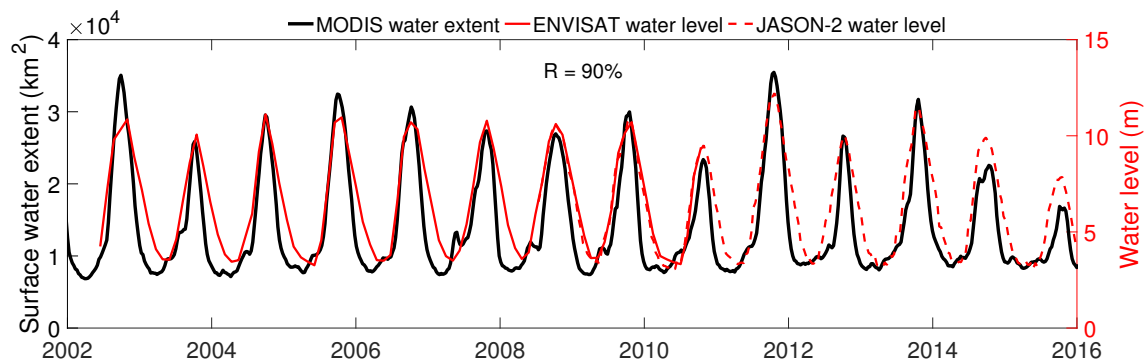


FIGURE 4.16: Time series of MODIS-derived surface water extent over the Mekong Delta and Cambodia, and satellite-based surface water height at the Tonle Sap Lake VS (latitude: 12.40°N, longitude: 104.20°E), over the common period (2002-2015).

## 4.5.2 Methodology

The variation of surface water volume in the Mekong delta and Cambodia corresponds to the difference of surface water heights integrated over the inundated areas (output of the Chapter 2). According to Papa et al. [2013]; Frappart et al. [2008, 2011], the variation  $\delta V(t_i, t_{i-1})$ , between two consecutive months numbered  $i$  and  $i - 1$ , over the floodplain  $S$ , are the sum of the products of the difference of surface water heights  $\delta h_j(i, i-1)$  with  $j = 1, 2, \dots$  inside  $S$ , by the elementary surfaces  $R_e^2 \sin(\theta_j) \delta\theta \delta\lambda$  and the percentage of inundation  $P_j$ :

$$\delta V(i, i - 1) = R_e^2 \delta\lambda \delta\theta \sum_{j \in S} P_j \delta h_j(\theta, \lambda, i, i - 1) \sin(\theta_j) \quad (4.3)$$

where  $\delta\lambda$  and  $\delta\theta$  are the sampling grid steps along longitude  $\lambda$  and latitude  $\theta$  ( $0.0045^\circ$ ), respectively, and  $R_e$  is the mean radius of the Earth (6378 km). The surface water volume variations are expressed in  $\text{km}^3/\text{month}$ . Results and comparisons with total water volume changes derived from the GRACE data are shown in the next section.

## 4.5.3 Results

The differences between two consecutive months of the surface water volume (blue) in the lower Mekong basin for the 2003-2009 period, calculated by equation 4.3, are shown in Figure 4.17. Compared to the differences between two consecutive months of the total land water volume (red) derived from GRACE data, they are in strong agreement ( $\sim 95\%$  of linear temporal correlation). Positive variations are observed from May to October, while negative variations are between November and April. By definition, the total water storage (TWS) is the sum of the surface water storage (SWS), the soil moisture and the groundwater. Therefore, the variation of soil moisture and groundwater (called "sub-surface water storage (Sub-SWS)") over the lower Mekong basin can be estimated by calculating the difference between the TWS and the SWS. Variation of the Sub-SWS is presented as the green line in Figure 4.17. Clearly, the contributions of the SWS and the Sub-SWS to the TWS are very close ( $\sim 50\%$  for each component). To better validate these results, monthly variation of the surface water volume in the

lower Mekong basin is evaluated with monthly variation of the in situ discharge data of the Mekong River over the 2003-2009 period (Figure 4.18). Monthly in situ discharge data are the sum of the discharge at the Tan Chau and the Chau Doc gauge stations. Monthly variation of the surface water volume agrees very well to monthly variation of the in situ discharge data of the Mekong River, with a linear temporal correlation of 96%.

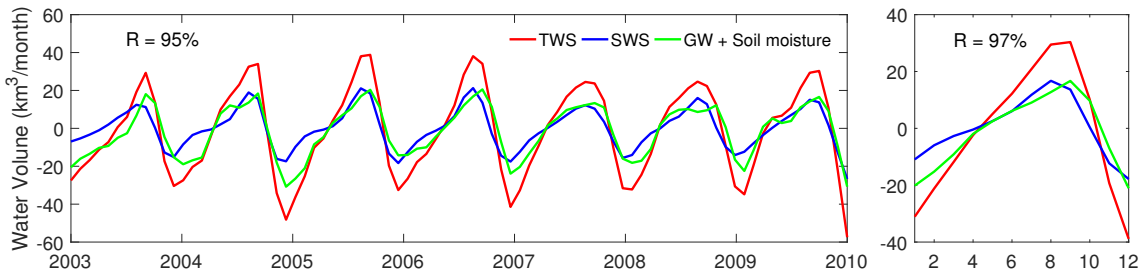


FIGURE 4.17: Left: Monthly variations of surface water volume change from ENVISAT radar altimetry and MODIS-derived surface water extent (blue), and monthly variation of the total land water volume change from GRACE data (red) in the lower Mekong basin over the 2003-2009 period. The difference, which represents the sum of soil moisture and groundwater, is represented by the green line. Right: Averaged over the 2003-2009 period.

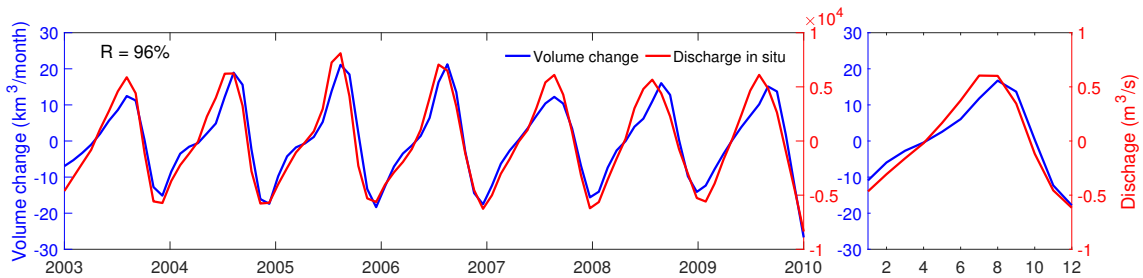


FIGURE 4.18: Left: Monthly variations of the surface water volume change from ENVISAT radar altimetry and MODIS-derived surface water extent (blue), and monthly variation of the in situ discharge of the Mekong River (red) in the lower Mekong basin over the 2003-2009 period. Right: Averaged over the 2003-2009 period.

## 4.6 Discussions and conclusions

This chapter focuses on the calculation of the variation of the surface water volume in the lower Mekong basin based on monthly variations of the surface water extent and the surface water height. The surface water extent are estimated

from optical MODIS observations (as presented in Chapter 2). It is then evaluated with SAR-derived surface water extent, precipitation and in situ water level data, as well as with an ENSO index. Results show high correlations between the MODIS-derived surface water extent and these datasets (94% with SAR-derived estimates, 93% with in situ water level data). Comparisons between the surface water extent and the MEI show that strong El Niño events can result in a reduction of the surface water extent compared to the normal state. Satellite-based surface water height data at satellite VSs are extracted from altimetry data, provided by CTOH/LEGOS. It is then evaluated by comparing to results in previous studies in the region, and to in situ water height estimates at the closest gauge stations, provided by the MRC. Satellite-based surface water height at VSs are linearly interpolated over inundated areas to construct maps of surface water height at 500 m spatial resolution. Surface water extent and surface water height data are compared directly to check their consistency before being used for surface water volume calculation. Results show a 90% linear temporal correlation between the two independent products. The monthly variation of the surface water volume over the 2003-2009 period is estimated based on a methodology described in [Frappart et al. \[2008, 2011\]](#). Results are evaluated with monthly variation of the total water volume derived from GRACE data, and with monthly variation of the in situ discharge data of the Mekong River over the same period. The surface water volume and the total water volume have similar seasonality, with 95% temporal correlation in their monthly variations. The surface water volume shares similar dynamics with in situ discharge estimates (96% temporal correlation).

Variations of groundwater and soil moisture could be estimated by removing the surface water volume from the total water volume. For validation, soil moisture variation from models (the ISBA-TRIP, for example) could be used. Then, it would be possible to quantify the contribution of the groundwater and soil moisture separately. Monitoring temporal and spatial variations of groundwater and soil moisture is also important because they are essential elements that drive land surface water. Information on their variations could be useful for agriculture applications, or drought prediction (like the 2014-2015 drought in the region due to a strong El Niño event).

Effects of dam constructions in Laos and China to the variation of surface

water height is also studied in this chapter. However, it cannot be done systematically because of the limitations of the satellite altimeters in this region. The river widths are too small ( $< 0.5$  km) and it limits the number of intersections between the satellite track and the river. As a consequence, it is impossible to extract good records of the surface water height there. However, satellite altimeters can capture significant increases of the surface water height at VSs located 40-50 km above electrical dams after they started their operations.

## **Chapter 5**

# **Comparisons between Global Terrestrial Surface Water Datasets**





## 5.1 Introduction

In the last three chapters, I focused on the use of satellite remote sensing data (visible/infrared, active microwave, and altimetry observations) for surface water extent and surface water height monitoring regionally over the lower Mekong basin. However, satellites can provide global observations. Therefore, in this chapter, the variation of global surface water extent is studied by comparing several global surface water datasets that derived mainly from passive and active microwave satellite data. Results from these comparisons are useful to the community in choosing the most suitable surface water product for global and regional studies at major river basins of the world.

Continental surface waters only cover a few percent of the land surface [Lehner and Döll, 2004; Downing et al., 2006; Prigent et al., 2007], but they have a strong impact on the environment, as well as on human life [Vörösmarty et al., 2010]. Surface freshwaters comprise wetlands, rice paddies, rivers, lakes, reservoirs, and episodically inundated areas. Note that the definition of wetlands varies with applications and there is not an overall consensus on the subject (e.g., Reichhardt [1995]). Surface waters play a key role in the biogeochemical and hydrological cycles, in biodiversity, and in climate variability. They show very diverse natures, from wetlands to inundated urban areas or rice paddies, associated with different dynamics from the tropics to the boreal regions. Wetlands are considered one of the most biologically diverse of all ecosystems since they support plant and wild animal species during important states of their life cycles. They are the world's largest natural source of methane ( $\text{CH}_4$ ), and they provide about one-third of the total global emission ( $\sim 165 \text{ Tg yr}^{-1}$ ; [Bousquet et al., 2006; Bridgman et al., 2013; Wania et al., 2013]). Consequently, monitoring surface freshwater extent and dynamics is a high priority in water management and climate research (e.g., Alsdorf et al. [2007]).

Nevertheless, our understanding about the global distribution of the surface waters and their dynamics is still limited, with only a few datasets providing information at the global scale. Efforts have been made to collect information on water surfaces to produce static maps of surface waters, for example, the Global Lakes and Wetlands Database (GLWD) from Lehner and Döll [2004] or the

[Matthews and Fung \[1987\]](#) wetlands dataset. These maps are representative of the maximum surface water extent, and they do not provide any information on the temporal dynamics. Satellite observations in the visible, infrared, or microwave domains have the potential to detect surface water extent and their variations ([Verpoorter et al. \[2014\]](#); [Yamazaki et al. \[2015\]](#); [Mueller et al. \[2016\]](#); [Feng et al. \[2016\]](#); [Pekel et al. \[2016\]](#)), with different degree of success depending on the environments. Optical and near-infrared satellite measurements provide good spatial resolution but are limited by their inability to penetrate clouds and dense vegetation. Microwave, passive or active, have the ability to penetrate clouds and vegetation, to a certain extent. With Synthetic Aperture Radar (SAR) data, high spatial resolution is obtained, but global products describing the surface water dynamics are not available yet. Passive microwave observations have long been used to detect surface water extents, but used alone, it is difficult to disentangle the vegetation contribution from the measured signal. [Prigent et al. \[2016\]](#) propose a review on the use of different satellite techniques to monitor surface water, discussing in detail their advantages and limitations regarding the diverse applications.

These considerations lead to the conclusion that there is currently not a unique satellite technique for detecting surface water dynamics globally from tropical to boreal regions. In the following, surface waters will include all surface water types (wetlands, rice paddies, rivers, lakes, reservoirs, and episodically inundated areas), as the satellite observations cannot distinguish between the different natures of the surface water. A multi-satellite methodology has been developed to derive surface water extent and dynamics at the global scale, benefiting from complementary strengths of satellite observations in the visible, passive and active microwave [[Prigent et al., 2001, 2007, 2012](#); [Papa et al., 2010](#)]: the Global Inundation Extent from Multi-Satellites (GIEMS). More recently, the Surface Water Microwave Products Series (SWAMPS) has been produced, also based on the merging of passive and active microwave satellite observations [[Schroeder et al., 2015](#)]. Long time series of global surface water estimates are necessary today to analyze the changes in the wetland-related methane emissions (e.g., [Melton et al. \[2013\]](#); [Wania et al. \[2013\]](#)), and climate modelers are in strong need of wetland extent information to understand the methane variability over past decades [[Ringeval et al., 2010](#); [Pison et al., 2013](#)] for a better prediction of its evolution

in the upcoming decades. GIEMS has been extensively evaluated and is limited to 1993-2007 (time extensions is underway). SWAMPS is a recent dataset that extends from 1992 to 2013. They have a similar spatial resolution ( $\sim 0.25^\circ \times 0.25^\circ$ ). As these two datasets are similar global surface water datasets with monthly time series, a thorough comparison of these estimates is needed for the user community.

In this chapter, global surface water datasets are systematically and objectively compared, including the two multi-satellite datasets, GIEMS and SWAMPS, along with two static datasets. The analysis covers the common period of the two satellite-derived products (1993-2007), and both the spatial and temporal variability of the databases are studied. The surface water datasets are described in section 5.2. Global and regional comparisons are described in section 5.3. A discussion is presented in section 5.4, and section 5.5 concludes this chapter.

## 5.2 Data

### 5.2.1 GIEMS

GIEMS was the first global surface water dataset that provided monthly distribution of wetland and surface water extent (including lakes, rivers, and irrigated agriculture). GIEMS data cover the period 1993-2007, and are mapped on an equal-area grid of  $0.25^\circ \times 0.25^\circ$  spatial resolution at the equator (pixels of 773 km<sup>2</sup>). Inundated surfaces were detected and their extent was estimated by the method developed by Prigent et al. [2001, 2007, 2012]. Global monthly inundation maps are derived from daily data of the following satellite observations: 1) passive microwave emissivity from the Special Sensor Microwave Image (SSM/I), 2) active microwave backscatter coefficients from the scatterometer on board the European Remote-Sensing (ERS) satellite, and 3) the normalized difference vegetation index (NDVI) derived from visible and near-IR reflectances of the Advanced Very High Resolution Radiometer (AVHRR). Instead of using directly the brightness temperature, surface emissivities are calculated to avoid modulation of the signal by atmospheric effects and surface temperature variations [Prigent

et al., 2006]. Passive microwave emissivities from SSM/I are primarily used to detect inundation of the land surface while active microwave backscatter is used to assess the vegetation contribution to the passive microwave signal. NDVI information is necessary to distinguish between bare surfaces and inundated surface in semi-arid regions where they can produce similar passive microwave signatures. Because of the lack of continuity and consistency in the ERS and AVHRR products, the current version of GIEMS uses monthly mean climatology of ERS and AVHRR, calculated over 1993-2000. Other solutions were carefully tested, such as using other instruments like QuikSCAT instead of ERS, but this was not satisfactory [Papa et al., 2010]. The snow-covered areas are filtered out using the National Snow and Ice Data Center (NSIDC) datasets ([Brodzik and Armstrong, 2013]; <http://nsidc.org/data/NSIDC-0046>). Inland seas (Caspian Sea and Aral Sea), big lakes (e.g., Great Lakes in North America), and coastal pixels (possibly contaminated by radiation from the ocean because of the large microwave fields of view) were also carefully suppressed from the GIEMS data. GIEMS has been thoroughly evaluated by comparisons with static surface water databases, and its consistency with other hydrological information (e.g., precipitation and river height) has been assessed [Papa et al., 2006, 2007, 2008a, 2010]. Recent works have been performed to downscale GIEMS using ancillary high-resolution data. GIEMS-D15 has a 15-arc-second resolution (nearly 500 m at the equator) for three different temporal states of the inundation extent [Fluet-Chouinard et al., 2015]. First, the original 12-year time series data of GIEMS (1993-2004) were aggregated to get mean annual minimum, mean annual maximum, and long-term maximum at the pixel level. Second, GLWD data [Lehner and Döll, 2004] were added to supplement missing data or to correct GIEMS underestimation for low water fraction. Finally, a global inundation probability map derived from the HydroSHED dataset [Lehner et al., 2008] was used to downscale GIEMS from the original resolution to 15-arc-second [Fluet-Chouinard et al., 2015]. In this chapter, GIEMS-D15 is averaged to the GIEMS grid (0.25° at the equator) for comparison with other datasets. More recently, a new version (GIEMS-D3) at 3-arc-second resolution [Aires et al., 2017] has been produced with an improved downscaling scheme.

## 5.2.2 SWAMPS

SWAMPS is a daily global surface water dataset, mapping open water areas and water under low-density vegetation, for the period 1992-2013 [Schroeder et al., 2015], on an equal-area grid of 25 km resolution at the equator (similar to GIEMS). It is derived from combined passive and active microwave observations: SSM/I and the Special Sensor Microwave Image/Sounder (SSM/I/S) for the passive microwave and the backscatter coefficient from ERS, QuikSCAT, and the Advanced Scatterometer (ASCAT). The microwave polarization difference index (MPDI), the ratio of the difference between the brightness temperatures in the two orthogonal polarizations over the sum, is the main indicator to detect the water bodies in this methodology. Backscatter coefficient from the ERS (5 GHz), QuikSCAT (10 GHz), and ASCAT (5 GHz) satellites are used to reduce the vegetation effect on the MPDI. The three instruments do not observe with the same angles and frequencies, and ad hoc corrections are performed to limit the effects of the changes on the time series [Schroeder et al., 2015]. Snow cover and frozen ground are filtered using the method developed by Grody and Basist [1996] and Chang et al. [1987]. Daily global SWAMPS data were averaged to obtain monthly global SWAMPS estimates at  $0.25^\circ \times 0.25^\circ$  for comparison with the other datasets.

## 5.2.3 GLWD

GLWD is a global open water and wetland dataset developed by Lehner and Döll [2004]. The dataset is derived from the combination of a variety of existing maps and information. Among these, the Digital Chart of the World (DCW) of the Environmental System Research Institute was the main source map to identify lakes and reservoirs. GLWD is not a satellite product, but a static map with three data levels: GLWD-1, GLWD-2, and GLWD-3. The GLWD-3 dataset used in this chapter includes three main types of open water (rivers, lakes, and reservoirs) and nine different natural wetland classes in the form of a global raster map at 30-second resolution. For each wetland category, (0%-25%, 25%-50%, 50%-100%), the averaged value (i.e., 12.5%, 37.5%, 75%) is used. GLWD is designed to present the maximum level of surface water extent regionally and

globally. GLWD has been used extensively and compared favorably with different satellite-based land-cover datasets [Nakaegawa, 2011]. For comparison with the other datasets, GLWD is aggregated to the GIEMS grid.

#### 5.2.4 Other ancillary datasets

Matthews and Fung [1987] wetland fractions come from aeronautical charts, the information for which is more likely acquired during warm seasons of maximum flooding [Matthews and Fung, 1987]. It is a static dataset on a  $1^\circ \times 1^\circ$  regular grid. It has been extensively used in the past by climate groups to estimate the methane emission from wetlands.

Satellite products such as GIEMS and SWAMPS detect all surface water, including inundated areas associated with cultivation (such as the rice paddies in Asia) that are not accounted for in GLWD dataset or in the Matthews and Fung [1987] wetland estimate. The global monthly irrigated and rain-fed crop areas (MIRCA) around the year 2000 [Portmann et al., 2010] provides information about irrigated and rain-fed agriculture globally. Irrigated rice is inundated during most of its growing season, and the MIRCA dataset provides an estimate of these inundated surfaces that are especially important in Asia [Adam et al., 2010]. MIRCA is averaged from the 5-arc-minute resolution to the common  $0.25^\circ$ -resolution grid of GIEMS dataset.

River discharge data in some important basins of the world (e.g., Amazon, Orinoco, Niger, Mississippi, Congo, Mekong, and Ob) are collected as an ancillary source of information to compare with time series of the surface water datasets. In this chapter, a total of 23 different basins are studied, but comparisons between wetland time series and river discharge over only 6 basins are shown, including Amazon, Orinoco, Niger, Ganges, Ob and Mississippi. For the Amazon River, I use the in situ monthly discharges observed at Obidos, Brazil, which is the closest gauge to the mouth of the river ( $\sim 800$  km from the ocean), and for which data are available for 1993-2007 at the Observation Service for the Geodynamical, Hydrological and Biogeochemical control of Erosion/Alteration and Material Transport in the Amazon, Orinoco, and Congo basins (SO HYBAM) website

(<http://www.ore-hybam.org/>). For the Orinoco River basin, river discharge data are also obtained from the HYBAM project, but available only for the 2003-2007 period. Daily river discharge at the Lokoja gauge (1998-2005) are used to get monthly river discharge data for the Niger River basin. The data are collected from the Global Runoff Data Center (GRDC; [http://www.bafg.de/GRDC/EN/Home/homepage\\_node.htm](http://www.bafg.de/GRDC/EN/Home/homepage_node.htm)). Over the Ganges River basin, river discharge data for the studied period (1993-2007) are obtained from a combination of in situ data and altimetry observations as in [Papa et al. \[2010\]](#). Over the Ob River basin, I use in situ river discharge records from the Russian Hydrometeorological Service that are available on a monthly basis until 2004 in the archives of the R-Arctic project (<http://www.r-arcticnet.sr.unh.edu/v4.0/index.html>). Finally, river discharge data over the Mississippi River basin are also collected from the GRDC database for the 1993-2006 period.

## 5.3 Comparisons of the surface water datasets

The satellite-derived surface water datasets are compared over their common period 1993-2007, first globally and then at basin scale. The 23 largest river basins in the world have been analyzed. However, the results are presented only for the lower Mekong basin, and 6 other river basins located in contrasted types of environments: the Amazon and the Orinoco River basins in the tropics of South America, the Mississippi River basin in North America, the Niger River basin in a semi-arid area of Africa, the Ganges River basin in South Asia, and the Ob River basin over the boreal region. The comparisons are also systematically performed with the two static datasets previously described: GLWD and the [Matthews and Fung \[1987\]](#) estimates.

### 5.3.1 Global comparisons

Figure 5.1a shows the GIEMS long-term monthly-mean maximum inundation for each pixel over the 1993-2007 period, along with the SWAMPS equivalent information (Figure 5.1b), for comparison with GLWD (Figure 5.1c). Even at this global



scale, large differences are evident between the three datasets. GIEMS and GLWD show much larger inland water fractions than SWAMPS. GLWD has particularly large inundation extent in Canada, where many small lakes are located. The major large river floodings (e.g., Amazon, Orinoco, and Ganges-Brahmaputra) appear clearly on both GIEMS and GLWD maps. The large water fractions in SWAMPS are concentrated on the coastal regions (see Indonesia or Japan, for example). That is very likely related to the contamination of the retrieval by the ocean. Close to the coastal line, part of the energy observed by the microwave instruments can come from the ocean, and the signal can be misinterpreted as coming from terrestrial inundated surfaces, if a careful filtering is not applied. Here I propose to filter the SWAMPS data to eliminate the ocean contamination close to the coast. Figure 5.1d represents the SWAMPS data where the contaminated coastal pixels are masked. SWAMPS also detects water almost everywhere on the globe, even in the North African desert. Histograms of the maximum fractional water surface are presented in Figure 5.2 for the four datasets in Figure 5.1. GLWD shows a large number of highly inundated pixels (>90%), mostly located in Canada (see Figure 5.1c). SWAMPS has a very large number of fractional water surfaces below 0.2, much more than the two other datasets. However, it has much less larger water fractions, especially after filtering of the coastal pixels.

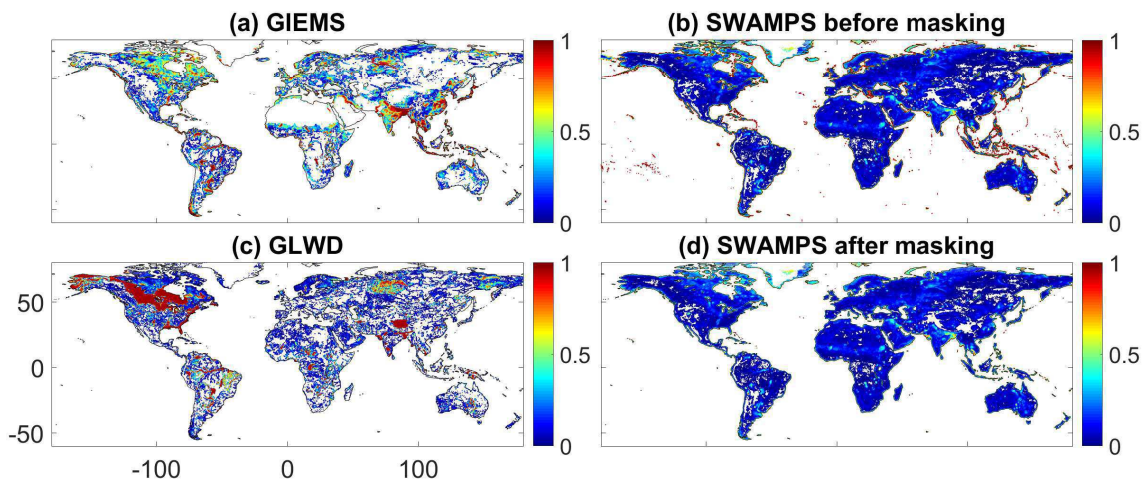


FIGURE 5.1: Global maps of fractional surface water for different datasets on the equal-area grid at  $0.25^\circ \times 0.25^\circ$  at the equator ( $773 \text{ km}^2$  pixels). (a) GIEMS long-term monthly-mean maximum over 1993-2007, (b) original SWAMPS long-term monthly-mean maximum over 1993-2007, (c) GLWD, and (d) SWAMPS long-term monthly-mean maximum over 1993-2007 after coastal filtering.

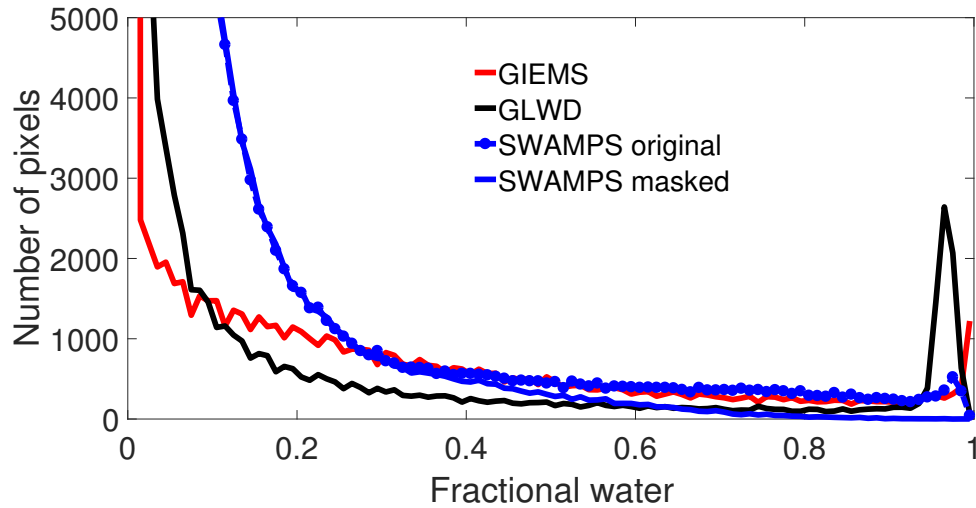


FIGURE 5.2: Histograms of long-term maximum fractional water surfaces for the four datasets in Figure 5.1.

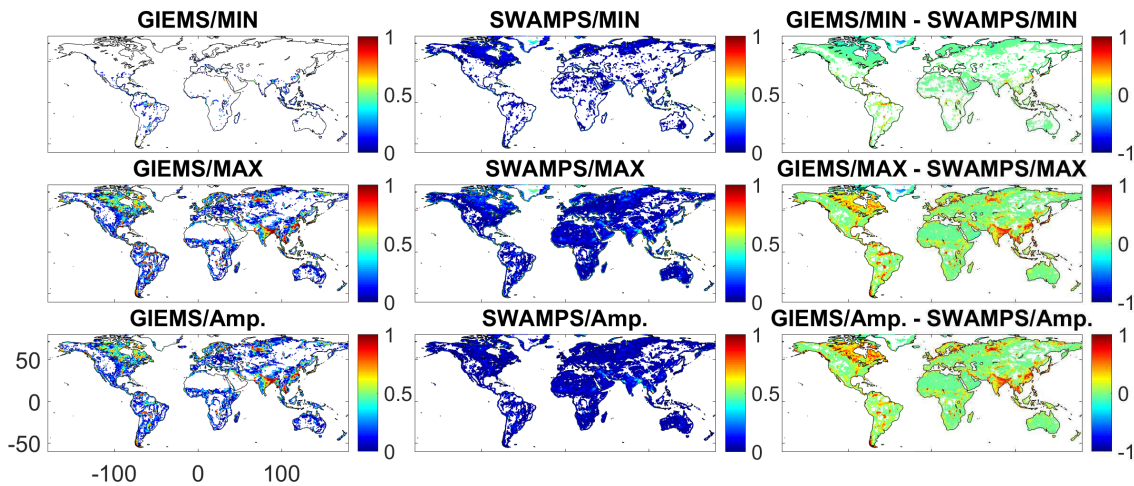


FIGURE 5.3: Global-mean annual minimum (top) and maximum (middle) of the fractional inundation and amplitude (bottom) for GIEMS (left), SWAMPS (center), and their differences (right). The information is presented on the 773 km<sup>2</sup> equal-area grid.

For each pixel and each satellite-derived dataset, the mean fractional inundation at annual maximum and minimum has been calculated, along with the mean yearly amplitude of the fractional inundation (see Figure 5.3; note that the coastal filtering is applied for SWAMPS). Maps of the differences of these values between GIEMS and SWAMPS datasets are also presented. Compared to GIEMS, SWAMPS shows very limited amplitude in the annual cycle of the inundation, even in regions where large seasonal variations are expected.

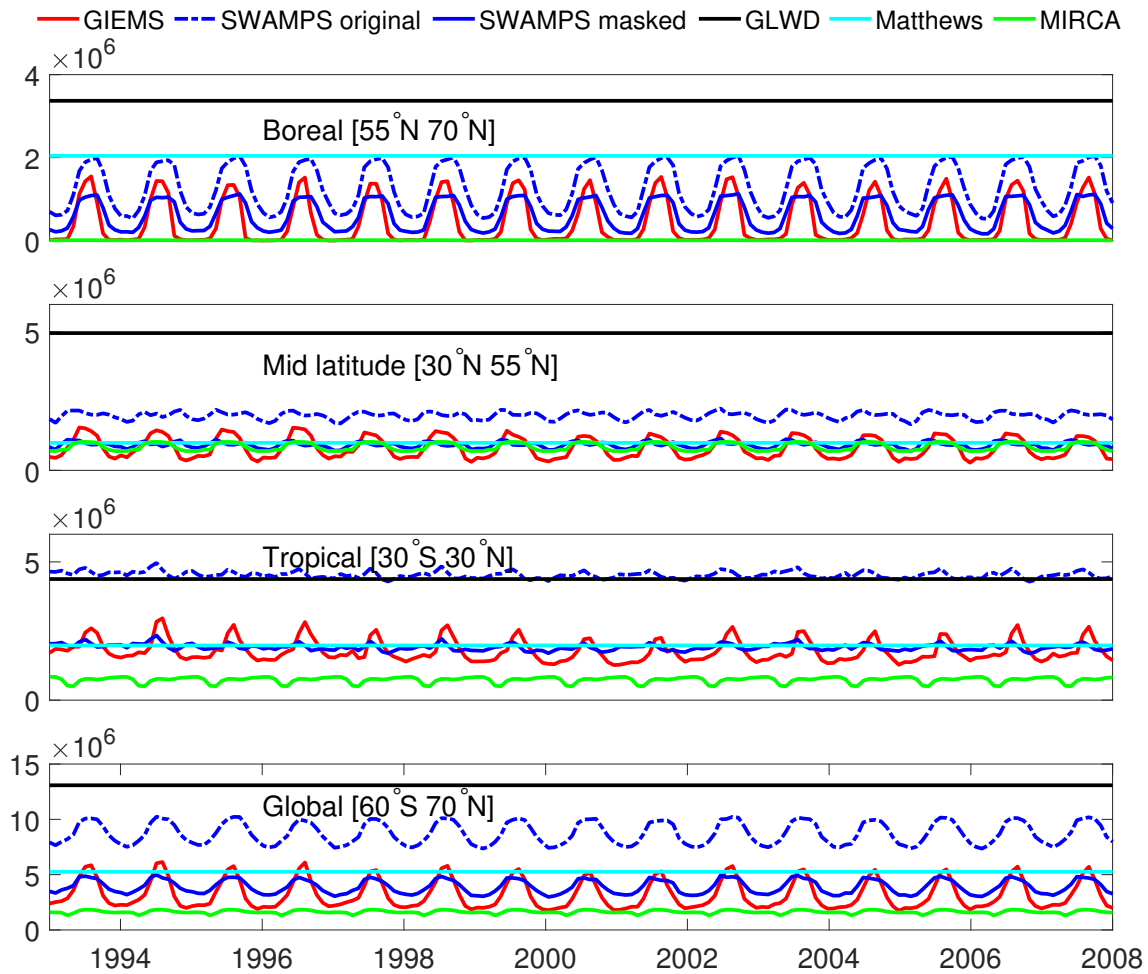


FIGURE 5.4: Time series of surface water extent derived from GIEMS (red), original SWAMPS (dashed blue), SWAMPS with coastal filtering (blue), GLWD (black), [Matthews and Fung \[1987\]](#) dataset (cyan), and MIRCA irrigated fields (green), at three latitude bands and globally, for the period 1993-2007.

Time series of the surface water from GIEMS and SWAMPS are compared globally and for three latitude bands [tropical (30°S-30°N), mid-latitude (30°-55°N), and boreal (55°-70°N)] in Figure 5.4, along with the corresponding values from GLWD, from the [Matthews and Fung \[1987\]](#) wetland dataset, and from the irrigated fields from MIRCA. GLWD (black) shows the maximum level of surface water both globally ( $\sim 13 \times 10^6$  km<sup>2</sup>) and regionally. This is expected as it is representative of the maximum inundation and it has a better spatial resolution than the [Matthews and Fung \[1987\]](#) dataset (that is also expected to present the maximum inundation, but with a much lower spatial resolution, and thus likely to miss the small water surfaces). The [Matthews and Fung \[1987\]](#) dataset (cyan)

has much less water extent in the boreal and the mid-latitude regions, as it is representative of the wetlands only and does not include the lakes and rivers ( $\sim 5.3 \times 10^6 \text{ km}^2$ ). The maximum global surface water extent derived from GIEMS is ( $\sim 6.2 \times 10^6 \text{ km}^2$ ), smaller than that derived from GLWD. From comparisons with high-resolution SAR estimates over the Amazon basin, the accuracy of the GIEMS products has been estimated at 10% of the  $773 \text{ km}^2$  equal-area grid: because of the low-spatial-resolution satellite sensors, GIEMS tends to miss the small water fraction below 10% of the pixels or to overestimate the large inundation fraction above 90% [Prigent et al., 2007]. However, this problem is partly corrected in GIEMS-D15 [Fluet-Chouinard et al., 2015], which merges GIEMS low-inundation pixels with GLWD. The maximum of the original SWAMPS inundation extent at global scale ( $\sim 10.3 \times 10^6 \text{ km}^2$ ) is close to the GLWD extent, as mentioned in Schroeder et al. [2015]. Over the tropical region, the maximum surface water derived from the original SWAMPS (dashed blue) is very close to that of GLWD (black),  $\sim 4.9 \times 10^6 \text{ km}^2$  and  $\sim 4.4 \times 10^6 \text{ km}^2$ , respectively. However, it is surprising to observe that the global maximum values is reduced to  $\sim 5 \times 10^6 \text{ km}^2$  (i.e., a reduction of nearly 51%) when the coasts are filtered out. Over the tropical region, SWAMPS decreases by nearly 53% after coastal masking to reach  $\sim 2.3 \times 10^6 \text{ km}^2$ . The original SWAMPS dataset detects more water surface than GIEMS dataset for all three latitude bands. After coastal filtering, the yearly-mean temporal SWAMPS water surface extent is similar to the GIEMS one, but the seasonal amplitude of SWAMPS is much lower than the GIEMS one. Table 5.1 shows monthly long-term minimum, long-term maximum, and long-term mean, as well as seasonal amplitude between monthly long-term minimum and maximum of surface water extent derived from GIEMS and SWAMPS after coastal filtering globally, and at three latitude bands, for the studied period (1993-2007). Long-term mean values between GIEMS and SWAMPS are close, but GIEMS shows larger amplitude than SWAMPS regionally and globally. From Figure 5.4 we can see that wetland surfaces derived from GIEMS (red) and SWAMPS (dashed blue) have strong temporal correlations globally (92%) and over the boreal region (88%), but lower temporal correlations over mid-latitude (58%) and over tropical areas (48%). Furthermore, in these two regions, GIEMS has a much stronger seasonal cycle than SWAMPS.

For the rest of this chapter, the coastal filter is applied to the SWAMPS dataset.

TABLE 5.1: Monthly long-term minimum, long-term maximum, long-term mean, and seasonal amplitude between monthly long-term maximum and minimum of surface water extent derived from GIEMS and SWAMPS after coastal filtering at three latitude bands and globally, for the period 1993-2007. Unit is  $\times 10^6$  km<sup>2</sup>.

	Min	Max	Mean	Amplitude
<b>Boreal</b>				
GIEMS	0.0	1.5	0.4	1.5
SWAMPS	0.2	1.2	0.6	1.0
<b>Mid-latitude</b>				
GIEMS	0.3	1.6	0.8	1.3
SWAMPS	0.7	1.2	0.9	0.5
<b>Tropical</b>				
GIEMS	1.3	3.0	1.8	1.7
SWAMPS	1.7	2.3	1.9	0.6
<b>Global</b>				
GIEMS	1.7	6.2	3.2	4.5
SWAMPS	3.0	5.0	3.9	2.0

## 5.3.2 Basin-scale comparisons

### 5.3.2.1 Comparisons over the lower Mekong basin

Long-term maximum inundation maps (1993-2007) over the lower Mekong basin, derived from GIEMS and SWAMPS, are shown in Figure 5.5. The distribution of surface water extent is very similar between the two datasets, with a spatial correlation of 83%, but surface water amplitude from SWAMPS is very limited compared to what observed from GIEMS. It is clear in Figure 5.6-top where the maximum surface water extent from SWAMPS is only one third of that derived from GIEMS, although the variations of surface water extent over the lower Mekong basin from GIEMS and SWAMPS share similar dynamics with high temporal correlations (96% for their time series, and 85% for their anomalies, respectively). The monthly-mean annual cycle from January to December is computed by averaging values from all Januaries to Decembers between 1993 and 2007. Then these values are subtracted from the time series for each given month, to obtain the anomaly time series. Combined monthly in situ river discharge data of the Mekong River in Vietnam (at Tan Chau and Chau Doc stations, from 2002-2007) are used to confirm the variation of surface water extent from the two datasets

over the lower Mekong basin. It is clear that GIEMS and SWAMPS capture correctly the seasonality of wetlands in this basin, compared to results derived from in situ measurements. Temporal correlations between GIEMS/SWAMPS and the discharge are around 80%, and it increases to more than 90% when calculated with 1-month lag.

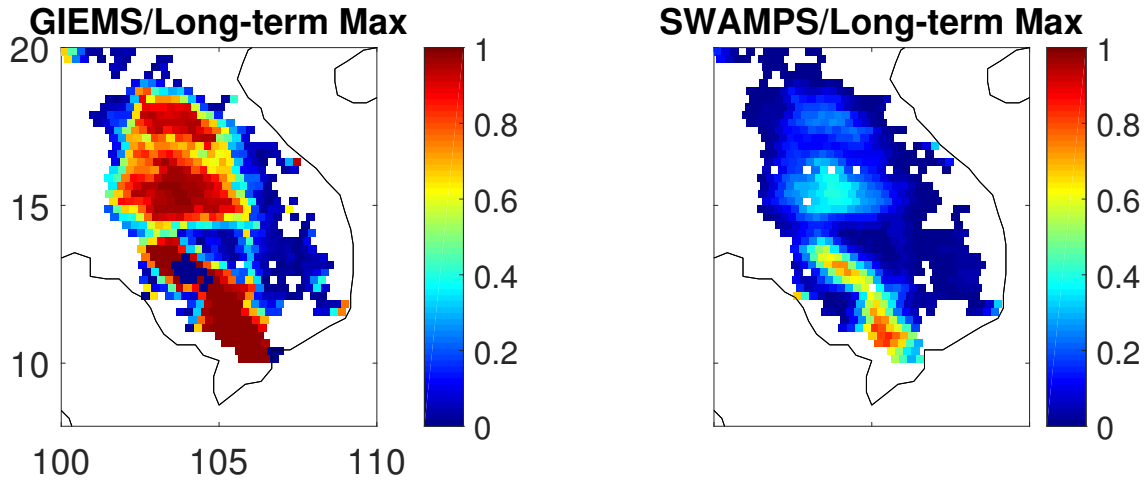


FIGURE 5.5: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS over the lower Mekong basin.

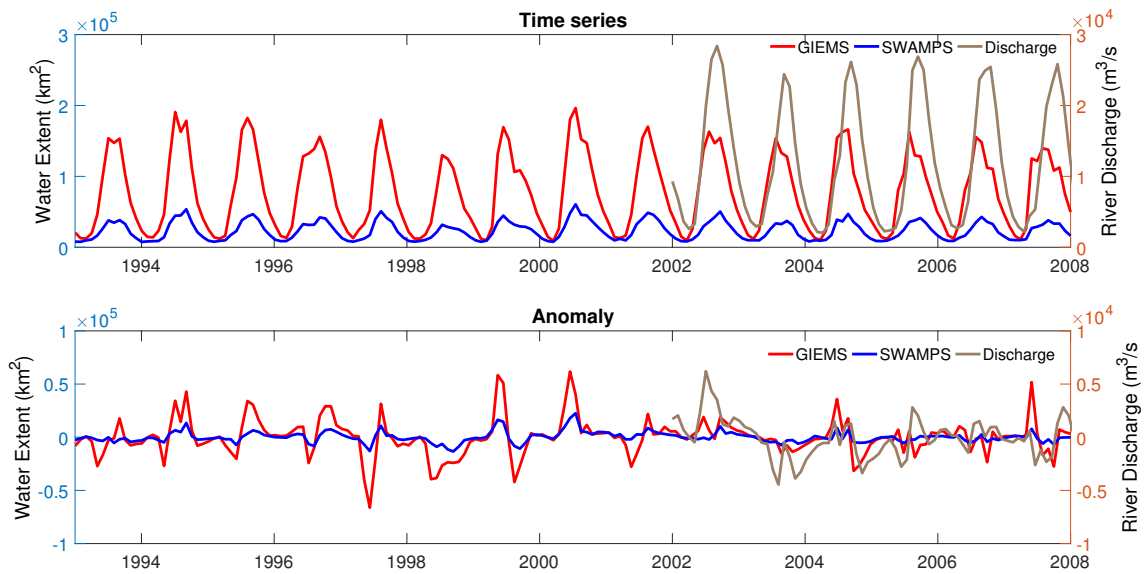


FIGURE 5.6: Time series (top) and anomaly (bottom) of surface water extent derived from GIEMS and SWAMPS (1993-2007), along with the river discharge (2002-2007), over the lower Mekong basin. Anomaly is calculated by removing the monthly-mean annual cycle from time series.

### 5.3.2.2 Comparisons over the Amazon basin

The Amazon basin is the largest drainage basin in the world with the largest discharge, and it is mostly located in the tropical rain forest. Figure 5.7 shows the long-term maximum inundation maps (1993-2007) over the Amazon basin, derived from the different wetland datasets. The spatial distribution of the surface water datasets are similar. Although the spatial correlation between long-term maximum inundated maps of GIEMS and SWAMPS is nearly 90%, SWAMPS fractional surface water is much lower than the GIEMS and GLWD ones. In [Schroeder et al. \[2015\]](#), it is noted that SWAMPS has problems detecting water underneath dense forest canopy. The ability of passive microwaves to detect surface water below dense forest was demonstrated early by [Giddings and Choudhury \[1989\]](#) or [Sippel et al. \[1994\]](#) in their pioneer works. It is rather surprising that SWAMPS cannot detect these surface waters. GIEMS-D15 corrects GIEMS by adding the small surface water fractions that are likely misses by GIEMS, and this is clearly seen when comparing GIEMS and GIEMS-D15 inundation maps in Figure 5.7.

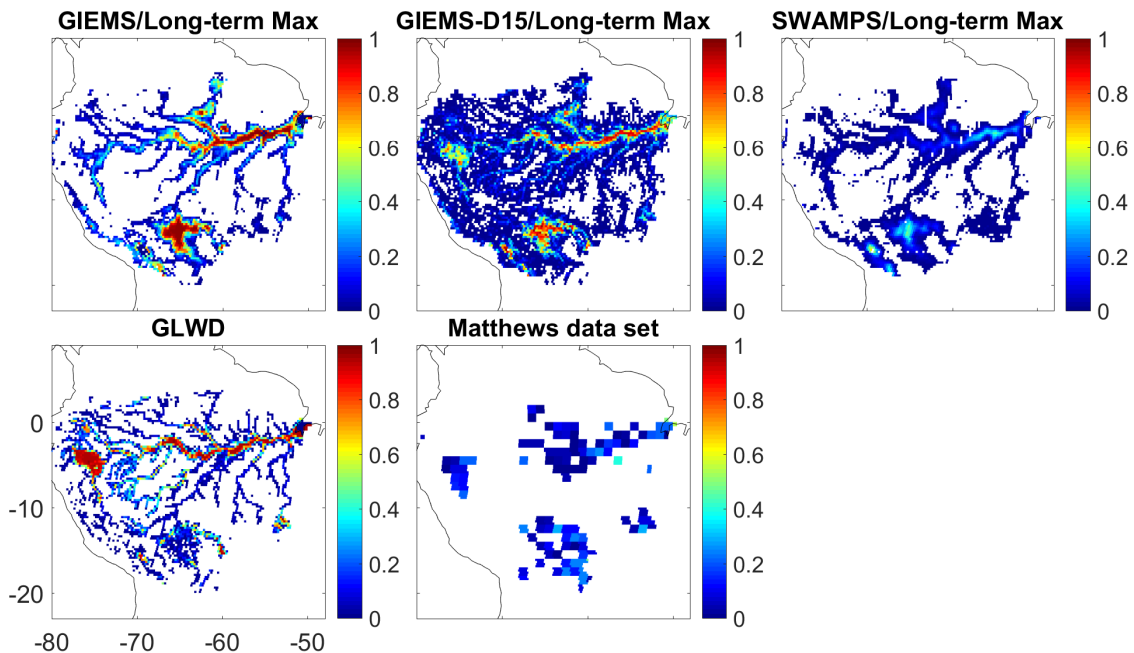


FIGURE 5.7: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS, and maximum inundation maps from the different surface water datasets over the Amazon basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \[1987\]](#) datasets.

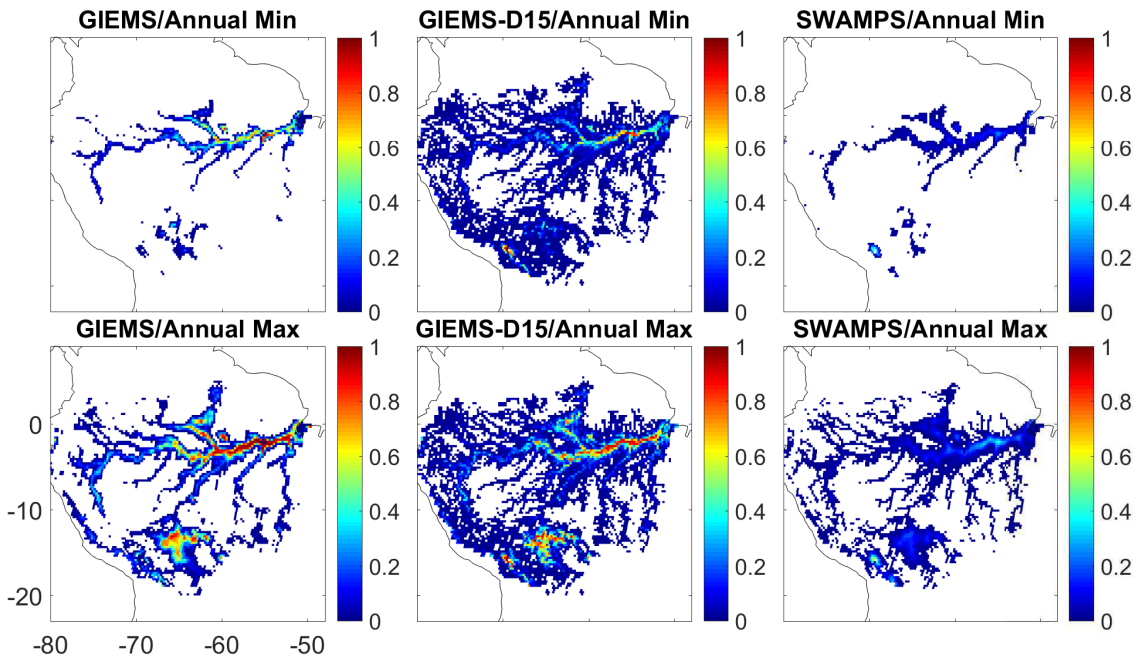


FIGURE 5.8: Mean annual minimum (top) and maximum (bottom) of the inundation derived from different surface water datasets over the Amazon basin from GIEMS (left), GIEMS-D15 (center), and SWAMPS (right).

Mean annual minimum and maximum inundation maps of GIEMS, GIEMS-D15, and SWAMPS are shown in Figure 5.8. Similar to the long-term maximum inundated maps, the spatial distributions of the GIEMS and SWAMPS datasets are similar (spatial correlation of 80% and 90% for the minimum and maximum, respectively), but SWAMPS detects much less surface water than the two GIEMS versions.

Monthly time series of the surface water extents have been calculated over the basin, along with the river discharge at the mouth of the river (Figure 5.9, top). Time series and anomaly correlations are shown in Table 5.2. The behaviors

TABLE 5.2: Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge  $Q$  over the Amazon basin for the period 1993-2007. Numbers in parentheses are calculated with 1-month lag between GIEMS/SWAMPS and  $Q$ .

	Time series correlation	Anomaly correlation
<b>GIEMS/SWAMPS</b>	95%	77%
<b>GIEMS/<math>Q</math></b>	78% (91%)	54% (58%)
<b>SWAMPS/<math>Q</math></b>	74% (88%)	57% (61%)



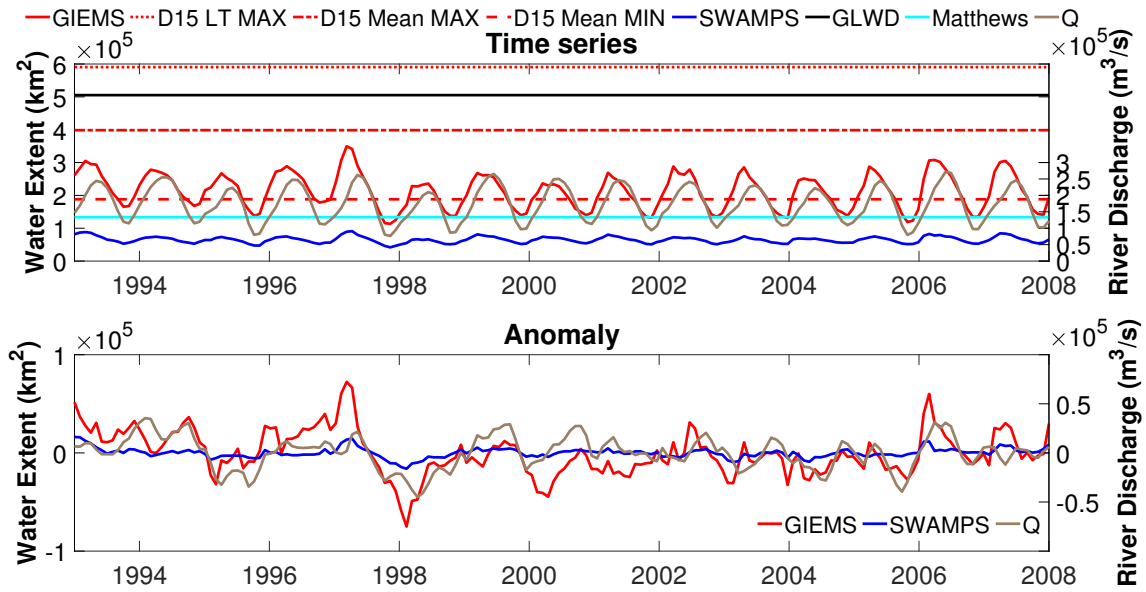


FIGURE 5.9: Time series (top) and anomaly (bottom) of surface water extent derived from the different wetland datasets, along with the river discharge over the Amazon basin from 1993 to 2007.

of time series of GIEMS (red) and SWAMPS (blue) are very similar to that of the river discharge (brown), with very high correlations. Time series correlation between GIEMS and SWAMPS is 95%. It is 78% and 74% with the river discharge for GIEMS and SWAMPS, respectively. Time series correlations with the river discharge increase when calculated with 1-month lag, as expected, reaching 91% and 88%, respectively. The anomaly correlation between GIEMS and SWAMPS is surprisingly high (77%).

### 5.3.2.3 Comparisons over the Orinoco basin

The Orinoco basin is also located in the South American tropical region, north of the Amazon basin. Similar analyses as over the Amazon basin are conducted in this river basin. Long-term maximum inundations maps for the 1993-2007 period are shown in Figure 5.10. Similar to what was observed over the Amazon basin, spatial correlations between these datasets are high (93% between GIEMS and SWAMPS), but SWAMPS fractional water is again much lower than the GIEMS and GLWD ones. Monthly time series and anomalies of the surface water extents of these datasets are shown in Figure 5.11, and their correlations are shown

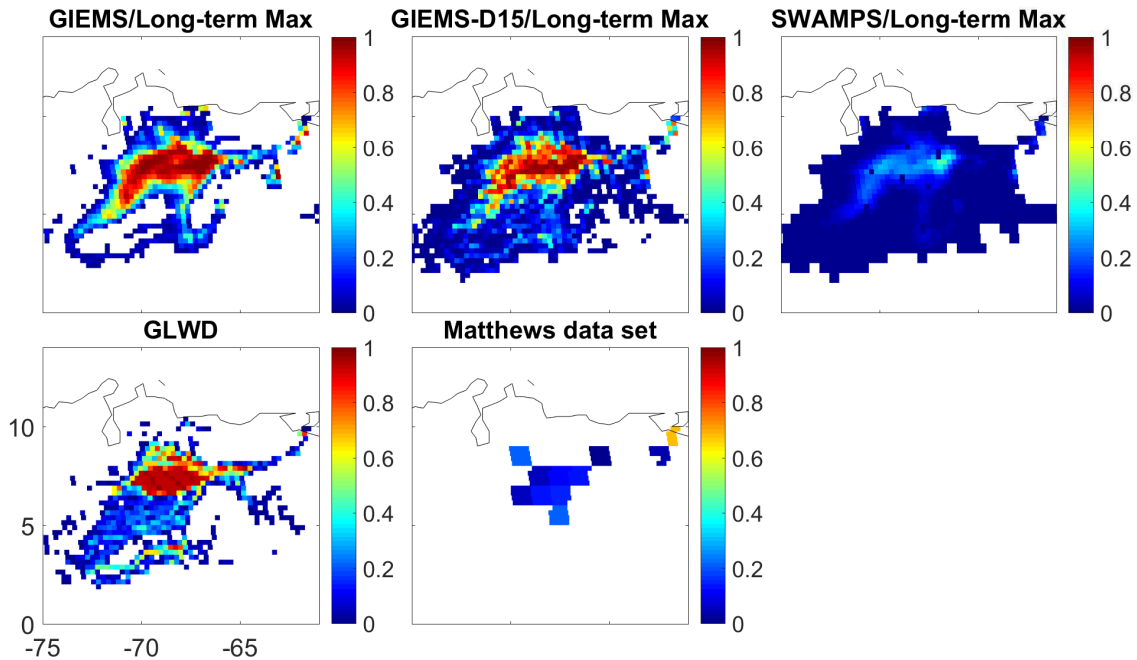


FIGURE 5.10: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS, and maximum inundation maps from the different surface water datasets over the Orinoco basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \[1987\]](#) datasets.

in Table 5.3. River discharge data for the 2003-2007 period (also obtained from the HYBAM project) are used to calculate time series and anomaly correlations. GIEMS and SWAMPS show exactly the same seasonal cycle for the 1993-2007 period, with an extremely high time series correlations (99%). Time series correlations between river discharge ( $Q$ ) and GIEMS/SWAMPS are 93% and 92%, respectively (for the common 2003-2007 period). Anomaly correlation between GIEMS and SWAMPS is even more surprising, being 97%, while it is only 62% and 60% between river discharge ( $Q$ ) and GIEMS/SWAMPS, respectively. Considering GIEMS and SWAMPS products were developed based on totally independent algorithms (although they used similar passive and active satellite observations as the input), it is suspected that this extremely high anomaly correlation should be almost impossible to achieve without fine tuning of the SWAMPS algorithm to the GIEMS temporal structure over this river basin.

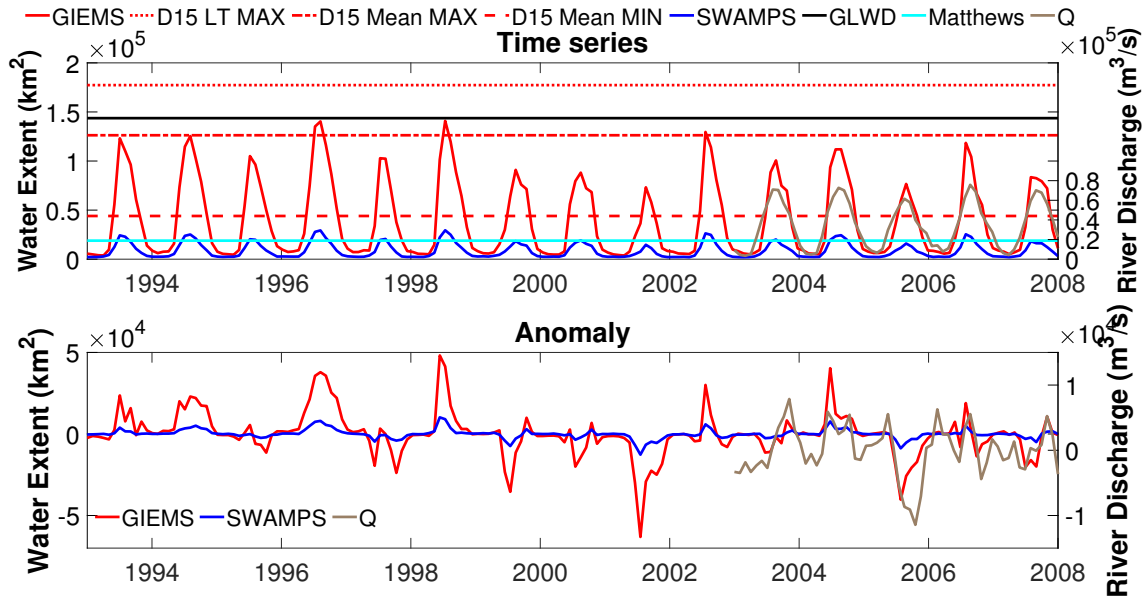


FIGURE 5.11: Time series (top) and anomaly (bottom) of surface water extent derived from the different wetland datasets (1993-2007), along with the river discharge (2003-2007) over the Orinoco basin.

TABLE 5.3: Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge Q over the Orinoco basin. Q is available for the 2003-07 period.

	Time series correlation	Anomaly correlation
<b>GIEMS/SWAMPS</b>	99%	97%
<b>GIEMS/Q</b>	93%	62%
<b>SWAMPS/Q</b>	92%	60%

### 5.3.2.4 Comparisons over the Niger basin

The Niger basin is characterized by a large inner delta that results in a region of braided streams with seasonal floods. Long-term maximum inundation maps over the Niger basin are shown in Figure 5.12, while their minimum and maximum are presented in Figure 5.13, and Figure 5.14 shows their time series and anomalies, as well as that derived from the river discharge data. Time series and anomaly correlations between GIEMS/SWAMPS and other ancillary datasets are shown in Table 5.4. Time series of GIEMS and SWAMPS are in opposite phase (Figure 5.14), making the time series correlation negative (-40%, see Table 5.4). Again, GIEMS shows a much stronger seasonal cycle than SWAMPS over the Niger River basin. GIEMS (red) and the river discharge (brown) show similar behavior with a time series correlation of nearly 81% (for the common 1998-2005

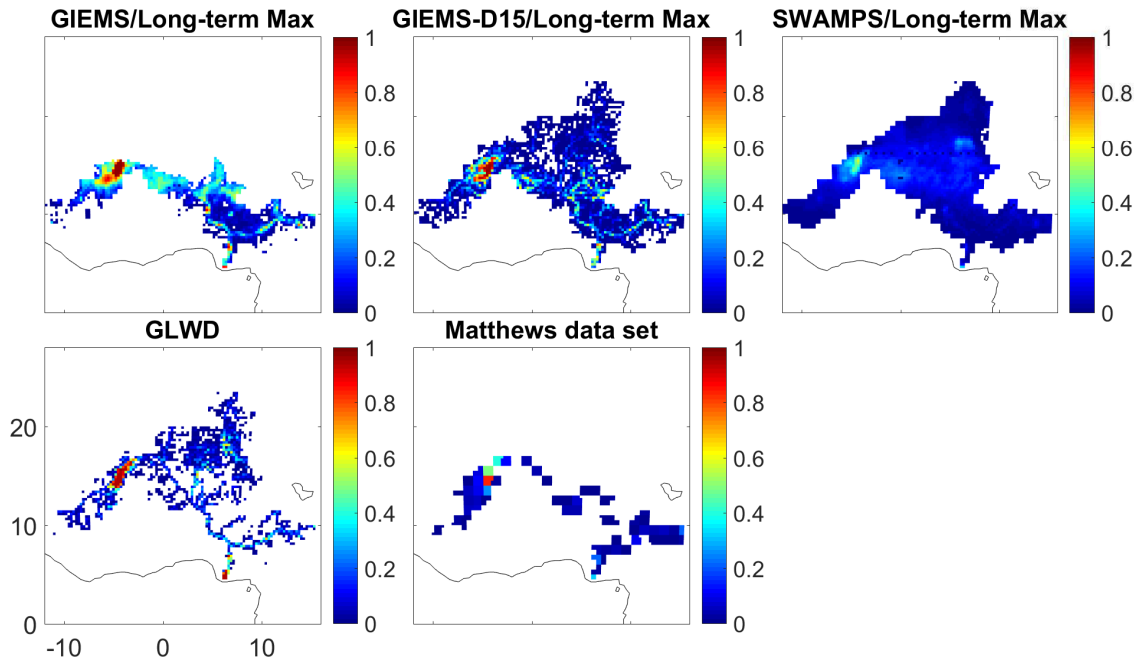


FIGURE 5.12: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS, and maximum inundation maps from the different surface water datasets over the Niger basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \[1987\]](#) datasets.

period). In contrast, SWAMPS does not capture the same seasonal cycle, making its time series correlation with the river discharge negative (-60%, see [Table 5.4](#)). Note that other studies have analyzed GIEMS over the Niger River basin, for example, [Pedinotti et al. \[2012\]](#) or [Aires et al. \[2014\]](#). [Pedinotti et al. \[2012\]](#) evaluated the ability of the ISBA-Total Runoff Integrating Pathways (TRIP) continental hydrologic system to represent key processes (surface water, rivers and floodplain dynamics, and water storage) related to the hydrological cycle of the Niger River basin. To this end, GIEMS is used to evaluate the long-term simulations which showed that the flooding scheme leads to a non-negligible increase of evaporation over large flooded areas, which in turns improved the Niger River discharge estimates at several locations. The objective of [Aires et al. \[2014\]](#) is to develop downscaling methodology to obtain a long time record of inundation extent over the inner Niger River delta at high spatial resolution (500 m) based on the existing low-spatial-resolution results of the GIEMS dataset and observations from MODIS.

Time series of the surface water derived from MODIS visible images over the

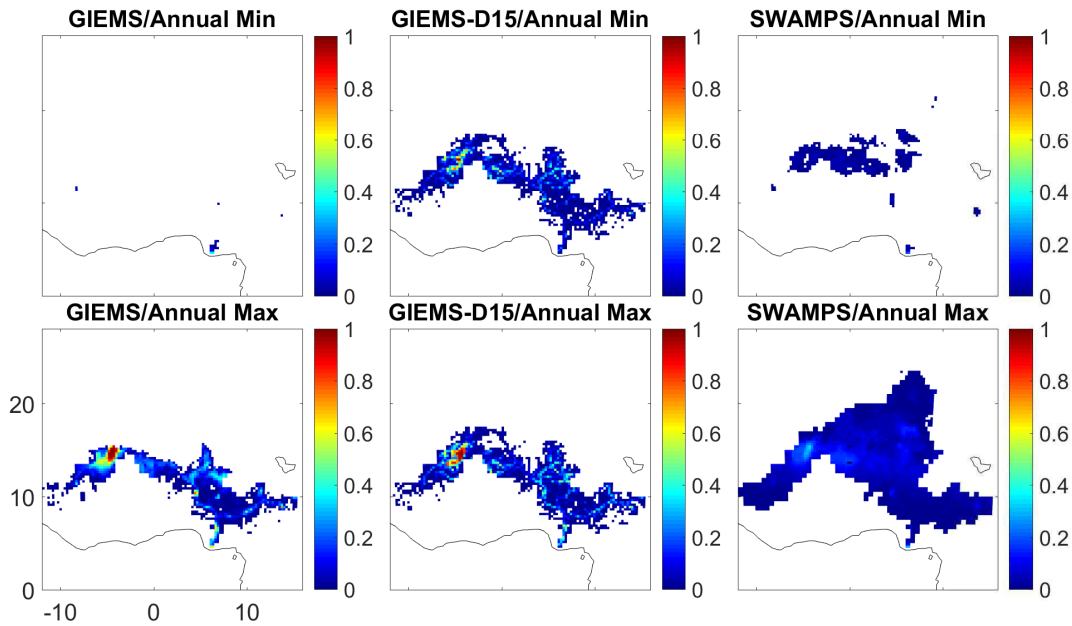


FIGURE 5.13: Mean annual minimum (top) and maximum (bottom) of the inundation derived from different surface water datasets over the Niger basin for GIEMS (left), GIEMS-D15 (center), and SWAMPS (right).

TABLE 5.4: Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge  $Q$  over the Niger basin for the period 1993-2007. Numbers in parentheses are calculated with 2-month lag between GIEMS/SWAMPS and  $Q$ .

	Time series correlation	Anomaly correlation
<b>GIEMS/SWAMPS</b>	-40%	20%
<b>GIEMS/<math>Q</math></b>	81% (51%)	-3% (44%)
<b>SWAMPS/<math>Q</math></b>	-60% (10%)	-2% (12%)

Niger River basin for the 2000-2007 period [Bergé-Nguyen and Crétaux, 2015] are also compared to the behavior of GIEMS and SWAMPS over this region. From Figure 5.14 (top) and Figure 5.15, it is clear that GIEMS and MODIS surface water time series have similar seasonal dynamics over the common period (2000-2007). However, GIEMS has higher maximum values than MODIS, which could suggest an overestimation from GIEMS over the Niger River basin. In addition, the interannual variability is not totally similar between GIEMS and MODIS. Similar passive microwave signatures can be observed over arid regions and over water: these two surface types have low emissivities with rather large emissivity polarization differences. As a consequence, reliable and accurate detection of surface water in arid and semiarid regions is not trivial. In GIEMS dataset, NDVI

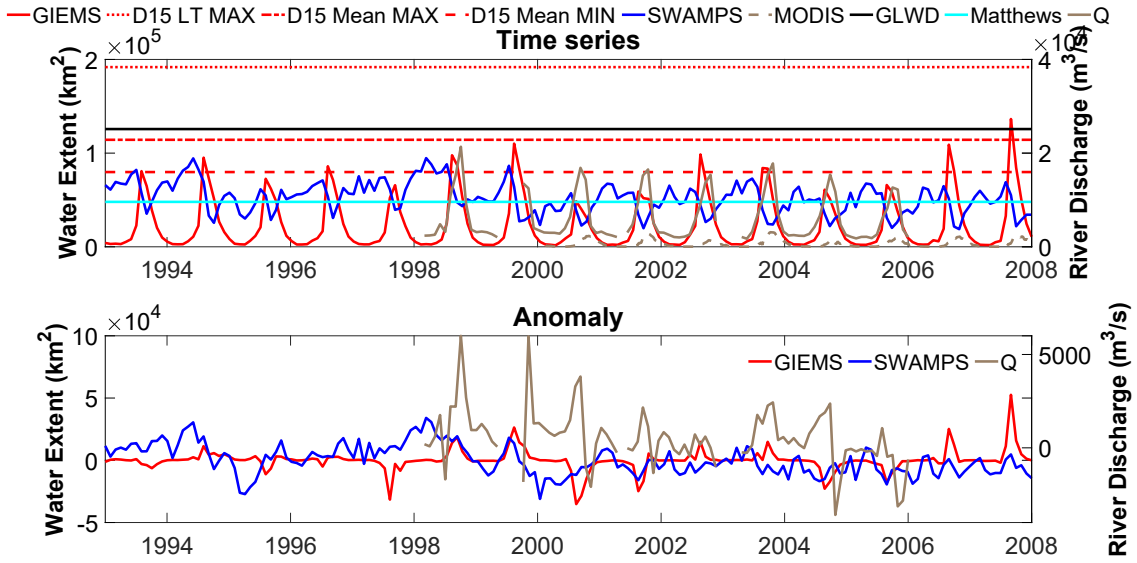


FIGURE 5.14: Time series (top) and anomaly (bottom) of surface water extent derived from the different wetland datasets (1993-2007), along with the river discharge (1998-2005) over the whole Niger basin. MODIS surface water extent time series for the inner delta only (dashed brown) is available from 2000.

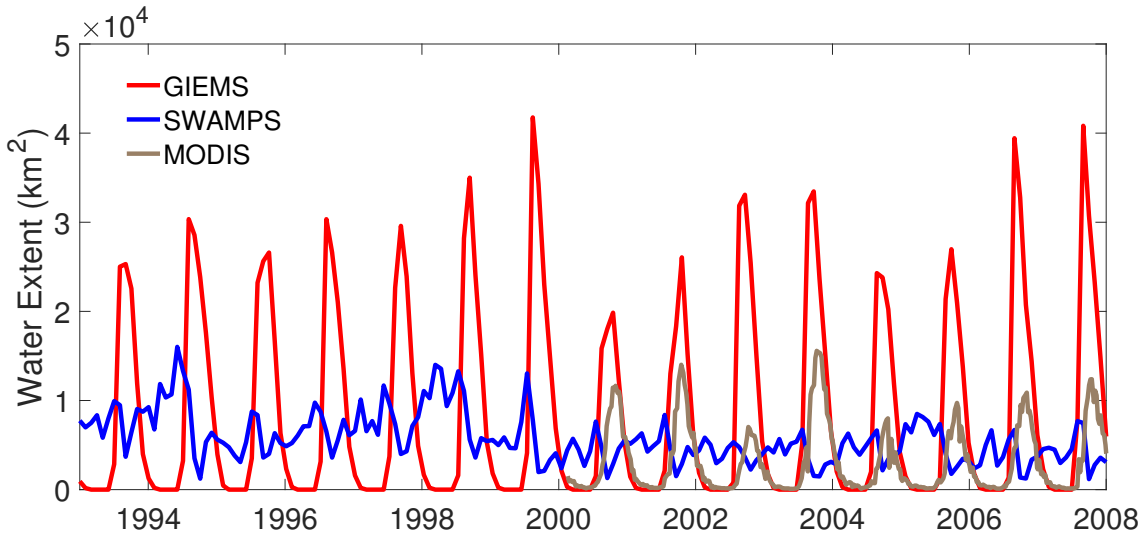


FIGURE 5.15: Time series of surface water extent derived from GIEMS, SWAMPS, and MODIS for the inner Niger delta.

information are used in the process to help solve these ambiguities. SWAMPS obviously encounters difficulties in this type of environment, with false detection of water in deserts and underestimation of water in inundated deltas. This is in agreement with [Schroeder et al. \[2015\]](#). As a result, over the Niger River basin, SWAMPS does not capture at all the water surface dynamics.

### 5.3.2.5 Comparisons over the Ganges basin

The Ganges basin is part of the Ganges-Brahmaputra-Meghna basin, located between India, Nepal, and China. It is one of the world major floodings. Figure 5.16 shows the long-term maximum inundation maps derived from different datasets, while their time series and anomalies are presented in Figure 5.17. Over the Ganges River basin, GIEMS surface water extent time series (red) is higher than the maximum surface water extent derived from GLWD dataset (black), as GIEMS also detects inundated areas associated to cultivation activities, such as rice paddies, which are not taken into account in GLWD. SWAMPS surface water extent (blue) is much lower, compared to both GIEMS and GLWD datasets. However, GIEMS and SWAMPS have a high time series correlation (87% in Table 5.5). River discharge data collected from the Ganges River basin also show high time series correlation with GIEMS and SWAMPS, being 83% and 64%, respectively. The same high correlation is observed when comparing GIEMS and SWAMPS anomalies data (74%), but it is much lower when calculated with river discharge data (Table

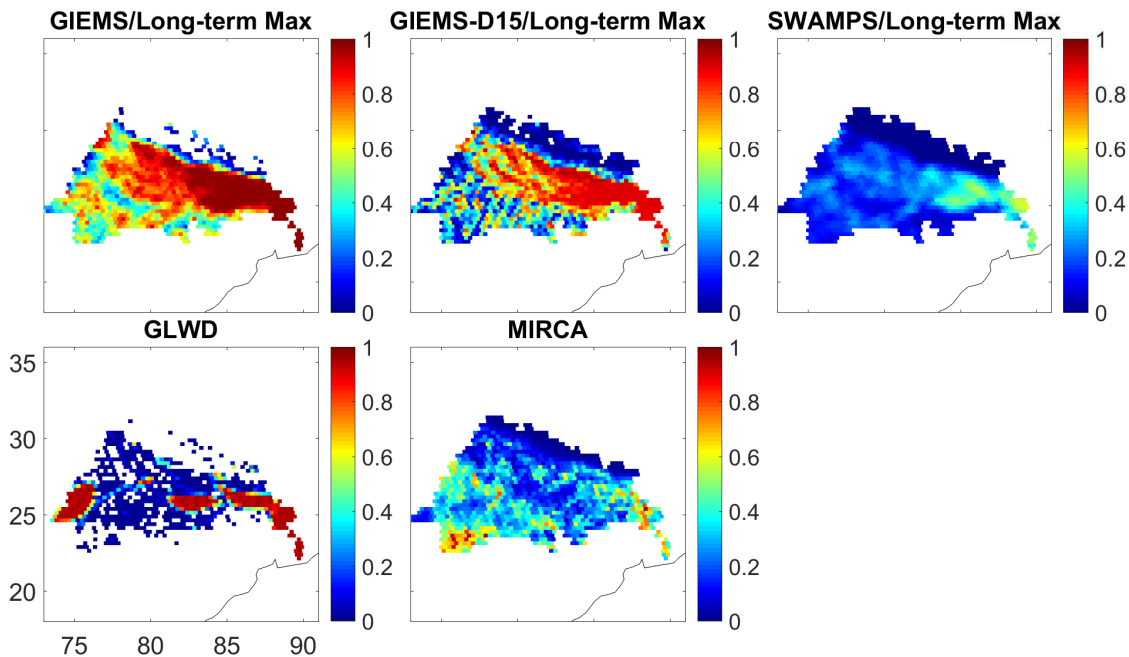


FIGURE 5.16: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS, and maximum inundation maps from the different surface water datasets over the Ganges basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \[1987\]](#) datasets.

5.5). This low correlation might also be explained by the human influence with several dams along the reaches.

TABLE 5.5: Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge Q over the Ganges basin for the period 1993-2007. Numbers in parentheses are calculated with 1-month lag between GIEMS/SWAMPS and Q.

	Time series correlation	Anomaly correlation
GIEMS/SWAMPS	87%	74%
GIEMS/Q	83% (80%)	23% (5%)
SWAMPS/Q	64% (88%)	35% (19%)

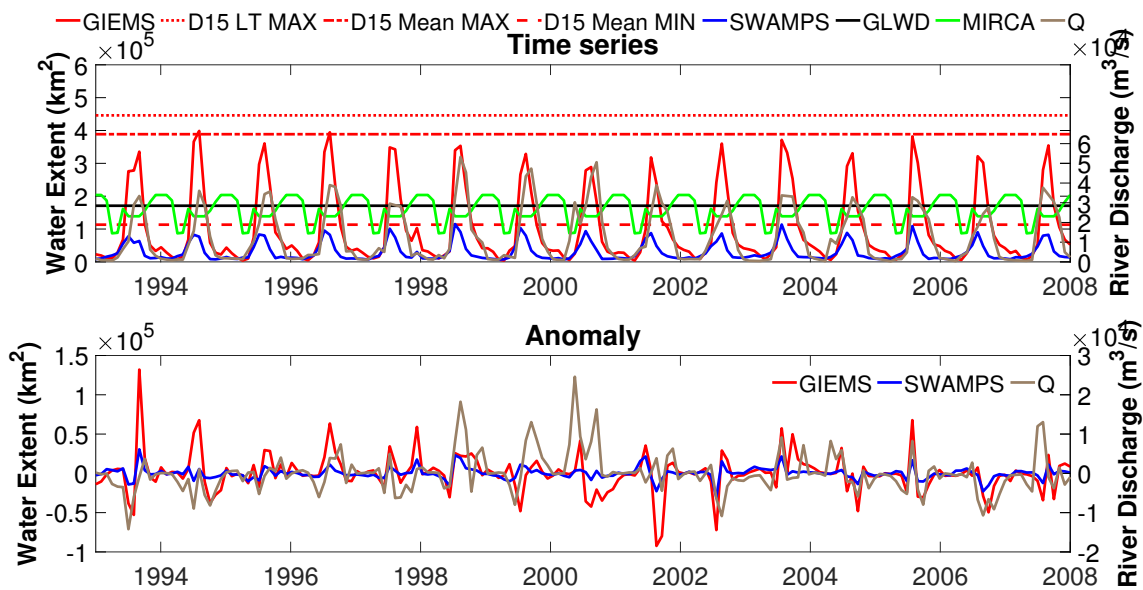


FIGURE 5.17: Time series (top) and anomaly (bottom) of surface water extent derived from the different wetland datasets, along with the river discharge over the Ganges basin from 1993 to 2007.

### 5.3.2.6 Comparisons over the Ob basin

The Ob River basin in western Siberia is selected to represent the boreal environments. Over the Ob River basin, SWAMPS surface waters are again much less extended than the other estimates (see Figure 5.18). SWAMPS surface water peaks generally in May, one month earlier than the GIEMS maximum peaks occurring generally in June (see Figure 5.19). Time series correlation between GIEMS (resp. SWAMPS) and the river discharge data for the 1993-2004 period is 91% (resp. 62%). When calculated with 1-month lag, the time series correlation decreases to



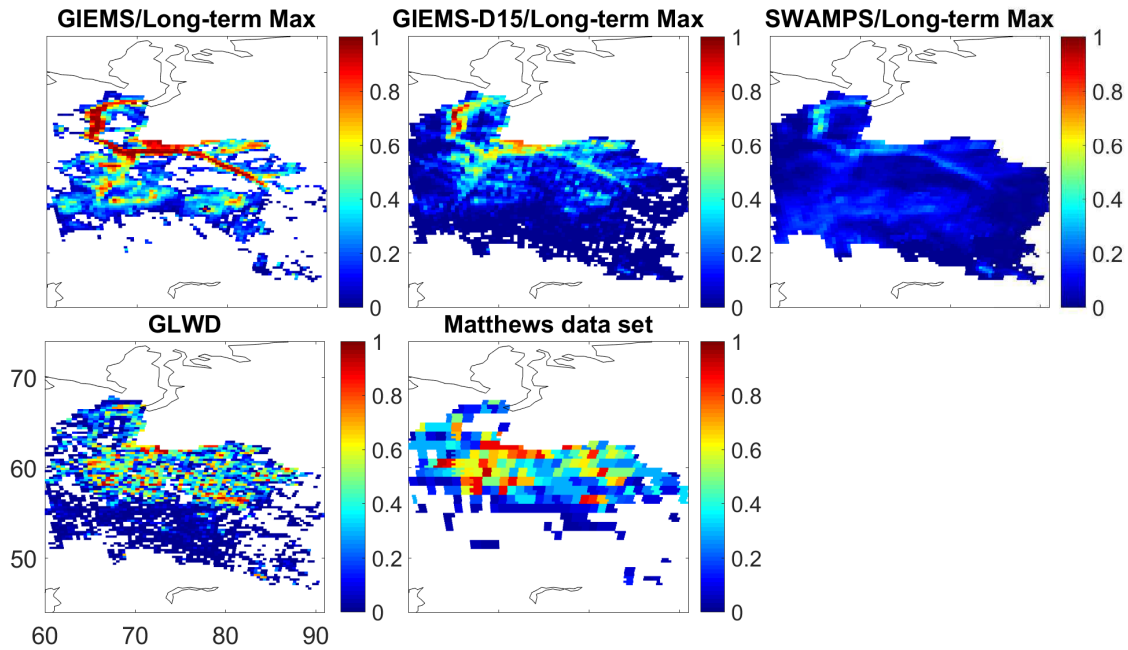


FIGURE 5.18: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS, and maximum inundation maps from the different surface water datasets over the Ob basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \[1987\]](#) datasets.

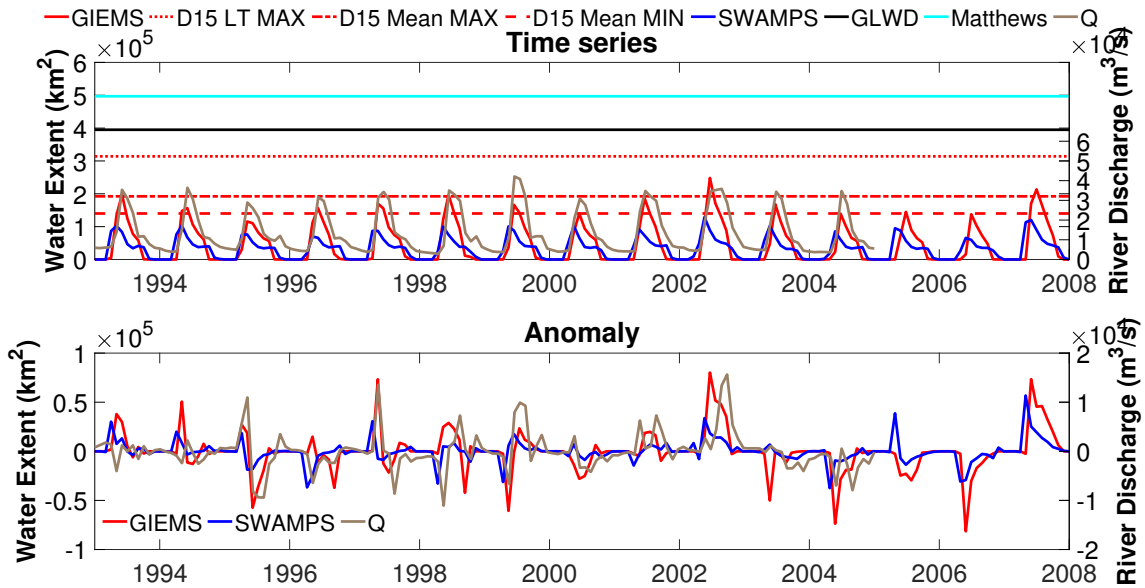


FIGURE 5.19: Time series (top) and anomaly (bottom) of surface water extent derived from the different wetland datasets, along with the river discharge over the Ob basin (available until 2004).

80% with GIEMS, but it increases to 91% with SWAMPS (see Table 5.6). The same conclusions can be found for the anomaly correlations between GIEMS (resp.

SWAMPS) and the river discharge data over the Ob River basin (also Table 5.6). The lag between GIEMS and SWAMPS could be partly related to differences in the snow filtering performed monthly with GIEMS and daily with SWAMPS. GIEMS estimates are flagged earlier in the season by the NSIDC snow mask, so the ends of the high-water-stage season (September-October) are missing. The river discharges from the Ob River basin have been compared in Papa et al. [2008b], and the use of the snow mask was well discussed. The snow flag in GIEMS is under analysis and will be refined for the next version of the dataset.

### 5.3.2.7 Comparisons over the Mississippi basin

The Mississippi River is the chief river of the largest drainage system in the North American continent. The long-term maximum inundation maps, and the time series and anomalies are shown in Figure 5.20 and Figure 5.21, respectively. Time series and anomaly correlations between GIEMS (resp. SWAMPS) and river discharge data are also shown in Table 5.6. Similar to most of the studied regions, SWAMPS fractional water extent over the Mississippi River basin is lower than that derived from the GIEMS dataset, but the two products still have high time series and anomaly correlations (84% for time series, and 74% for anomaly).

TABLE 5.6: Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge Q over the Ob basin (left) and over the Mississippi basin (right), for the period 1993-2004. Numbers in parentheses are calculated with 1-month lag between GIEMS/SWAMPS and Q.

	Time series correlation		Anomaly correlation	
	Ob	Mississippi	Ob	Mississippi
GIEMS/SWAMPS	70%	84%	38%	74%
GIEMS/Q	91% (80%)	75%	49% (40%)	59%
SWAMPS/Q	62% (91%)	68%	28% (59%)	39%

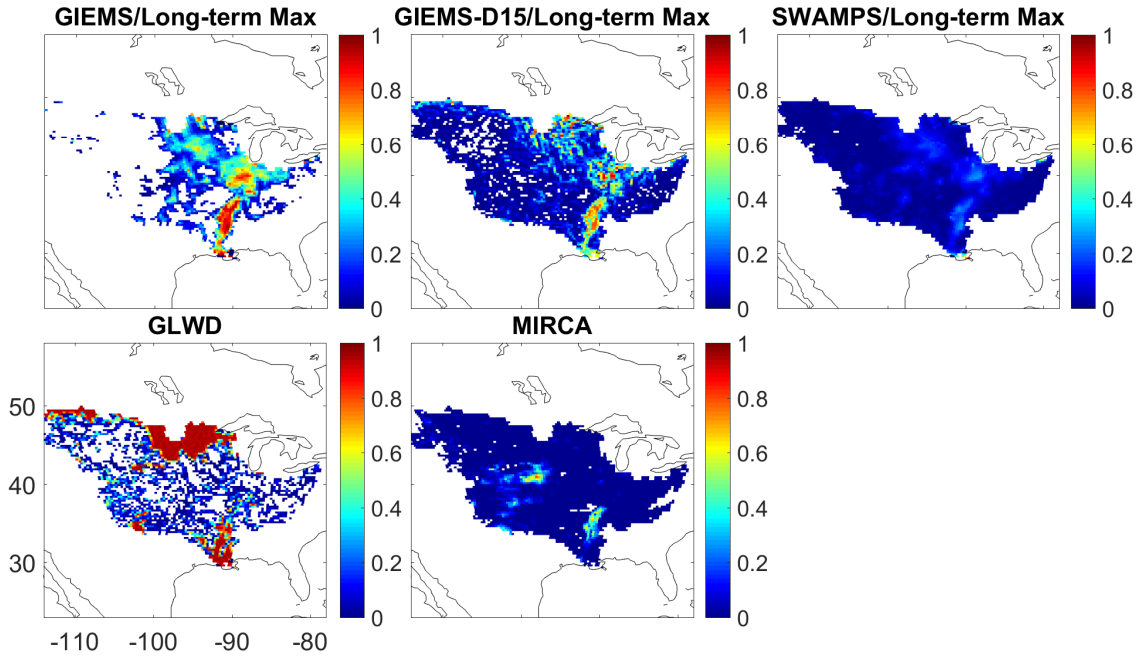


FIGURE 5.20: Long-term maximum inundation maps averaged over 1993-2007 from GIEMS and SWAMPS, and maximum inundation maps from the different surface water datasets over the Mississippi basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \[1987\]](#) datasets.

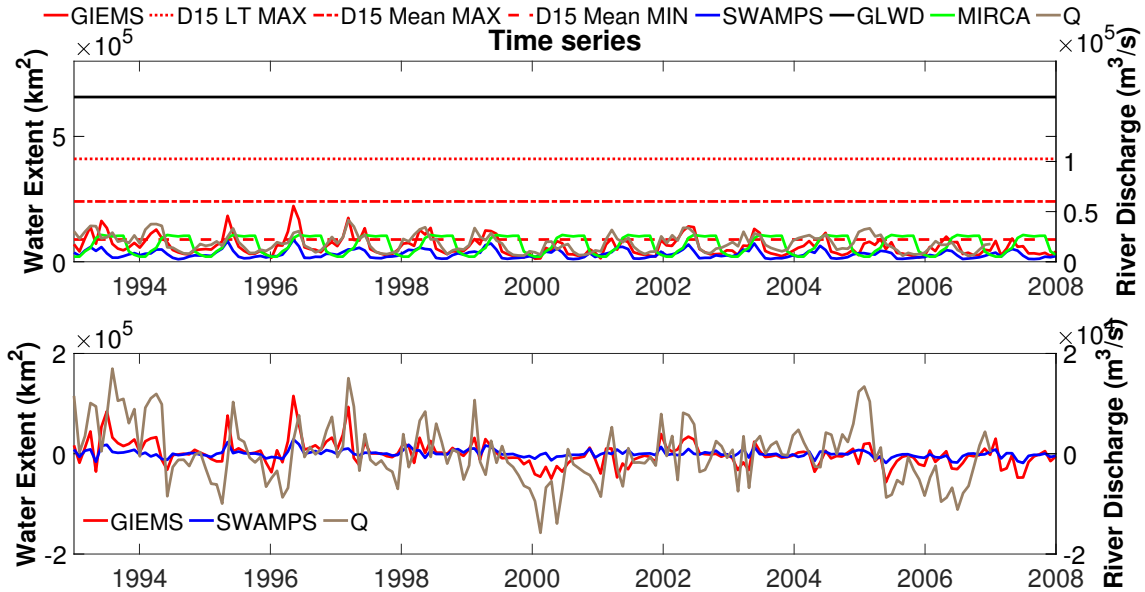


FIGURE 5.21: Time series (top) and anomaly (bottom) of surface water extent derived from the different wetland datasets, along with the river discharge over the Mississippi basin (available until 2006).

## 5.4 Discussions

As already observed in Figure 5.1, the maximum surface water extent estimated by SWAMPS for the major basins is limited, compared to the other estimates. The annual maximum SWAMPS surface extent (including the coasts) and the GLWD surfaces are similar, and this feature is advertised in Schroeder et al. [2015]. However,  $\sim 50\%$  of the surface water with SWAMPS are located along the coasts. This clearly relates to contamination by the ocean in the observation field-of-views of the passive microwave observations and to the lack of adequate filtering in the retrieval algorithm. The SWAMPS algorithm fits the global maximum water extent of GLWD, but as the coastal waters were mistakenly included in the tuning, the inland water extent is therefore strongly underestimated. The range of seasonal variability of SWAMPS is also strongly reduced because coastal regions do not evolve in time.

The underestimation of the SWAMPS extent under dense vegetation is particularly significant, as observed in the Amazon and the Orinoco basins. In the Amazon basin, the well-established SAR estimates from Hess et al. [2003] is 243,000 km<sup>2</sup> at the high stage, very close to the GIEMS values, and very different from the PALSAR estimates (40,000 km<sup>2</sup>) provided by Schroeder et al. [2015] in their paper.

For the passive microwaves, the surface emissivity of water and desert surfaces are both rather low, with large polarization differences. As a consequence, there can be confusion between deserts and surface waters. This is typically what happens over deserts with SWAMPS, with anomalous detection of surface water over arid regions. In GIEMS, the systematic use of visible and near-infrared observations helps suppress these ambiguities.

Figure 5.22 (top) shows the time series correlations between the two datasets and Figure 5.22 (bottom) shows time correlations between their anomalies, for the major 23 river basins in the world. The correlation is important for most basins, for the time series as well as for their anomalies. It is even very high for some tropical basins (Orinoco, Mekong). This tends to confirm the seasonal variations

of the surface water estimates, despite their different seasonal amplitudes. Note that correlations on time series or anomalies can be high despite large bias errors.

Over Asia (i.e., North East India, Bangladesh, Vietnam), GIEMS estimates large surface water extents that are related partly to rice paddies (see the MIRCA estimation of the rice paddies extent over the Tropical region in Figure 5.4). It is also suspected that GIEMS is very sensitive to saturated soil, and as a consequence might overestimate the surface water extent in these regions.

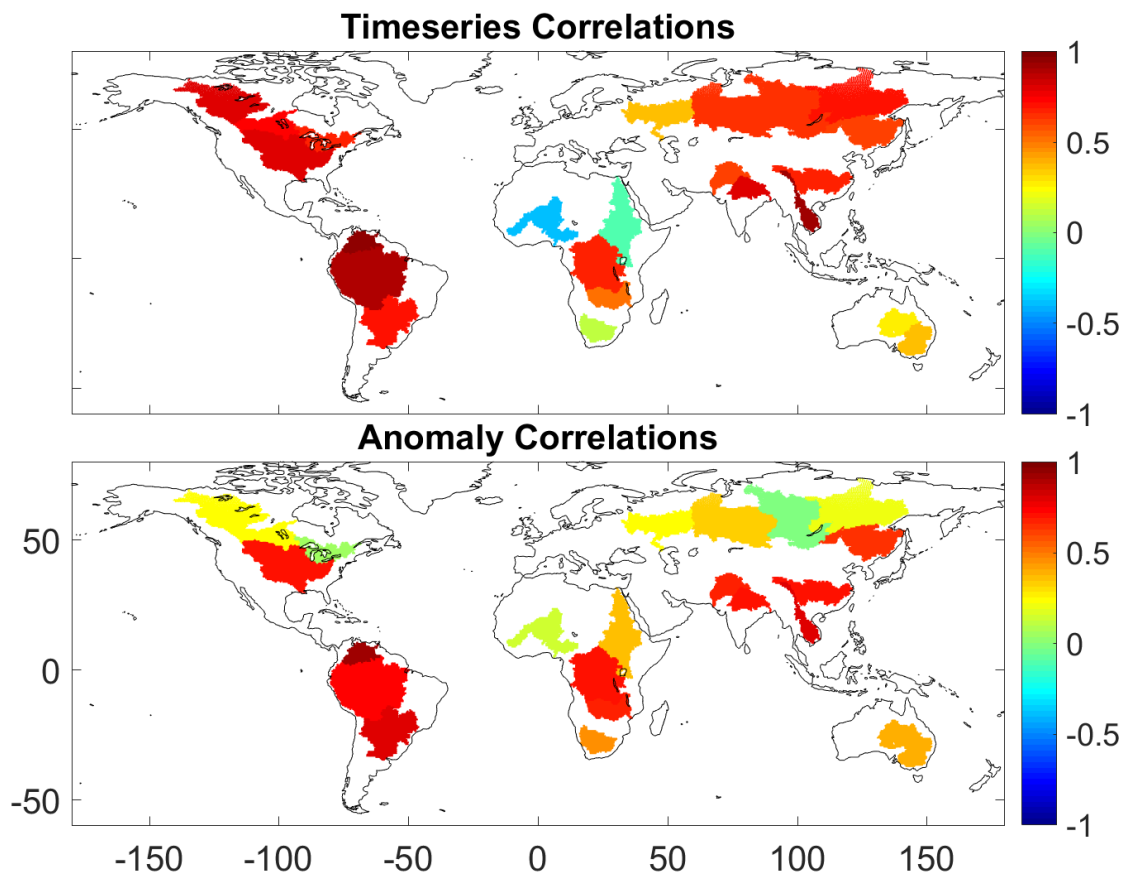


FIGURE 5.22: Basin-wide global (top) time series and (bottom) anomaly correlations for the 23 largest basins in the world between GIEMS and SWAMPS datasets.

## 5.5 Conclusions

Two global satellite-derived surface water datasets are compared on a monthly mean basis from 1993 to 2007 (GIEMS and SWAMPS), along with two widely used static maps of the surface water (GLWD and [Matthews and Fung \[1987\]](#)). The 23 largest basins in the world have been studied and 6 basins representative of different environments (the Amazon and the Orinoco basins in the Tropics, the Mississippi in North America, the Niger basin in semi arid environment, the Ganges basin in South Asia, and the Ob River basin in the boreal region) have been presented. Although they are based on similar observations, mainly passive and active microwaves, the satellite-derived datasets show large differences, globally and regionally, in terms of surface extents both at minimum and maximum inundation in the year. The global maximum inundation extent over the 1993-2007 period is  $\sim 6.2 \times 10^6$  km<sup>2</sup> for GIEMS,  $\sim 10.3 \times 10^6$  km<sup>2</sup> for SWAMPS, to compare to  $\sim 13 \times 10^6$  km<sup>2</sup> for GLWD and to  $\sim 5.3 \times 10^6$  km<sup>2</sup> for the [Matthews and Fung \[1987\]](#) wetland dataset.  $\sim 50\%$  of the SWAMPS inundated surfaces are located along the coast at the maximum annual inundation. This is clearly related to contamination by the ocean in the observation field-of-views. Once this problem is filtered out, the long-term maximum surface water from SWAMPS is reduced to  $\sim 5 \times 10^6$  km<sup>2</sup>. Globally and for the studied basins, the annual amplitude of the inundation extend is very limited in SWAMPS compared to GIEMS (47% lower). Despite their large difference in the seasonal amplitude, GIEMS and SWAMPS have similar temporal dynamics for most parts of the globe.

Over the Amazon River basin, GIEMS and SWAMPS show a very high temporal correlation for water surface time series (95%), but with SWAMPS maximum water extent a half of what is observed from GIEMS and from previous SAR estimates. Similar conclusions are found when comparing GIEMS and SWAMPS datasets over the Orinoco River basin, showing extremely high correlations for both their time series (99%) and anomaly values (97%). Over the Niger River basin, SWAMPS seasonal cycle is out of phase with both GIEMS and MODIS-derived estimates, as well as with river discharge data. This confirms the fact that SWAMPS fails to capture the seasonal dynamic of wetlands here. GIEMS and MODIS surface water time series agree in the seasonal variability, but GIEMS

water extent is significantly larger than the MODIS one. Over the Ganges River basin, GIEMS and SWAMPS datasets show similar seasonal wetland dynamics, but are very different in term of fractional water. Comparing to GLWD data over the Ganges River basin, it might suggest that GIEMS is overestimated while SWAMP is underestimated. In the Ob region, the different snow detection method could explain part of the difference in the seasonal cycle.

A clear advantage of the SWAMPS dataset today is its longer time period, up to 2013. The current version of the GIEMS algorithm requires a large quantity of ancillary satellite products to run, including outputs from the GEWEX International Satellite Cloud Climatology Project [Rossow and Schiffer, 1999] that stopped in 2008. Efforts are underway to extend the GIEMS time series to current times, to provide the community with a long time record of carefully evaluated surface wate extent all over the globe, using a reduced number of ancillary parameters for more robustness.

Part of the work in this chapter was selected to write a paper named "**Comparisons of Global Terrestrial Surface Water Datasets over 15 Years**", published in *Journal of Hydrometeorology* in April 2017.

## **Chapter 6**

# **Conclusions and Perspectives**





## 6.1 Conclusions

Monitoring surface water and their dynamics, as well as understanding the link between variations of surface water and other hydrological components (precipitation, discharge, or groundwater) are important for water management and climate change studies [Alsdorf et al., 2007]. This thesis has been developed to investigate potential of using different types of satellite observations to study the variations of hydrology from space over the last two decades, focusing on the lower Mekong basin located in Vietnam and Cambodia as this region is vulnerable to climate change impacts, and to extreme weather events like floods and droughts. This is also a preparation to the exploitation of Vietnamese satellite data (VNREDSat and JV-LOTUSat), as well as other satellite observations that are now free available to the community (such as Sentinel data). With the results from this thesis, I provide to the community surface water dynamics of the lower Mekong basin, and it can be used to predict the variation of local hydrology in the future. This thesis focuses on the following aspects:

- potential of using visible/infrared MODIS/Terra satellite observations to monitor surface water extent within the lower Mekong basin;
- potential of using SAR Sentinel-1 satellite observation to monitor surface water extent within the lower Mekong basin;
- potential of combining satellite-based products (surface water extent and height) to analyze the variation of the surface water volume within the lower Mekong basin;
- analyzing global surface water variations from 1993 to 2007, based on two global surface water datasets with monthly time series (GIEMS and SWAMPS) at global scale and basin scale.

### 6.1.1 Surface water monitoring with visible/infrared MODIS satellite observations

The first part of this thesis (Chaper 2) focuses on the use of visible/infrared MODIS satellite observations to detect and monitor surface water over the Mekong

Delta and Cambodia. As there are many well-known remote sensing techniques using different water indices (NDVI, NDWI, MNDWI, EVI, or AWEI) for this purpose, I do not try to develop a new methodology. A modified version of the Sakamoto et al. [2007] methodology is applied to produce regular surface water estimates at 500 m spatial resolution. This surface water product has a 8-day temporal resolution, starting from January 2001 until the present time. Each completed classified map contains three states: 0 for non-water pixels, 1 for mixed pixels, and 2 for water pixels. The Sakamoto et al. [2007] methodology is chosen because it has been specifically designed to detect flood and to monitor surface water over tropical regions like the Lower Mekong River Delta. Compared to the original reference Sakamoto et al. [2007] data for the common period (2001-2007), my reproduced surface water product is very similar. They share similar surface water seasonality, similar surface water amplitudes, and the total water-related areas detected are very close. This methodology is specially designed to work in tropical regions, it cannot be applied globally or in different environments (for example, over the boreal regions where surface water bodies and snow/ice surfaces have similar signatures for the EVI index that is the main indicator in this methodology). Regular MODIS-derived surface water maps (since January 2001) for the Mekong Delta and Cambodia are updated monthly, then provided free of charge to the community at the following address: <http://space.usth.edu.vn/en/news/projects/>.

### **6.1.2 Surface water monitoring with SAR Sentinel-1 satellite observations**

The Mekong Delta is often covered by clouds, especially during the rainy season causing difficulties to visible/infrared satellite sensors (like MODIS ones). SAR satellite sensors can overcome this problem with their ability to work day and night, under all weather conditions. In the second part of the thesis (Chapter 3), SAR Sentinel-1 satellite observations are used for the same purposes as in the first part to detect and monitor surface water over the Mekong Delta and Cambodia. A Neural Network (NN) classification trained with optical Landsat-8 satellite observations is designed to classify each SAR Sentinel-1 pixel into non-water or

water pixel. SAR Sentinel-1 images are converted to 30-m spatial resolution to match with Landsat-8 spatial resolution. Sensitivity tests are carried out to define the number of the NN input parameters, then to optimize the performance of the NN classification. Comparisons between the SAR predicted surface water maps and the reference Landsat-8 surface water maps over the Mekong basin show a true positive water detection of  $\sim 90\%$ . SAR predicted surface water maps are also compared to the floodability maps derived from high spatial resolution topography data. Results show a high consistency between the two independent products, with 98% of SAR-derived surface water pixels located in areas with a high probability of inundation. Results from the second part highlight the applicability of the SAR Sentinel-1 observations for surface water monitoring, especially in tropical regions where cloud cover can be very high during the rainy seasons. The output SAR Sentinel-1 derived surface water maps (since February 2015) for the Mekong Delta and Cambodia are also updated monthly, and provided freely to the community at the same location as the MODIS-derived products previously described (<http://space.usth.edu.vn/en/news/projects/>).

Work presented in this chapter were selected to write a paper that has been published in *Water* in May 2017 (see Appendix A).

### 6.1.3 Toward the variation of the surface water volume

The third part of the thesis (Chapter 4) is an development of the two previous chapters. It is dedicated to analyze monthly variation of the surface water volume in the Mekong Delta and Cambodia. Surface water volume changes are calculated as the product of the MODIS-derived surface water extent and the satellite-based surface water height, based on the methodology described in Frappart et al. [2008, 2011]. The surface water extent are validated by compared to SAR-derived surface water extent, precipitation and in situ water level data, while the surface water height are validated with results from previous studies, as well as with in situ water level data. Then the two components are directly compared, and results always show high temporal correlations between them ( $> 90\%$ ). Surface water height at VSs are linearly interpolated over inundated areas to construct monthly maps of surface water height at the same spatial resolution as the MODIS-derived

surface water extent (500 m), before the monthly variation of the surface water volume is extracted. Results show high temporal correlations with both GRACE-derived total land surface water volume changes (95%) and changes of discharge estimates of the Mekong River derived from in situ measurements (96%). Variation of soil moisture derived from models could be used to quantify the contribution of the surface water volume, the groundwater and the soil moisture on the total land surface water volume.

#### **6.1.4 Comparisons between two global terrestrial surface water datasets**

The last part of the thesis expands the satellite analysis of surface water extent to the globe by comparing mainly two global multi-satellite surface water extent datasets with monthly time series (GIEMS and SWAMPS), for the common 1993-2007 period. Global comparisons between GIEMS and SWAMPS show that they have similar temporal dynamics for most parts of the globe, but ~50% of the SWAMPS inundated surfaces are located along the coast at the maximum annual inundation. This problem is related to the contamination by the ocean as the SWAMPS algorithm did not use a careful land/ocean filter to remove signals emitted by the ocean. Compared to GIEMS, SWAMPS detects surface water almost everywhere even in areas where surface water are not expected to be presented like the North African desert. In addition, SWAMPS also shows very limited amplitude in the annual inundation cycle in many regions where large seasonal variations are expected (for example over the Amazon, the Ganges, or the Mekong basins).

Comparisons between surface water time series and anomalies of GIEMS and SWAMPS over the 23 largest river basins in the world confirm the surface water seasonal variations in most locations, but there are still differences between the two products. Over the Amazon and Orinoco basins, GIEMS and SWAMPS have similar surface water spatial distributions at both low and high water states. They also share the same dynamics for their surface water time series, and even their anomalies for the 15-year period. However, the total surface water area detected from SWAMPS is 2-3 times lower than that detected from GIEMS, and SWAMPS

surface water amplitude is much more limited than GIEMS. Similar conclusions can be made over the Ganges basin, but GIEMS seems to overestimate the surface water extent while SWAMPS seems to underestimate it, compared to the maximum derived from GLWD. Over the Niger basin, SWAMPS fails to capture correctly the surface water dynamics since its seasonal surface water time series is out of phase with GIEMS, MODIS-derived estimates from [Bergé-Nguyen and Crétaux \[2015\]](#), or in situ river discharge data.

Work presented in this chapter were selected to write a paper that has been published in *Journal of Hydrometeorology* in April 2017 (see Appendix B).

## 6.1.5 Other activities during my PhD

### 6.1.5.1 Participation to another paper

During my PhD, I also participated to the work of [Aires et al. \[2017\]](#) on producing the high-spatial-resolution (90 m) version of the GIEMS dataset (GIEMS-D3). My contribution focuses on validation of the GIEMS-D3 dataset to other independent satellite observations from MODIS, Landsat, and SAR over the Amazon and the Mekong basins.

### 6.1.5.2 Teaching experience

Since the 2016/2017 academic year, I started to give practical training for the module 21.16 (Remote Sensing of Continental Surfaces) to master students at the University of Science and Technology of Hanoi (USTH). This has been a great opportunity to train myself to prepare for the future when I come back to work as a lecturer and a researcher at USTH after finishing my PhD training in LERMA. Through the practical training, students can see the great potential of using satellite products to study Earth sciences since more and more high quality remote sensing data are provided free of charge to the community.

## 6.2 Perspectives

This thesis has shown various applications of satellite observations in studying hydrology variation from space. More work could be done to benefit from all different types of satellite products in several directions. First, same techniques as in Chapters 2,3 can be applied for surface water monitoring, or flood detection in other types of environment in Vietnam (over the Red River Delta in the North Vietnam, for example) and in other countries in Southeast Asia. Second, the closure of the water cycle over the lower Mekong basin can be checked using different satellite products. Last but not least, some techniques presented in this thesis can be further developed to adapt to data from the JV-LOTUSat-1 satellite (planned to be launched in 2020), or from the US-French Surface Water and Ocean Topography (SWOT) mission (planned to be launched in April 2021).

### 6.2.1 Surface water monitoring over the Vietnam Red River Delta

The Red River Delta (located in the North Vietnam) is the second largest river delta in Vietnam, after the Mekong Delta, as well as the second most important rice-producing area in the country (accounting for ~20% of the national crop) [Tran, 2003]. Monitoring surface water, and especially floods over this delta is important because floods can cause serious damages for this highly dense population area (almost a third of the Vietnam population lives in the Red River basin, including over 17 million inhabitants in the delta itself [Tran, 2003]). Using optical satellite observations for this task is very challenging because this area is cloud-covered most of the year. During one year (November 2016 - November 2017), there is no optical Landsat-8 observations with less than 5% cloud contamination available in this area, and there are only 4 imagery with cloud contamination less than 10%. Therefore, SAR satellite observations (like Sentinel-1 or JV-LOTUSat) are the best choice for monitoring applications over the Red River Delta.

Due to heavy local precipitation in July, August and October 2017, there were several floods occurring in the North Vietnam that caused serious damages and deaths to local communities (more than 100 people dead or missing during the floods from 9-14 October 2017, according to reports from the Vietnam National

Hydro-Meteorological Service). Some tests have been done using SAR Sentinel-1A observations (data from the Sentinel-1B satellite are not available in this area) to have an overview of areas affected by the floods. Primary results of flooded areas are shown in Figures 6.1 & 6.2 for June-August, and October 2017 periods, respectively. Figure 6.1 shows the evolution of inundation over the North Vietnam from 28 May to 1 September 2017. Floods occurred mostly over the northeast

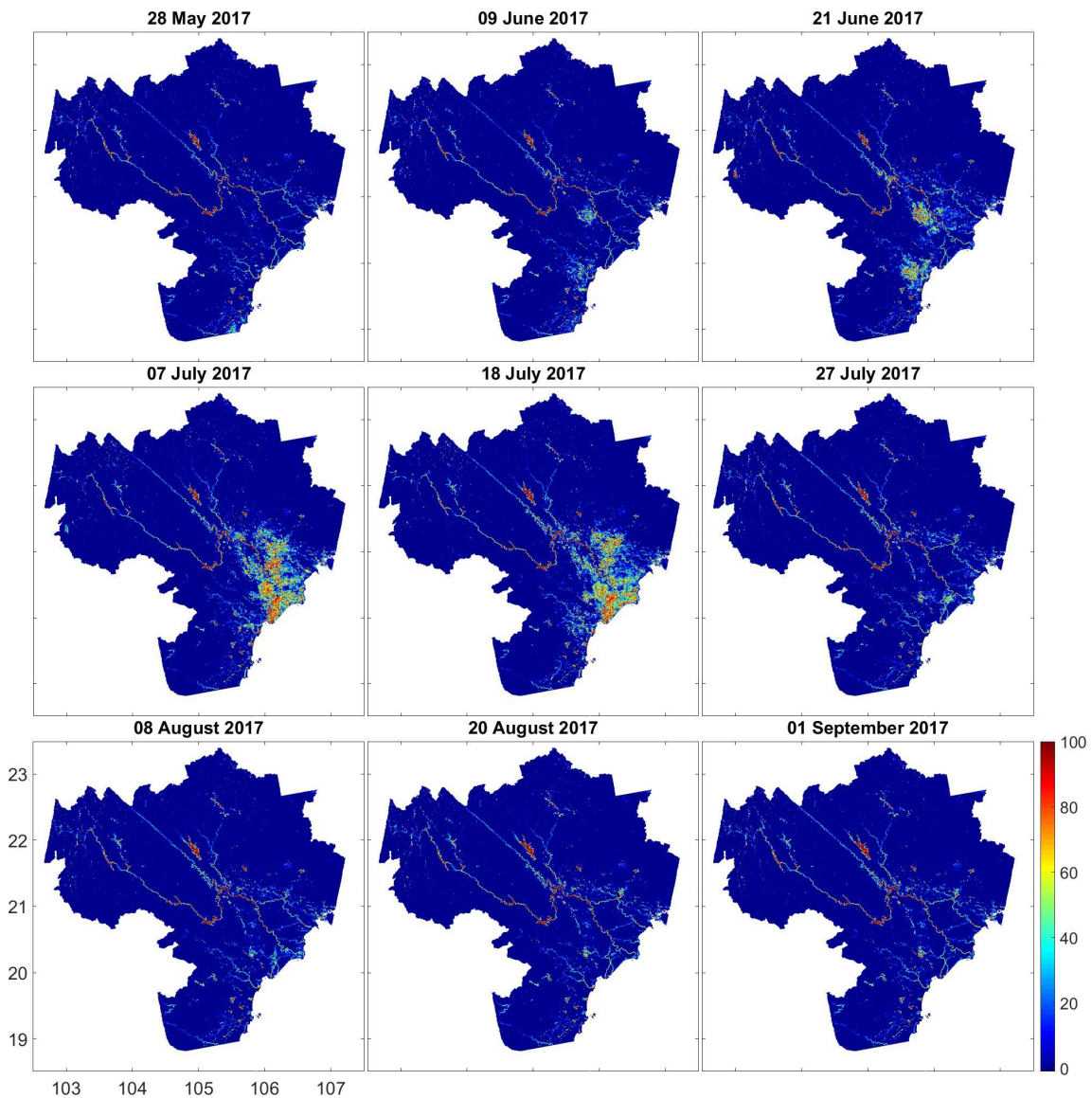


FIGURE 6.1: Flooded areas over the North Vietnam from June to August 2017, derived by applying a threshold ( $T = -21.18$  dB) directly to SAR Sentinel-1A VH polarization observations. The spatial resolution is converted to 550 m for a better visualization of the effected areas.



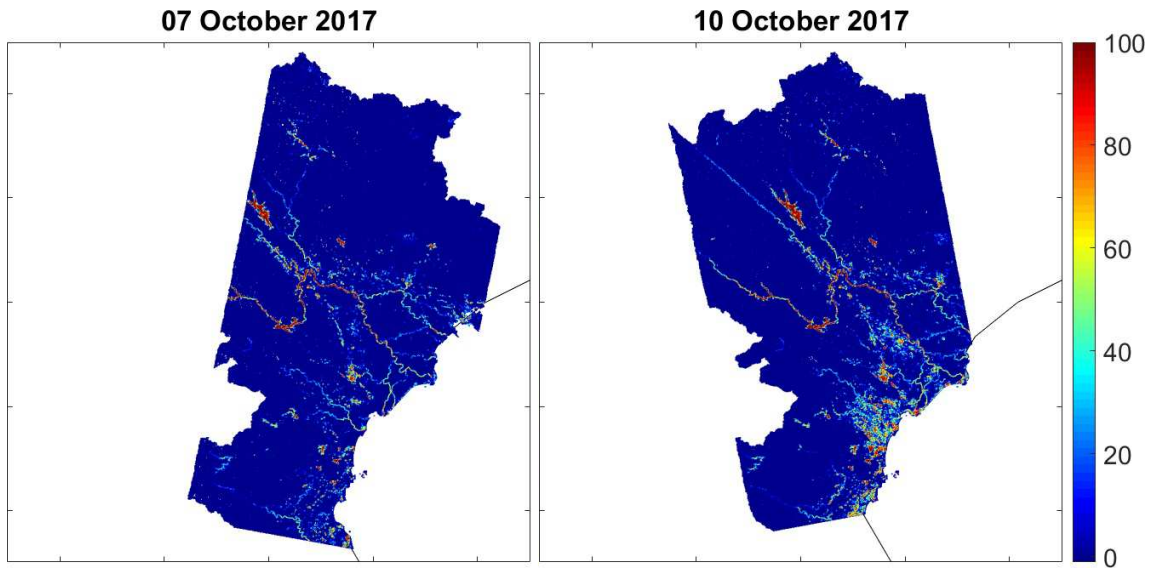


FIGURE 6.2: Same as Figure 6.1, but for the floods from 9-14 October 2017

provinces (Hai Phong, Thai Binh, Nam Dinh, and Ninh Binh). The soil started to saturate in late June, then the food peak occurred middle of July, before going back to a normal state at the beginning of August. Figure 6.2 shows the second flood event occurring in October 2017 before and after the starting day of the flood (9 October).

As these flood maps are produced based on the simple threshold method applied directly to the SAR Sentinel-1 VH polarized backscatter coefficient ( $T = -21.18$  dB), results are very sensitive to change of the threshold. As seen in Table 6.1, the inundated area will increase (decrease) from 14% to 16% when the threshold is increased (decreased) by 0.5 dB. More work need to be done to find a robust threshold working well for this area. There are many mountains located in North Vietnam, that can affect the backscattering signal, and the method needs to be refined to better distinguish surface water and mountain shadows. In situ measurements of local precipitation and water level can be collected for validation. Results from this technique can be useful for management activities. Warnings can be sent to local communities if soil saturation is observed, the next precipitation likely leading to floods.

TABLE 6.1: Threshold sensitivity when increasing (decreasing) the threshold by 0.5 dB. The total area is  $\sim 138,233 \text{ km}^2$ . Numbers in parenthesis indicate inundation percentages for area shown in Figure 6.1.

Date	Total inundated area ( $\text{km}^2$ )		
	T = -21.68 dB	T = -21.18 dB	T = -20.68 dB
2017 May 28	1276 (0.92%)	1430 (1.04%)	1621 (1.17%)
2017 June 09	1336 (0.97%)	1550 (1.12%)	1796 (1.30%)
2017 June 21	1861 (1.35%)	2303 (1.67%)	2813 (2.03%)
2017 July 07	3269 (2.36%)	4027 (2.90%)	4843 (3.50%)
2017 July 18	3403 (2.46%)	4223 (3.06%)	5094 (3.69%)
2017 July 27	1571 (1.14%)	1813 (1.31%)	2123 (1.54%)
2017 August 08	1522 (1.10%)	1722 (1.25%)	1971 (1.43%)
2017 August 20	1588 (1.15%)	1745 (1.26%)	1916 (1.38%)
2017 September 01	1630 (1.18%)	1784 (1.29%)	1948 (1.40%)

## 6.2.2 Closure of the water cycle over the Mekong Delta and Cambodia

An other research direction after this thesis is to check the closure of the water cycle over the lower Mekong basin, using satellite products. By definition, the terrestrial water budget is composed from four water budget components, including precipitation ( $P$ ), evapotranspiration ( $E$ ), runoff (surface and sub-surface flows -  $R$ ), and surface and sub-surface water storage variation ( $\Delta S$ ). The relationship between the four water budget components is shown in equation 6.1, meaning that the fluxes of precipitation, evapotranspiration and runoff are balanced by the change of the total water storage.

$$dS/dt = P - E - R \quad (6.1)$$

Some of the water budget components have been collected and studied in this thesis. Evapotranspiration will be calculated from the Global Land Surface Flux Product Dataset (LandFlux) and the Global Land-surface Evaporation (GLEAM). For other variables ( $P$ ,  $E$ , and  $R$ ), different global datasets exist, but they show significant differences in time and spatial variation, with strong difficulties to evaluate the accuracy of each individual dataset. The consistency of each single water budget component will be examined first to identify the most reliable dataset.

Then, the closure of the water cycle over the Mekong Delta and Cambodia will be checked.

### 6.2.3 The future with the SWOT and JV-LOTUSat satellites

Data from the future SWOT and JV-LOTUSat satellites can be used for hydrology and water management applications. The SWOT mission is specially designed to provide major improvement in oceanography and inland hydrology thank to its new altimeter instrument concept working in a Ka-band radar interferometer (KaRIn). The SWOT mission configuration is presented in Figure 6.3. All the past and present satellite altimeters emit microwave pulses in the nadir direction, the SWOT satellite emits microwave pulses in near nadir incidence angles ( $0.6^\circ - 3.9^\circ$ ) [Fjrtoft et al., 2014]. As a consequence, altimetry observations are not limited just along the satellite track, the SWOT satellite can provide altimetry observations on both sides of the satellite track up to 50 km wide. In addition, each point inside the swath will be observed by two different sensors onboard the satellite, that will provide very precise estimations of the elevation at each point. The SWOT will

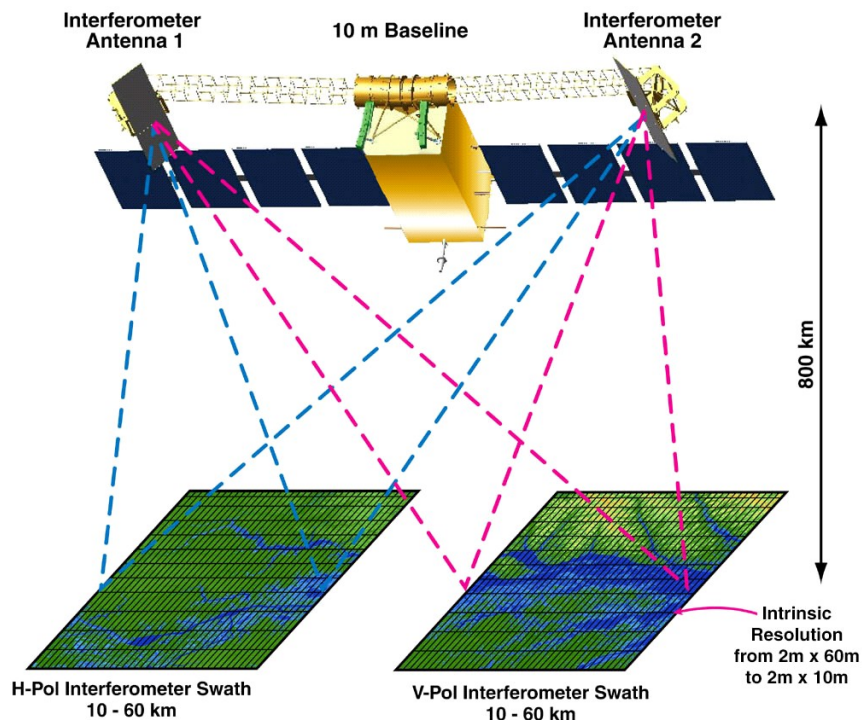


FIGURE 6.3: The SWOT mission configuration. Figure is credited by NASA.

provide the water heights of water bodies larger than  $250\text{ m} \times 250\text{ m}$ , and the water heights and discharge of rivers wider than  $100\text{ m}$ . Global long-term high-spatial-resolution surface water product can be produced by combining SWOT data with GIEMS dataset.

The configuration of the JV-LOTUSat-1,2 satellites is presented in Figure 6.4. LOTUSat-1 will be the first SAR satellite of Vietnam to operate in the X-band, and provides high resolution observations (maximum at  $1\text{ m}$  spatial resolution). The JV-LOTUSat-1,2 satellites will help Vietnam to take the initiative in acquiring satellite imagery for management of natural resources and environment.

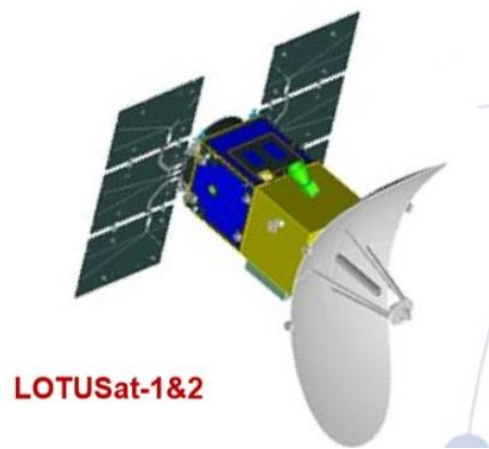


FIGURE 6.4: The JV-LOTUSat-1,2 configuration. Figure is credited by the Vietnam National Space Center.

Satellite data, especially from SAR satellites, are very important for countries located in tropical regions (like Vietnam). Satellite SAR data are the most effective source of imagery from space for monitoring applications over for these highly cloudy regions. All exercises in this thesis can be considered as training for the use of the different types of satellite data in the future (Sentinel-2,3, SWOT, JV-LOTUSat-1,2, or Jason-3, for example).



# **Appendix A**

## **Paper A: Surface Water Monitoring within Cambodia and the Vietnamese Mekong Delta over a Year, with Sentinel-1 SAR Observations**

**Authors:**

B. Pham-Duc, C. Prigent, F. Aires

**Paper published in:**

Water, 2017



Article

# Surface Water Monitoring within Cambodia and the Vietnamese Mekong Delta over a Year, with Sentinel-1 SAR Observations

Binh Pham-Duc <sup>1,2,\*</sup>, Catherine Prigent <sup>1</sup> and Filipe Aires <sup>1</sup>

<sup>1</sup> Laboratoire d'Études du Rayonnement et de la Matière en Astrophysique et Atmosphères, UMP 8112, l'Observatoire de Paris, 61 Avenue de l'Observatoire, 75014 Paris, France; catherine.prigent@obspm.fr (C.P.); filipe.aires@obspm.fr (F.A.)

<sup>2</sup> Space and Aeronautics Department, University of Science and Technology of Hanoi, 18 Hoang Quoc Viet, Cau Giay, 10000 Hanoi, Vietnam

\* Correspondence: pham.binh@obspm.fr; Tel.: +33-07-8243-5237

Academic Editor: Frédéric Frappart

Received: 24 March 2017; Accepted: 19 May 2017; Published: 23 May 2017

**Abstract:** This study presents a methodology to detect and monitor surface water with Sentinel-1 Synthetic Aperture Radar (SAR) data within Cambodia and the Vietnamese Mekong Delta. It is based on a neural network classification trained on Landsat-8 optical data. Sensitivity tests are carried out to optimize the performance of the classification and assess the retrieval accuracy. Predicted SAR surface water maps are compared to reference Landsat-8 surface water maps, showing a true positive water detection of ~90% at 30 m spatial resolution. Predicted SAR surface water maps are also compared to floodability maps derived from high spatial resolution topography data. Results show high consistency between the two independent maps with 98% of SAR-derived surface water located in areas with a high probability of inundation. Finally, all available Sentinel-1 SAR observations over the Mekong Delta in 2015 are processed and the derived surface water maps are compared to corresponding MODIS/Terra-derived surface water maps at 500 m spatial resolution. Temporal correlation between these two products is very high (99%) with very close water surface extents during the dry season when cloud contamination is low. This study highlights the applicability of the Sentinel-1 SAR data for surface water monitoring, especially in a tropical region where cloud cover can be very high during the rainy seasons.

**Keywords:** SAR; Sentinel-1; surface water monitoring; neural network; Mekong Delta; Landsat-8; MODIS

## 1. Introduction

Studying the spatial and temporal distribution of surface water resources is critical, especially in highly populated areas and in regions under climate change pressure. With an increased number of Earth-observation satellites providing a large diversity of remote sensing data, there is now the potential to monitor the surface water at regional to global scale. However, mapping surface water is still challenging. It is difficult to provide products with the accuracy required for a large range of applications (e.g., agriculture, disaster management, and hydrology).

Several methods have already been proposed to detect and monitor surface water with visible and Near-Infrared (NIR) images. Ref. [1] used positive values of the Normalized Difference Water Index (NDWI) to classify water bodies. Ref. [2] applied a threshold on NIR reflectances of the NOAA/AVHRR satellite to delineate lakes. Ref. [3] detected surface water by identifying the positive values of the Modification of Normalized Difference Water Index (MNDWI). Ref. [4] combined NIR data and the Normalized Difference Vegetation Index (NDVI) to detect surface water bodies. However, cloud



contamination is a stringent constraint for these methods, limiting their application to cloud-free conditions which is very restrictive in some regions (e.g., in the Tropics). Vegetation can also mask the surface water partly or totally. This makes the water detection difficult or impossible under canopy. In addition, the NIR reflectance over highly turbid water can be higher than the red reflectance, introducing confusions in the indices used for the water detection.

Synthetic Aperture Radar (SAR) have become an important source of data to detect flood or monitor surface water as they allow observations regardless of the cloud cover, day and night, with spatial resolution comparable to visible and near-infrared satellite images [5]. SAR instruments have been available on many sensors and platforms (Envisat ASAR, PALSAR, or RADARSAT, for example) providing observations for different areas all over the globe (but normally with a limited number of images available per year in some regions). Flood detection using different SAR observations has been studied by many authors, showcasing the advantages of SAR instruments compared to optical instruments in monitoring floods. Ref. [6] used a single decision tree classifier on two sets of JERS-1 SAR data to classify surface water within the states of North Carolina and South Carolina into five land cover types (water, marsh, flooded forest, field, and non-flooded forest). Although the classifier was simple, they reported an overall classification accuracy of nearly 90%. Ref. [7] showed the potential of the COSMO-SkyMed data for flood detection by showing case studies in several locations all over the globe (e.g., Tarano River overflow, Italy, April 2009; Pakistan inundation, July–September 2010; Thailand flood, October 2010; and Australia flood, January 2011). COSMO-SkyMed instruments provided very high resolution X-band SAR images, but covered limited areas (the highest spatial resolution is  $\sim 1$  m for an observation area of  $10 \text{ km} \times 10 \text{ km}$ ). X-band data from TerraSAR-X instrument were also reported suitable for flood mapping under forest canopy in the temperate forest zone in Estonia [8]. Ref. [9] compared four flood detection approaches over five areas (Vietnam, the Netherlands, Mali, Germany, and China) using SAR data from the TanDEM-X mission. Although these four approaches were designed according to different requirements, their performances were satisfactory over the studied areas (17 out of 20 water masks reaching an overall accuracy larger than 90%). Other studies using SAR data for water monitoring locally and regionally under different environments can be listed, such as [10–12]. Mapping water bodies at global scale using SAR data was limited due to the lack of global observations, and the fact that SAR data are not easy to access freely. Ref. [13] used multi-year (2005–2012) Envisat ASAR observations to create, for the first time, a global potential water body map at a spatial resolution of 150 m. Errors concentrated along shorelines and coastline, but this global water map has an accuracy of  $\sim 80\%$  compared to the reference data.

The Mekong Delta in Southeast Asia (one of the largest deltas in the world) is a vast triangular plain of approximately  $55,000 \text{ km}^2$ , most of it lower than 5 m above sea level. The seasonal variation in water level results in rich and extensive wetlands. For instance, the Mekong Delta region covers only 12% of Vietnam but produces  $\sim 50\%$  of the annual rice (with two or three harvests per year depending on the provinces), represents  $\sim 50\%$  of the fisheries, and  $\sim 70\%$  of the fruit production. In the Delta, the dry season extends from November to April and the rainy season from May to October. Many researches have been carried out to monitor the surface water in the Delta, using both optical and active microwave satellite images. Ref. [14] produced a monthly mean climatology of the water extent from 2000 to 2004 with a spatial resolution of 500 m, using visible and NIR MODIS/Terra data. However, with 85% to 95% cloud cover during the wet season over the Mekong Delta [15], remote sensing methods derived from visible and NIR images present some limitations. Different SAR observations have also been exploited to study floods and wetlands over the Delta. Ref. [16] mapped flood occurrence for the year 1996 over the Delta using five ERS-2 observations. Ref. [17] used 60 Envisat ASAR observations during the years 2007–2011 to study the flood regime in the Delta. Thanks to the launch of the Sentinel-1A & B satellites, as well as the free data policy of the European Space Agency (ESA), Sentinel-1 SAR observations are now regularly and freely accessible for scientific and educational purposes, over large parts of the globe. Similar to previous SAR instruments, Sentinel-1 instruments show strong potential for detecting open water bodies at high spatial resolution [18,19].

With the advantage of higher temporal resolution than previous SAR instruments, Sentinel-1 has the ability to monitor the seasonal cycle of water extent every six days over Europe and the boreal region, and with slightly reduced temporal sampling elsewhere. In this study, we propose a methodology using Sentinel-1A SAR observation for monitoring water surface extent within Cambodia and the Mekong Delta for the year 2015. It is based on a Neural Network (NN) algorithm, trained on visible Landsat-8 images (30 m spatial resolution). At the time of this study, the temporal resolution of Sentinel-1 over the Delta was 12 days: it reduced to 6 days after the launch of the Sentinel-1B in April 2016.

The Sentinel-1 SAR data and the ancillary observations are described in Section 2, including the pre-processing steps. Section 3 presents the NN methodology, along with sensitivity tests. Results and comparisons with other products are provided and discussed in Section 4. Section 5 concludes this study.

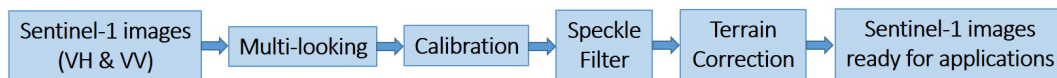
## 2. Sentinel-1 SAR Data and the Ancillary Datasets

### 2.1. Sentinel-1 SAR Data

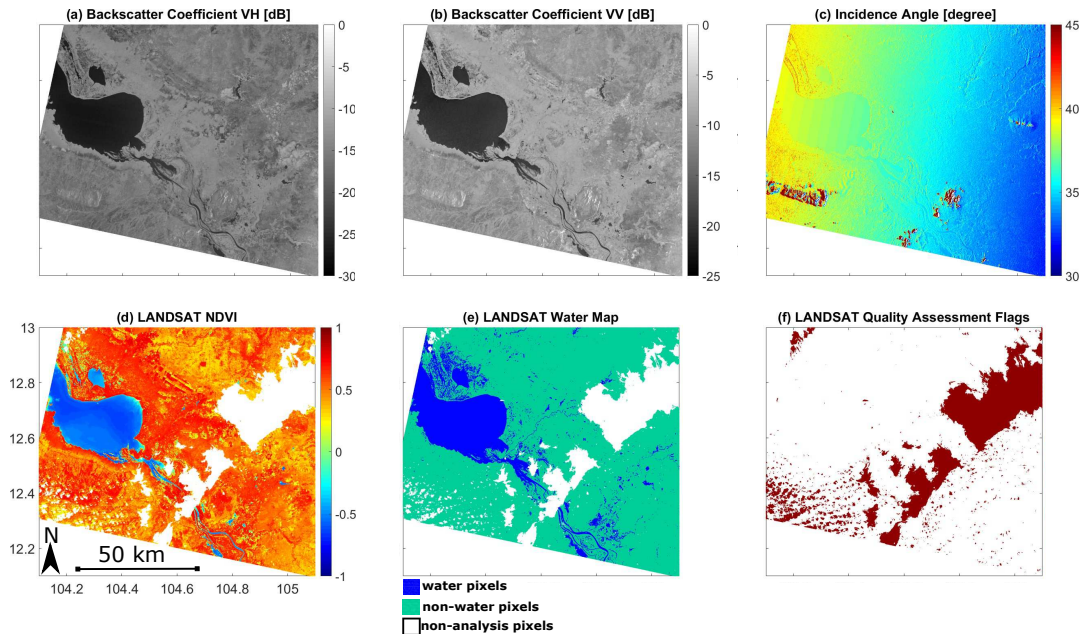
Sentinel-1 is a satellite project funded by the European Union and carried out by the European Space Agency. It is a two satellite constellation working at C-band (5.405 GHz). The major objective of the satellites is the observations and monitoring of land and ocean surfaces day and night, under all weather conditions [20]. The satellite operates in four exclusive imaging modes with different spatial resolutions (the highest being 5 m) and swaths (up to 400 km). The first Sentinel-1A satellite of the pair was launched on 3 April 2014, while the second Sentinel-1B satellite was launched on 22 April 2016. The Sentinel-1 satellites fly along a sun-synchronous, near-polar circular orbit at an altitude of  $\sim 693$  km. Incidence angle varies between  $29^\circ$  and  $46^\circ$ . The two satellites provide a re-visiting time of 6 days (it was 12 days before the launch of the Sentinel-1B satellite). Sentinel-1 satellites have dual polarization capabilities (HH, VV, HH + HV and VV + VH), giving final users the ability to access a large variety of applications, including the monitoring of surface water. SAR images from Sentinel-1 satellites are freely downloaded from the sentinel scientific data hub [21].

In this project, 20 m resolution (10 m pixel spacing) Level-1 Ground Range Detected (GRD) Sentinel-1 images are used, from the Interferometric WideSwath (IW) mode. These images have been detected and projected to ground range using an Earth ellipsoid model provided by ESA. Over the Mekong Delta, there are two polarizations available: the VH and VV polarizations. Some pre-processing steps have to be carried out using the free Sentinel Application Platform (SNAP) software developed by ESA, before moving to the analysis steps (see Figure 1). These pre-processing steps are described in the “SAR Basics with the Sentinel-1 Toolbox in SNAP tutorial” [22].

First, multi-looking processing is applied to each single Sentinel-1 image (both polarizations) to convert to 30 m spatial resolution (to match with Landsat-8 images). Applying multi-looking at the beginning of the chain reduces the processing time for the next steps since the size of the image is several times smaller than the original one. Second, the image is calibrated to convert values of the raw image from digital number to radar backscatter coefficient ( $\sigma^0$ ). Third, the Refined Lee filter is applied to reduce the speckle noise and to smooth the radar backscatter coefficient data because this filter maintains details of the standing water boundary [23]. Other filters (Lee, Lee Sigma or Median, for example) were tested, and results showed little differences in terms of water detection. Next, the “terrain correction” tool is used to compensate for distortions in the SAR images, so that the geometric presentation of the image will be as close as possible to the real world. At the end of this step, the image is also re-projected from the satellite projection to the Earth geographic projection, and is ready for applications. To fully cover Cambodia and the Vietnamese Mekong Delta, at least five Sentinel-1 SAR images are needed. Figure 2 (top) provides examples of the SAR backscatter coefficients for VH (a) and VV (b) polarizations, along with the incidence angle (c), over the Tonle Sap Lake, on 17 December 2015.



**Figure 1.** Pre-processing steps for Sentinel-1 Synthetic Aperture Radar (SAR) images.



**Figure 2.** Examples of satellite observations from Sentinel-1 (top) and from Landsat-8 (bottom), over the lower part of the Tonle Sap Lake (Cambodia) after the pre-processing steps: (a) SAR backscatter coefficient at VH polarization; (b) SAR backscatter coefficient at VV polarization; (c) SAR incidence angle; (d) The Normalized Difference Vegetation Index (NDVI) from Landsat-8; (e) Surface water estimated from Landsat-8; and (f) Landsat-8 quality flags. The white areas are cloud-covered pixels detected by the Landsat quality flags, and have been removed. Both Sentinel-1 and Landsat-8 images were taken on 17 December 2015.

## 2.2. Ancillary Datasets

### 2.2.1. Inundation Maps Derived from Landsat-8 Data

Landsat-8 satellite collects visible and shortwave images (30 m spatial resolution). NIR wavelength reflects less solar radiation than the red wavelength over water bodies [1,3], and surface water maps can be derived from the NDVI maps (water pixels and non-water pixels correspond to negative and positive values of NDVI, respectively) [24,25]. Other indices have been used to detect water, but the NDVI is effective when properly corrected from the atmospheric contamination. In this study, official and reliable atmospheric corrected Landsat-8 NDVI images are ordered directly from the U.S. Geological Survey (USGS) website (<https://espa.cr.usgs.gov/index/>). To limit cloud effects, only images with less than 10% of cloud contamination are used. The selected images are further filtered using the Landsat-8 quality assessment to remove pixels that might be affected by instrument artifacts or subject to cloud contamination. Figure 2 (bottom) shows the NDVI from Landsat-8 (d), the resulting surface water map based on negative NDVI values (e), and the quality flag (f), for the same regions and the same day (17 December 2015) as previously presented. Over the Lower Mekong Delta (lower than latitude number 15), there are ~250 Landsat-8 images available between January 2015 and January 2016. However, there is only ~10% (27 images) with less than 10% cloud contamination. Among the remaining images, only 1/3 was selected for this study since they were observed with a time difference of less than 3 days from a Sentinel-1 image.

### 2.2.2. Inundation Maps Derived from MODIS/Terra Data

In this study, the surface reflectance 8-Day L3 Global 500 m products from MODIS/Terra (MOD09A1) are used to create surface water maps, mainly based on values of the Enhanced Vegetation Index (EVI), the Land Surface Water Index (LSWI), and the difference between EVI and LSWI by a methodology described in [14]. MODIS surface water maps (500 m spatial resolution) over the Mekong Delta will be used to compare to the corresponding surface water maps derived from SAR Sentinel-1 observations for 2015. MODIS/Terra data can be downloaded from <http://reverb.echo.nasa.gov/reverb/>.

All Sentinel-1, Landsat-8 and MODIS/Terra observations used in this study are listed in Tables 1 and 2. Sentinel-1 and Landsat-8 training observations are used to train the NN (Section 3.2). Sentinel-1 and Landsat-8 test observations are used to test, optimize, and evaluate the performance of the NN (Sections 3.3 and 4.1). NN evaluation is also based on comparisons with MODIS surface water estimates (Section 4.3).

**Table 1.** List of 9 Sentinel-1 and corresponding Landsat-8 training (top) and test (bottom) observations used in this study over Cambodia and the Vietnamese Mekong Delta. Maximum gap between Sentinel-1 and Landsat-8 observations is only 3 days. The cloud cover percentage is indicated for each Landsat-8 observation.

Sentinel-1 and Landsat-8 Training Observations			
Image No	Sentinel-1	Landsat-8	Clouds
1	16 April 2015	14 April 2015	6.29%
2	21 April 2015	21 April 2015	0.05%
3	19 August 2015	18 August 2015	7.94%
4	17 December 2015	17 December 2015	4.84%
5	29 March 2016	31 March 2016	6.22%
6	9 June 2016	10 June 2016	3.94%
Sentinel-1 and Landsat-8 Test Observations			
Image No	Sentinel-1	Landsat-8	Clouds
1	5 January 2016	2 January 2016	0.16%
2	3 February 2016	3 February 2016	7.5%
3	22 February 2016	19 February 2016	0.29%

**Table 2.** List of 20 Sentinel-1 and corresponding MODIS/Terra observations used in this study over Cambodia and the Vietnamese Mekong Delta.

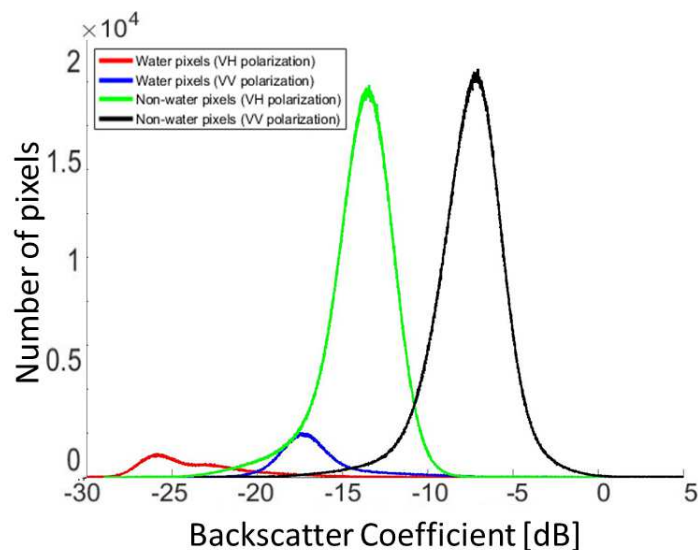
Sentinel-1 and MODIS/Terra Observations			
Image No	Date	Image No	Date
1	10 January 2015	11	14 August 2015
2	3 February 2015	12	26 August 2015
3	15 February 2015	13	7 September 2015
4	11 March 2015	14	19 September 2015
5	4 April 2015	15	1 October 2015
6	28 April 2015	16	13 October 2015
7	15 June 2015	17	25 October 2015
8	27 June 2015	18	6 November 2015
9	9 July 2015	19	30 November 2015
10	21 July 2015	20	24 December 2015

### 3. Methodology

#### 3.1. Surface Water Information from the Sentinel-1 SAR Images

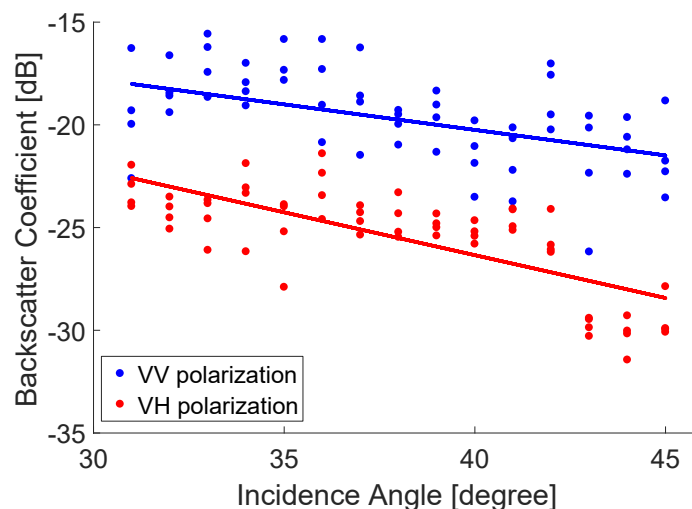
Flat water surfaces act like mirrors and reflect almost all incoming energy in the specular direction, thus providing very low backscatter. With this physical principle, detection of surface water is often based, at least partly, on the application of a threshold on the SAR backscatter coefficient, with the low backscatter values attributed to water bodies [6,7,16,17]. However, SAR backscatter coefficients over water surfaces are also affected by several mechanisms related to the interaction of the signal with vegetation or with possible surface roughness. The backscattered signals over flooded vegetation in wetlands can be enhanced due to the double-bounce scattering mechanism [26–28]. On the other side, the backscatter coefficients can be affected by vegetation canopy (e.g., rice) above the water surfaces due to volume scattering from the plant components (stems or leaves) [29]. The backscatter coefficients (especially the VV polarization) can also be influenced by the wind-induced surface roughness over open water [17,30]. Finally, there might be ambiguities between surface water and other very flat surfaces (such as arid regions), that could provide very similar backscatter signatures [31].

Based on a reference water mask derived from Landsat-8 NDVI, Figure 3 presents the histograms of the backscatter coefficients for VH and VV polarizations, separately for water and non-water pixels over the incidence angle range of  $30^{\circ}$ – $45^{\circ}$  for the area shown in Figure 2. For both polarizations, the water and non-water histograms are rather well separated, with thresholds of  $-22$  dB and  $-15$  dB for the VH and VV polarizations, respectively. Using these thresholds, the surface water has been classified separately for each polarization. The classification derived from the VH polarized image had a stronger spatial linear correlation with the reference water mask than the one derived from the VV polarized image (72% compared to 62%), confirming a higher sensitivity of the VH polarization to the presence of surface water [19]. Using both polarizations for the classification increased the correlation (76%), confirming that the two polarizations carry different information and that using both of them increases the retrieval accuracy. These findings confirmed the study by [32] where water detection with VV polarization was further refined using multiple-polarization.



**Figure 3.** For surface water delineated with Landsat-8, histograms of the water and non-water pixels for the SAR backscatter coefficients in VH and VV polarizations for the area shown in Figure 2 (over the incidence angle range of  $30^{\circ}$  to  $45^{\circ}$ ).

The effect of the backscatter incidence angle is also tested here. For a collection of pixels located over water (rivers, reservoirs, or lakes), the backscatter coefficient is plotted as a function of the incidence angle between  $30^\circ$  and  $45^\circ$  (Figure 4). Similar negative correlations between incidence angle and backscatter coefficients can also be found in [13] with ASAR data over water bodies (from  $\sim -5$  dB at  $20^\circ$  to  $\sim -20$  dB at  $45^\circ$  of incidence angle).



**Figure 4.** The SAR backscatter coefficients (VH and VV polarizations) from the Sentinel-1 as a function of the incidence angle over water bodies. The linear regression lines are also plotted.

As a conclusion, the SAR backscatter coefficients (VH and VV polarizations) are both sensitive to the presence of water, but with slightly different sensitivities. The effect of the incidence angle, although rather limited within the  $29^\circ$ – $46^\circ$  range of Sentinel-1 SAR, has to be accounted for if a high detection accuracy is required. Simple tests on thresholding techniques illustrated the limitations of these approaches and here we suggest developing a new scheme to delineate the surface water based on Neural Networks.

The temporal dynamics of the backscatter coefficients can also be a source of information and can help disentangle the influence of the other surface parameters [13]. However, this temporal information will not be investigated here.

### 3.2. A Neural Network-Based Classification

Here, we propose training a NN to produce surface water maps from SAR images, over the Mekong Delta. In the remote sensing field, NNs are often used as a regression tool to estimate a quantity. For each pixel, NN input satellite observations are represented by a vector  $x$ , and the network outputs (i.e., the retrieval) is represented by a vector  $y$ . However, NNs can also be used as classifiers. In this case, when trained with binary output values ( $y = 0$  for non-water, 1 for water surfaces), the NN becomes a statistical model for the conditional probability  $y = P(\text{surface} = \text{water} / x)$ , i.e., the probability of the surface being covered by water knowing the satellite observations  $x$ . The NN output can then directly be used as an index for water presence probability, but a threshold can also be applied to classify the state as being covered by water or not. The threshold needs to be optimized in order to satisfy some quality criteria, such as overall accuracy or false alarm rates.

The NN classifier needs to be trained in order to perform an optimal discrimination between water and non-water states. A supervised learning is chosen: the NN will be designed to reproduce an already existing classification. A dataset including a collection of SAR information  $x$  and associated surface water state  $y$  is first built. Part of it is then used during the training stage in order to determine the optimal parameters of the NN model. The reference dataset in the selected area is provided here by

a Landsat-8 surface water map (NN outputs), in spatial and temporal coincidence with the Sentinel-1 SAR data (NN inputs). A maximum time difference of 3 days is tolerated, as the two satellites do not fly in phase. Six Landsat-8 surface water maps are selected, along with the corresponding Sentinel-1 SAR observations (see Table 1 for more details on the training dataset). The selection process for the Landsat-8 images has been described in Section 2.2.1. The images cover parts of the lower Mekong Delta in Vietnam and Cambodia. For each image in the training dataset, the number of non-water pixels is much higher than the number of water pixels. To avoid giving too much weight to the non-water pixels, an equalization of the training dataset is performed: an equal number of non-water and water pixels is selected in the training dataset. For this purpose, non-water pixels are selected randomly in the images, to match the number of water pixels. The total number of training samples is  $\sim 10$  million pixels, half water pixels, half non-water pixels. It takes  $\sim 5$  h to train the NN (with the use of a personal computer), but when the training is completed, a surface water map can be produced quickly (after  $\sim 3$ – $4$  min) from any new set of satellite inputs  $x$ . A test dataset is chosen to measure the performance of the NN retrieval scheme with data not used in the training process. The NN methodology is summarized in Figure 5.

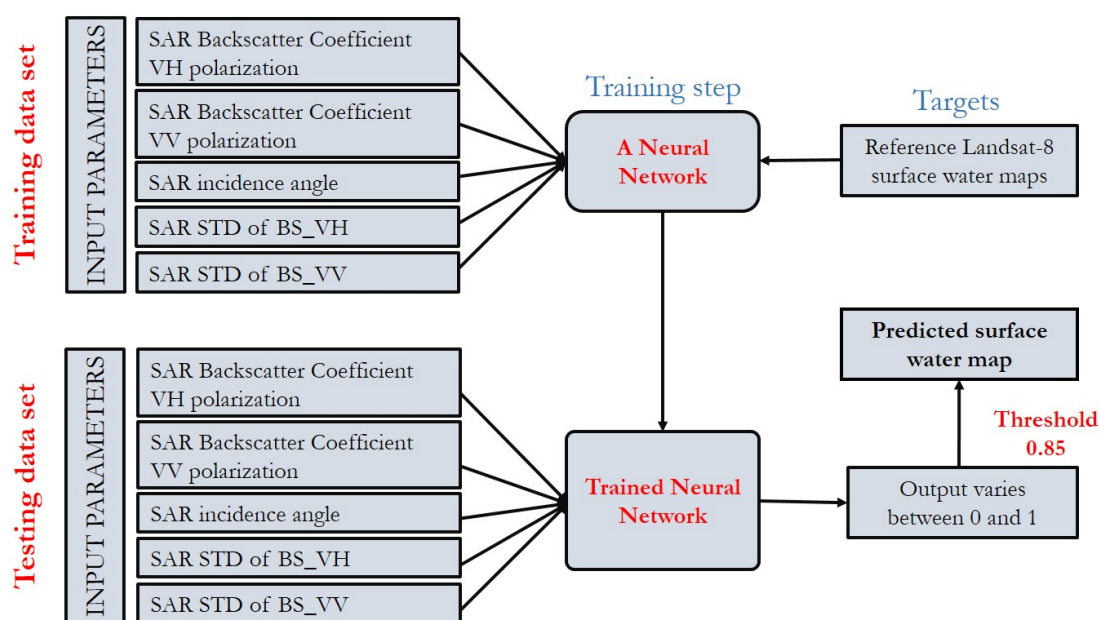
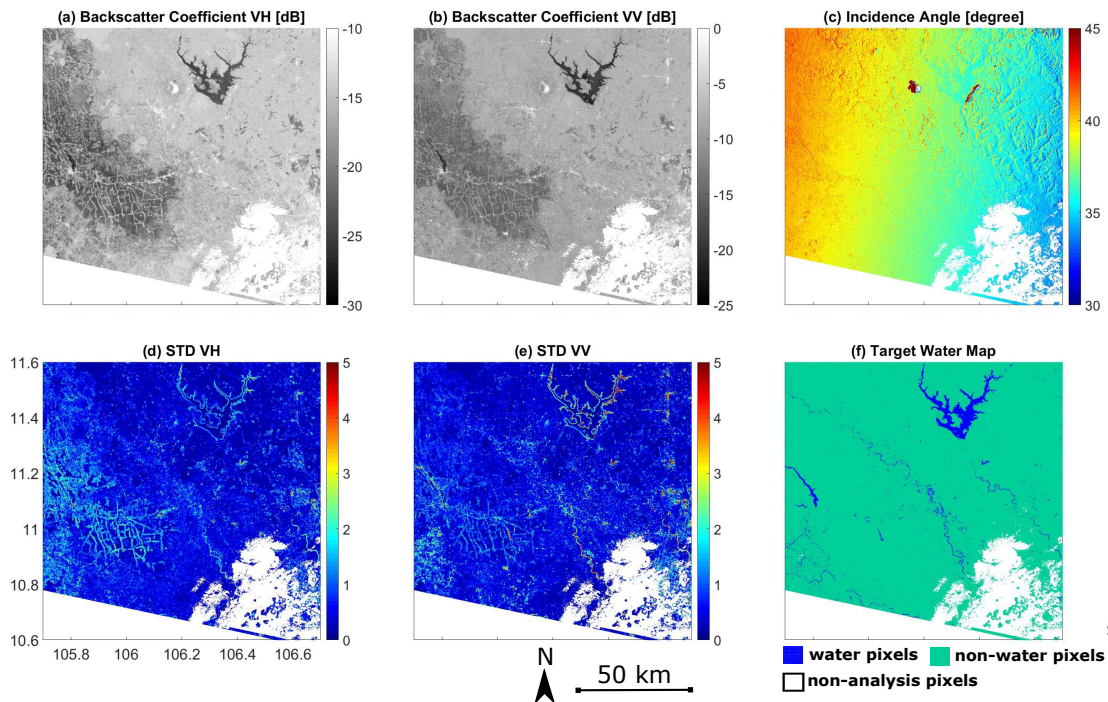


Figure 5. The block diagram of the proposed Neural Network (NN) algorithm.

Several tests were necessary to determine the optimum inputs to the NN, in addition to the obvious ones, i.e., the backscatter coefficients for both polarizations. To limit ambiguities between flat arid surfaces and surface water, and to better capture small rivers, the spatial homogeneity of the backscatter coefficients appeared to be a relevant parameter. The standard deviation of the backscatter coefficients are computed locally over  $100 \text{ m} \times 100 \text{ m}$  boxes. As a result, the NN uses a maximum of five different inputs  $x$ :

- SAR backscatter coefficient VH polarization (BS\_VH);
- SAR backscatter coefficient VV polarization (BS\_VV);
- SAR incidence angle;
- SAR standard deviation of backscatter coefficient VH over  $100 \text{ m} \times 100 \text{ m}$  (STD\_VH);
- SAR standard deviation of backscatter coefficient VV over  $100 \text{ m} \times 100 \text{ m}$  (STD\_VV);

Figure 6 presents an example of the set of five input images and the target surface water map used to train the NN. Missing areas in the maps correspond to Landsat-8 low quality pixels and are excluded from the training. The NN model is asked to find a relationship between these five input parameters and the corresponding water and non-water state.



**Figure 6.** Examples of the five inputs and the target for the NN. (a) SAR backscatter coefficient VH polarization; (b) SAR backscatter coefficient VV polarization; (c) SAR incidence angle; (d) SAR standard deviation of backscatter coefficient VH polarization; (e) SAR standard deviation of backscatter coefficient VV polarization; and (f) Target surface water map based on NDVI from Landsat-8. The white areas are cloud-covered pixels detected by the Landsat quality flags, and they have been removed. Sentinel-1 and Landsat-8 images were acquired on 16 and 14 April 2015, respectively.

### 3.3. NN Sensitivity Tests

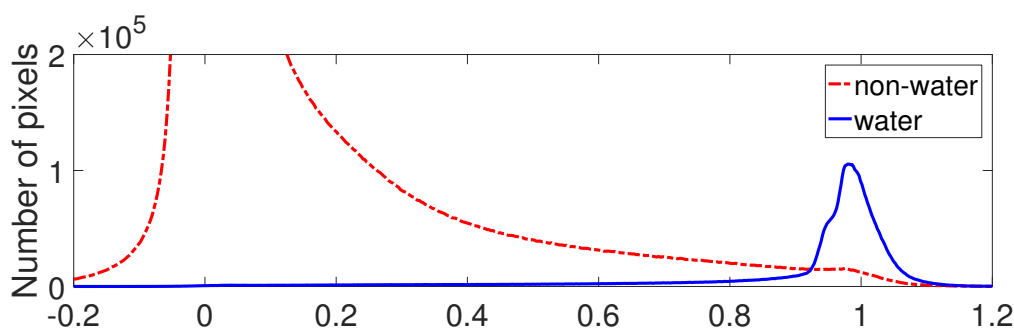
In this section, we use a test dataset of three SAR Sentinel-1 images and three corresponding Landsat-8 reference surface water maps to make several sensitivity tests in order to optimize the performance of the NN classification (see details of the test data sets in Table 1). Three different sensitivity tests were carried out: (1) selecting the best threshold of the NN output to classify land/water surface; (2) understanding the effect of the equalization of the water and non-water pixels in the NN training dataset; (3) finding the most important satellite NN inputs. The NN performances have been evaluated based on: spatial correlation between the SAR and Landsat-8 surface water maps, overall accuracy of the NN, as well as higher values of true positive (TP) and true negative (TN) percentages. True positive value indicates the NN ability to correctly detect water pixels, while true negative value illustrates its ability to correctly detect non-water pixels (compared to the Landsat-8 surface water maps).

#### 3.3.1. Selection of an Optimized Threshold for the NN Output

The first test is conducted to optimize the output threshold to distinguish water from non-water pixels. Figure 7 shows the histogram of the output of the NN, separating the water and non-water pixels according to the related Landsat-8 surface water map. The histograms of the water and non-water clusters intersect around 0.9, meaning that the optimal threshold to separate water from non-water pixels is close to this number. Different thresholds on the NN output values were tested (0.80, 0.85, and



0.90): for each one, the confusion matrix and the overall accuracy are calculated, with the corresponding Landsat-8 images as references. The overall accuracy and the spatial correlation increase from 98% to 99% when the threshold increases from 0.80 to 0.90 (Table 3), but the true positive pixel detection decreases from 92% (with threshold 0.80) to 89% (with threshold 0.90) and the false negative pixel detection increases from 8% to 11%. A threshold of 0.85 is selected here because of its good water detection performance and because it results in the predicted water surface closest to the reference map: 4430 km<sup>2</sup> from the Landsat-8 versus 4420 km<sup>2</sup> from the SAR results, i.e., a limited overestimation of 0.4% as compared to the reference map.



**Figure 7.** Histograms of the NN outputs, for water (blue) and non-water (dashed red) pixels separately, according to the corresponding Landsat-8 surface water maps. The NN uses the five initial inputs and the training dataset is equalized. The y axis range is selected to illustrate the peak of the water histogram.

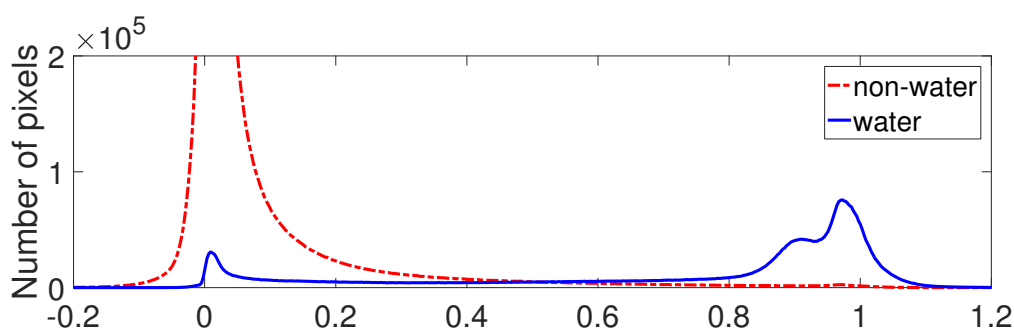
**Table 3.** Confusion matrix of the NN classification for different thresholds. The NN uses the five initial inputs and the training dataset is equalized.

Output Threshold: 0.80				
	Non-Water(0) (Predicted)	Water(1) (Predicted)	Overall Accuracy	Spatial Correlation
Non-water(0) (Actual)	99.3%	0.7%	98%	91%
Water(1) (Actual)	8%	92%		
Output Threshold: 0.85				
	Non-Water(0) (Predicted)	Water(1) (Predicted)	Overall Accuracy	Spatial Correlation
Non-water(0) (Actual)	99.5%	0.5%	99%	92%
Water(1) (Actual)	9%	91%		
Output Threshold: 0.90				
	Non-Water(0) (Predicted)	Water(1) (Predicted)	Overall Accuracy	Spatial Correlation
Non-water(0) (Actual)	99.6%	0.4%	99%	91%
Water(1) (Actual)	11%	89%		

### 3.3.2. Equalization of Water and Non-Water Pixel Number

For this test, instead of using an equal number of water and non-water pixels in the training dataset, 10% of each Sentinel-1 image is selected randomly to train the neural network, meaning that the number of non-water pixels is several times higher (10–15 times depending on each image in the training dataset) than the number of water pixels (as seen in Figure 7). The intersection between histograms of the NN outputs for water pixels (blue) and non-water pixels (red) moves to 0.5 (see the histogram in Figure 8), meaning that the value 0.5 should be selected to separate water from non-water

clusters. As shown in Table 4, the resulting NN is very efficient at detecting non-water pixels with a true negative detection of 99.7%, but it misses 14% of the actual water pixels (86% of true positive detection only, compared to 91% with the equalized training dataset—Table 3). The true positive detection of water pixels decreases because in the training database, the non-water pixels are more numerous and as such have more weight in the retrieval than the water pixels. As a consequence, the NN is more effective at detecting non-water pixels, and less effective at detecting water pixels. It is concluded that the use of an equalized training data set is very important in this classification framework.



**Figure 8.** Histograms of the NN outputs, for water (blue) and non-water (dashed red) pixels separately, according to the corresponding Landsat-8 surface water maps. The NN uses the five initial inputs but the training dataset is not equalized. The y axis range is selected to illustrate the peak of the water histogram.

**Table 4.** Confusion matrix of the NN without equalization of the training dataset, for an optimum threshold of 0.5 on the NN outputs.

	Non-water(0) (Predicted)	Water(1) (Predicted)
Non-water(0) (Actual)	99.7%	0.3%
Water(1) (Actual)	14%	86%

### 3.3.3. Analyzing the Weight of Each NN Satellite Input

To identify the most relevant inputs for the NN classification of the water surface, 15 NNs are trained based on all 15 different combinations of five input parameters, and their performances are evaluated following various criteria. Table 5 presents the best results with one to five inputs and illustrates how the overall accuracy of the NN classification increases when the number of satellite inputs increases, as compared to the reference Landsat-8 dataset. The NN trained with only the VH backscatter coefficient has a spatial correlation of 78% and a true positive accuracy (correctly detecting water pixels) of 77% compared to the reference data. The spatial correlation increases to 79%, and the true positive accuracy rises to 85% when the standard deviation of the VV backscatter coefficient is added as an input to the NN. The VV backscatter coefficient helps to increase the performance of the NN since both spatial correlation and true positive accuracy increase to 87% and 90%, respectively. The standard deviation of the VH backscatter coefficient does not significantly improve the accuracy of the NN classification. This is due to the strong linear correlation (88%) between the spatial standard deviations of the VH and the VV backscatter coefficients (the other linear correlations among the five input parameters of the NN are provided in Table 6). Similar to the standard deviation of the VH backscatter coefficient, the incidence angle does not have a strong impact on the performance of the NN since its accuracy remains nearly the same after adding the incidence angle as a new input. The input parameters of the NN classification are listed below, from the most important to the least important one in the NN processing:

- Backscatter coefficient VH polarization (BS\_VH)
- Standard deviation of backscatter coefficient VV polarization (STD\_VV)
- Backscatter coefficient VV polarization (BS\_VV)
- Incidence angle
- Standard deviation of backscatter coefficient VH polarization (STD\_VH)

**Table 5.** The NN classification performances when adding input parameters, one at a time.

<b>One Input: BS_VH</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	98%	2%	<b>78%</b>
<b>Water(1) (Actual)</b>	23%	77%	
<b>Two Inputs: BS_VH + STD_VV</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	98%	2%	<b>79%</b>
<b>Water(1) (Actual)</b>	15%	85%	
<b>Three Inputs: BS_VH + STD_VV + BS_VV</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99%	1%	<b>87%</b>
<b>Water(1) (Actual)</b>	10%	90%	
<b>Four Inputs: BS_VH + STD_VV + BS_VV + Angle</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-water(0) (Actual)</b>	99.5%	0.5%	<b>91%</b>
<b>Water(1) (Actual)</b>	10%	90%	
<b>Five Inputs: BS_VH + STD_VV + BS_VV + Angle + STD_VH</b>			
	<b>Non-Water(0) (Predicted)</b>	<b>Water(1) (Predicted)</b>	<b>Spatial Correlation</b>
<b>Non-Water(0) (Actual)</b>	99.5%	0.5%	<b>92%</b>
<b>Water(1) (Actual)</b>	9%	91%	

**Table 6.** Linear correlations among the five potential NN inputs.

	BS_VH	BS_VV	STD_VH	STD_VV	ANGLE
BS_VH	100%				
BS_VV	84%	100%			
STD_VH	24%	20%	100%		
STD_VV	21%	21%	88%	100%	
ANGLE	25%	22%	11%	6%	100%

To conclude, the water detection ability of the proposed NN increased when the input parameters are carefully selected and when an optimal output threshold is selected. An equal number of water and non-water pixels should be used in the training dataset to ensure that the NN performs equally well in classifying water and non-water clusters. The STD\_VH provides limited additional information to the NN due to its strong linear correlations with the other NN inputs. The incidence angle also plays a limited role in the NN performance. This is partly due to the rather narrow range of incidence angle, from 29° to 46°.

#### 4. Results and Comparisons with Other Surface Water Products

The following results and comparisons involve the optimized version of the NN classification with five input parameters (an equalization of water and non-water pixels, and the output threshold is 0.85). In Section 4.1, the SAR-predicted surface water maps are calculated for two test areas in the Mekong Delta, and compared to Landsat-8 surface water maps over the Tonle Sap Lake in Cambodia and over the Mekong river in Vietnam (see test dataset in Table 1). Other regions were tested but the results are not shown here. Due to the lack of in-situ local inundation maps at the time of this study, we do not have a reference dataset to confirm the accuracy of the Landsat-8 based maps. Therefore, an inter-comparison between Sentinel-1 estimate and other existing estimates is the only way to evaluate the new wetland product based on SAR Sentinel-1 data. First, the results are evaluated with respect to the floodability map derived mainly from the HydroSHEDS topography dataset [33], developed by [34] (Section 4.2). Second, time series of the SAR-derived surface water over the Mekong Delta is compared to the MODIS/Terra-derived inundation maps based on the methodology described by [14], for 2015 (Section 4.3)

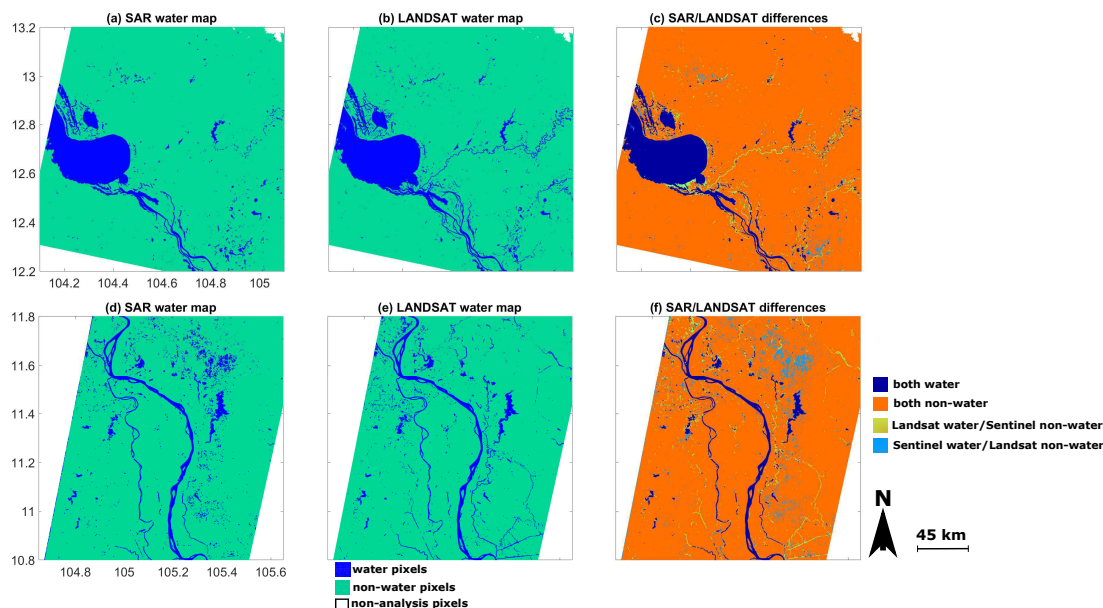
##### 4.1. Evaluation of the SAR NN Classification Method

Figure 9 shows the results of the classification applied over the Tonle Sap Lake in Cambodia (top) and over the Mekong river in Vietnam (bottom), in February 2016. Figure 9a,d show the SAR-predicted surface water maps, Figure 9b,e present the reference Landsat-8 surface water maps, whereas the differences between these two surface water maps are shown in Figure 9c,f.

Over the Tonle Sape Lake, both Sentinel and Landsat images were acquired on the same day (3 February 2016). The spatial correlation between the two surface water maps is 94%. The confusion matrix for this area is given in Table 7 (left). Overall accuracy of the classification is 99%, with a true positive water detection of 93.5%, and a false negative percentage of 6.5%. The classification correctly detects more than 99.6% of non-water pixels compared to the reference map. The classification slightly underestimates the surface water coverage by ~2.5%. This is 961 km<sup>2</sup> compared to the reference surface water map derived from the Landsat-8 images of 986 km<sup>2</sup>.

The second case study is carried out over the Mekong river and its surrounding areas (latitude range [10.8°N–11.8°N] and longitude range [104.6°E–105.6°E]). The optical Landsat-8 images were taken on 19 February 2016 and the SAR Sentinel-1 images were taken 3 days later, on 22 February 2016. These Sentinel and Landsat images were not acquired on the same day, but within 3 days in the middle of the dry season when land surfaces in this area are not expected to change much. Similar to the first case study, the classification works well, even though the environment here is rather complex, with rivers and vegetated wetlands. The overall accuracy is 98.8%, with a spatial correlation of nearly 82%

with the Landsat-8 reference surface water map. Confusion matrix for this area is shown in Table 7 (right) where the true positive percentage is 85.7%, the false negative percentage 14.3%, and 99.2% of non-water pixels are classified correctly. The total surface water area derived from Landsat data is 325 km<sup>2</sup>, and it is 355 km<sup>2</sup> predicted from the NN.



**Figure 9.** (a,d) SAR surface water maps; (b,e) Landsat-8 surface water maps; and (c,f) their differences; over the Tonle Sap Lake (left), and over the Mekong river (right), for February 2016. Blue color presents water pixels while orange color presents non-water pixels detected by both Sentinel and Landsat, green color is Landsat water/Sentinel non-water pixels, and light blue color is Sentinel water/Landsat non-water pixels.

**Table 7.** Confusion matrices (in numeric and percentage forms) of the SAR-predicted surface water maps and the Landsat-8 reference surface water maps, over the Tonle Sap Lake (Left) and over the Mekong River (Right).

	Tonle Sap Lake		Mekong River		
	Non-water(0) (Predicted)	Water(1) (Predicted)	Non-water(0) (Predicted)	Water(1) (Predicted)	
<b>Non-water(0)</b> (Actual)	11,641,078 (99.6%)	44,493 (0.4%)	<b>Non-water(0)</b> (Actual)	10,983,583 (99.2%)	85,096 (0.8%)
<b>Water(1)</b> (Actual)	71,884 (6.5%)	1,023,457 (93.5%)	<b>Water(1)</b> (Actual)	51,611 (14.3%)	309,982 (85.7%)

The same results are found when applying the NN classification to other areas. To conclude this comparison, the proposed NN methodology correctly detected ~90% of the water pixels observed by Landsat-8, with a spatial correlation of ~90%. The NN works better over open water bodies than over other heterogeneous environments. For instance, the NN has difficulties detecting small river branches (Southeast of the Tonle Sap Lake in Figure 9—top panel) although they are clearly detected with Landsat-8 images. The NN can provide water maps with high accuracy compared to the reference Landsat-8 water maps; there are differences between them. Errors could come from the following factors:

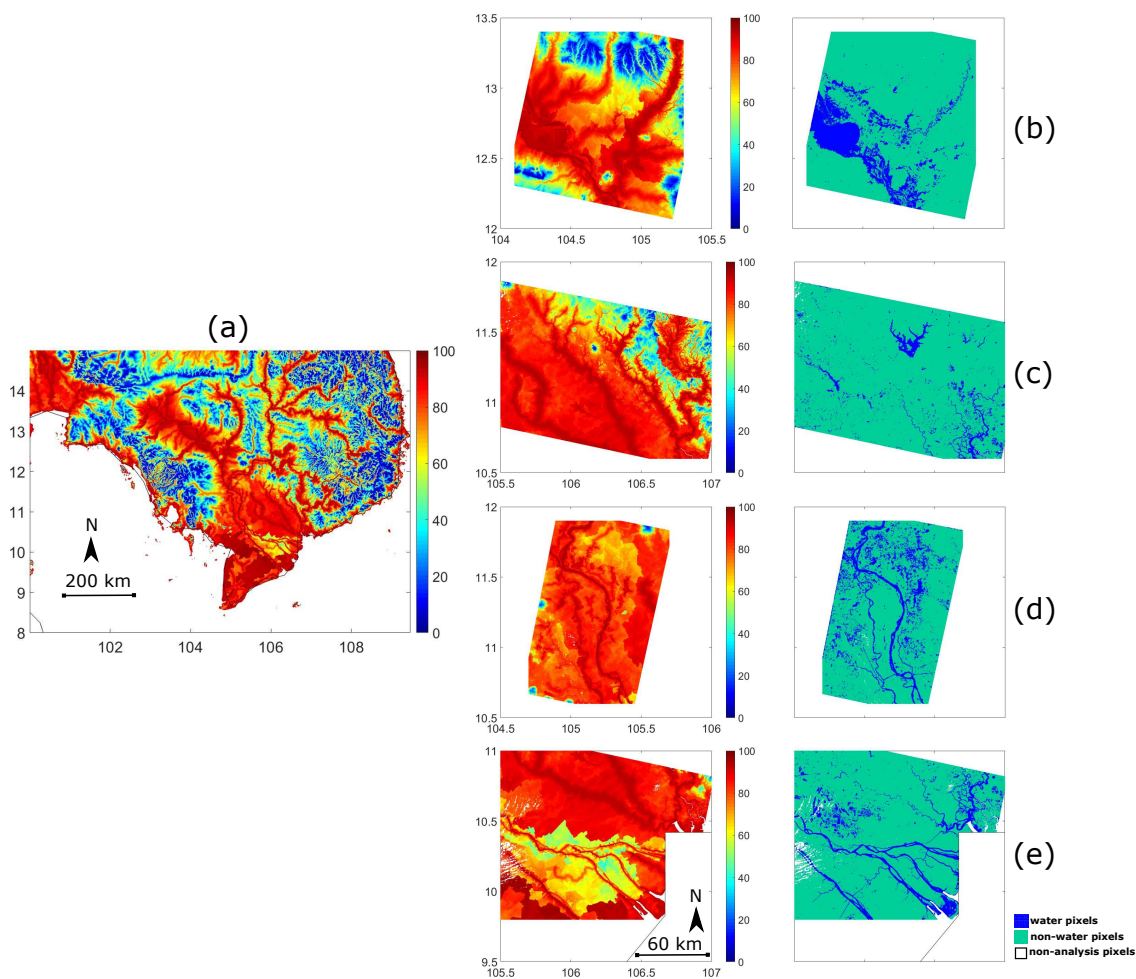
- The SAR responses can be affected by complex interactions with the terrain and the vegetation, especially along small river banks. It can be difficult to account for this local complexity in the methodology.
- In the SAR water detection method, as in any other classifications method scheme, different parameters were selected to optimize the overall performance of the method, but local ambiguities can still exist.
- Sentinel-1 and Landsat-8 data are not always acquired on the same day.
- Using Landsat-8 quality flags, we can remove cloud-covered pixels, but we cannot detect cloud-shadow pixels causing ambiguities in the NN training dataset.
- Reference surface water maps derived from negative NDVI values on the Landsat-8 images are not always perfect. Water under vegetation can be difficult to detect with Landsat-8 observations. The NDVI values can also be impacted for highly turbid waters where the NIR reflectance can be higher than the red reflectance.

#### 4.2. Evaluation Using a Topography-Based Floodability Index

A global floodability index based on topography has been developed by [34]. It uses mainly the Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS) dataset [33] that has been derived from elevation measured by the Shuttle Radar Topography Mission (SRTM) satellite. This floodability index provides a static map of an estimate of the probability for a pixel to be inundated (between 0% and 100%) at the spatial resolution of 90 m, based only on topography information (such as slope in the pixel, distance to the closest river, difference of elevation with the closest river). Figure 10a presents this floodability index map over the whole Mekong Delta. As expected, all rivers and lakes in this area have a very high probability of being inundated (over 80%). Since this index is based only on topography, its reliability is higher for natural environments and it can be less precise over regions with strong anthropic impact such as irrigated areas. The floodability data is upscaled from 90 m to 30 m spatial resolution to compare with predicted SAR surface water maps over the Tonle Sap Lake and the Vietnamese Mekong Delta. Each floodability pixel is divided into a  $3 \times 3$  matrix with the same value, and projected onto the Sentinel-1 grid. By comparing these two products, we can see where and how Sentinel-1 water pixels are located with respect to the floodability index, and test the consistency between two independent products. Figure 10b–e show floodability maps at 30 m spatial resolution and predicted Sentinel-1 water maps, over four different areas in the Mekong Delta. SAR surface water areas are generally located in areas with high predicted inundation probabilities, as expected (see Table 8). A total of 98% of the SAR surface water pixels are located in areas where the floodability index is greater than 60%, while only 2% of the SAR surface water pixels are located in areas with a lower floodability index ( $\leq 60\%$ ). As mentioned earlier, the floodability index only relies upon topography information, and it can be less precise over regions with strong anthropic activities, such as irrigation. There are many rice paddies in the Lower Mekong Delta, and these irrigated fields can be missed by the floodability index, contributing to the 15% errors of SAR water pixels located in areas with a floodability index less than 80%. In the future, in complex-topography environments where SAR only data could not provide the required accuracy for the water classification (the Red River Delta in the North of Vietnam, for example), the floodability index information could be added as another input to the NN to improve the classification performance.

**Table 8.** Performance of the SAR surface water classification for different ranges of floodability index.

Floodability Index	$\leq 40$	40–60	60–80	$\geq 80$
Percentage of surface water pixels detected by the NN classification	1%	1%	13%	85%



**Figure 10.** (a) Topography-based floodability index map over the Mekong Delta from [34]. (b–e) Comparisons of floodability index maps and SAR-predicted surface water maps for four areas over Cambodia and the Vietnamese Mekong Delta.

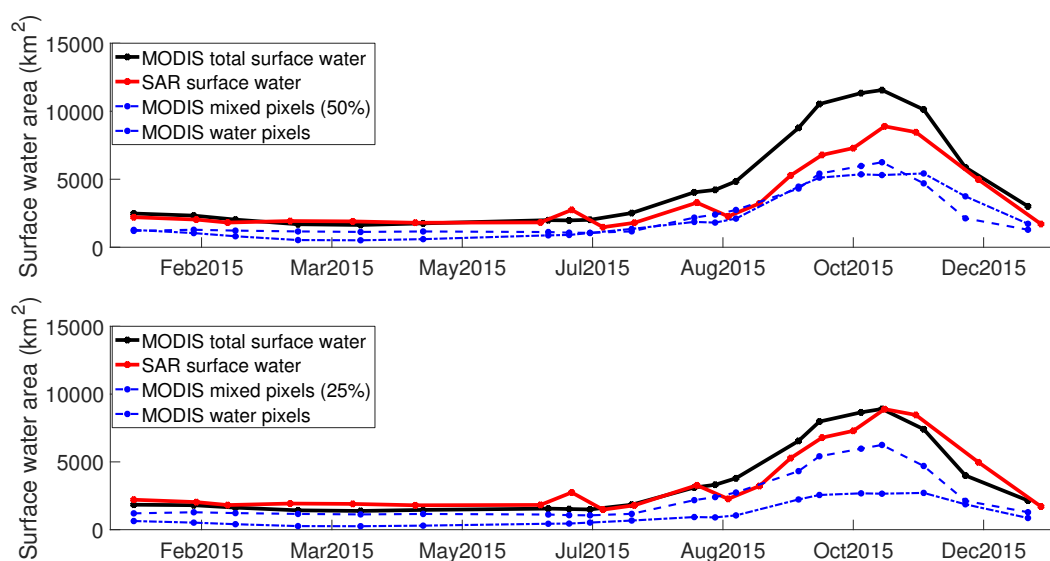
#### 4.3. Comparisons with MODIS/Terra-Derived Inundation Maps

In this section, the 30 m SAR surface water maps are compared to the 500 m MODIS/Terra-derived inundation maps, for a region in the Mekong Delta. One year (2015) of SAR Sentinel-1 and MODIS/Terra data are extracted, over the same region (latitude [9.8°N–11.3°N]; longitude [104.75°E–107°E]). The MODIS inundation maps are derived from the method described by [14]. We re-produced their methodology to calculate inundation maps with three different states of non-water, water, and mixed pixels, respectively. The total MODIS surface water is the sum of the water pixels (100% area is inundated) and mixed pixels (part of these pixels is inundated). For a mixed pixel, we tested two hypothesis: 25% or 50% of the pixel is inundated.

Twenty Sentinel-1 SAR observations are available over the selected region for the year 2015 (less than two images per month—see Table 2). The surface water extent calculated from the SAR and MODIS data are presented in Figure 11. With the first assumption (25% of a mixed MODIS pixel is covered by water), the two surface water extents have very similar seasonal cycles and amplitudes, with a correlation of ~99% (Figure 11-bottom). For the second assumption (the surface water extent of a mixed pixel is increased to 50%), the difference in surface water areas increases (without significant changes in the seasonal cycle with still high correlation with the SAR surface water time series). With both hypotheses, the SAR and MODIS surface water extents reach their maximum at the same time (around 20 October 2015). Total inundated areas derived from SAR and MODIS are very close

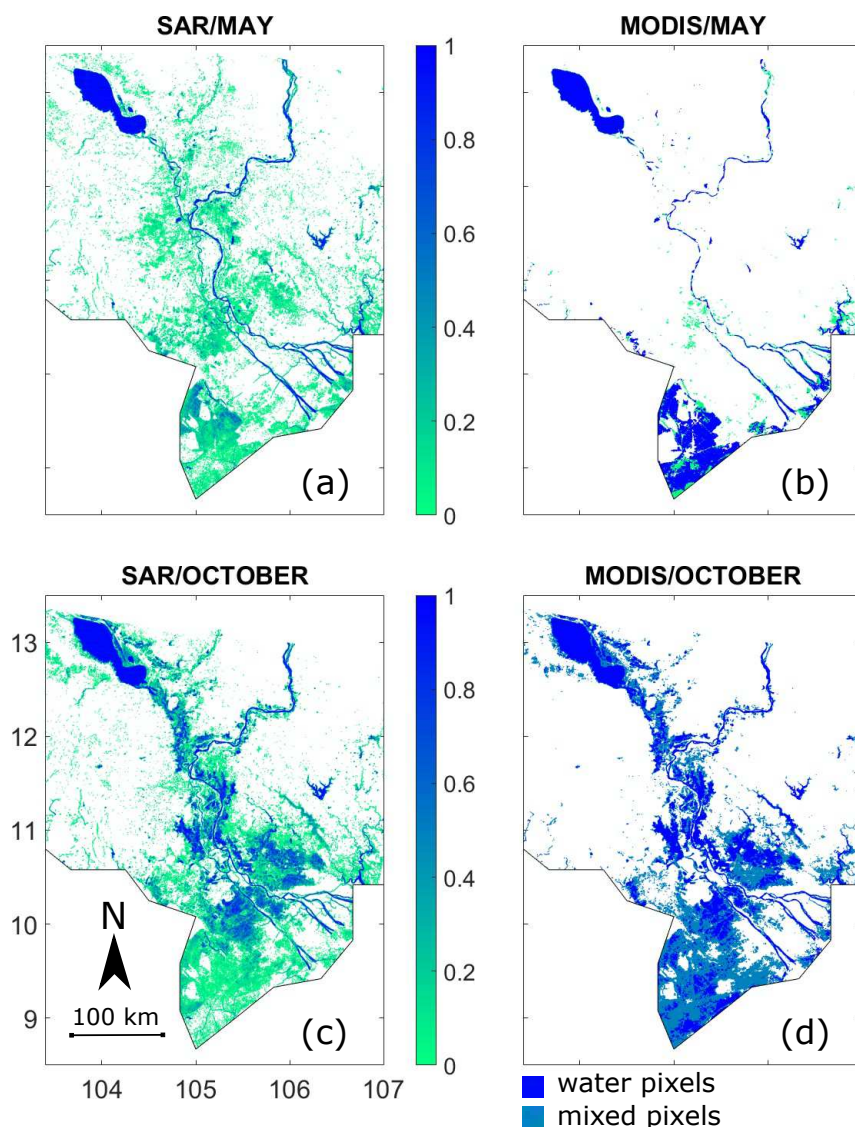
during the dry season (January to July). The cloud contamination of the MODIS estimate is low during that season. During the rainy season, more cloud contamination is expected in the MODIS estimates, and the discrepancies between the two surface water extents increase. The SAR-derived surface water estimate is expected to be more reliable due to its insensitivity to the cloud cover, but at this stage there is no convincing dataset at this spatial resolution to confirm it, as mentioned before.

To evaluate the consistency of the spatial structure between the SAR-derived and the MODIS-derived surface water maps, 10 SAR Sentinel-1 images were downloaded to cover the whole Mekong Delta and the Tonle Sap Lake (five images in May and five images in October 2015). For comparison purposes and to calculate the spatial correlation, the SAR surface water maps are aggregated from the 30 m resolution to the 500 m resolution of the MODIS-derived inundation maps (see Figure 12a,c). As a consequence, Sentinel-1-derived inundation maps are not binary (0 for non-water pixels or 1 for water pixels), but they are converted into a percentage of surface water at 500 m spatial resolution. For the dry season (Figure 12a,b—May 2015), the spatial correlation between the two surface water maps is 68%. A total of 4% of the area is inundated for the SAR estimation, while it is 5% for the MODIS estimates. For the rainy season (October 2015), the spatial correlation of the two maps increases to 78%, with 8% inundated area with the SAR and 11% with MODIS. For these calculations, we used the hypothesis of 25% inundation of the MODIS mixed pixels. Although SAR-derived and MODIS-derived water maps have a very similar seasonal cycle and similar spatial distribution of the water bodies, confirming the wetland seasonal cycle over this region, there are differences in the total surface of inundated areas. It comes mainly from the difference of spatial resolution between the two satellites. First, MODIS sensors cannot detect very small surface water fractions due to their spatial resolution. Second, the MODIS mixed pixels include water surfaces, vegetation surface and bare soil, and the percentage of each surface type within the pixel is not quantified.



**Figure 11.** Time series of the surface water detected by SAR (red) and MODIS data (black), over the Mekong Delta (Latitude [9.8°N–11.3°N]; Longitude [104.75°E–107°E]), for 2015. Two hypotheses are tested for the MODIS mixed pixels: 50% inundated (**top Panel**), and 25% inundated (**bottom Panel**).





**Figure 12.** (a,c) SAR and (b,d) MODIS surface water maps at 500 m resolution over the Mekong Delta in May (a,b) and October (c,d) 2015.

## 5. Conclusions and Perspectives

This study presents a methodology to monitor and quantify surface water under all weather conditions within Cambodia and the Mekong Delta in Vietnam, using high quality Sentinel-1 SAR observations, freely available online. The methodology is based on a neural network classification trained with optical Landsat-8 images at 30 m spatial resolution. The information content of each satellite input is analyzed and the inputs are selected to optimize the performance of the classification. This method allows the detection of surface water with good accuracy when compared to visible and NIR data under clear sky conditions, as well as when compared to a floodability map derived from topography data. Surface water maps derived from the proposed NN show a spatial correlation of  $\sim 90\%$  when compared to Landsat-8 water maps, with a true positive water detection of  $\sim 90\%$ . Compared to MODIS/Terra water maps over the Delta in 2015, our products share the same wetland seasonal cycle and dynamics, with a temporal correlation of  $\sim 99\%$ .

In the future, we will first apply the method to other areas under similar environments in southeast Asia and in other parts of the globe, and second we will test it in more vegetated environments. The final goal is to develop a general method capable of performing at the global scale to exploit the

full spatial coverage of the Sentinel-1 mission. For this purpose, several approaches will be tested to improve the retrieval scheme. First, the introduction of a priori information from a topography-based floodability index will increase information on flooding and reduce ambiguities in the SAR signal with other surface parameters. Second, with the launch of the optical Sentinel-2 satellite, Sentinel-2 observations could be used to replace Landsat-8 data, and to train the SAR surface water classification under clear sky conditions. The classification could then be extended to the cloudy areas using the SAR data. Third, the temporal information in the SAR backscatter could also be exploited (i.e., minimum or standard deviation of the time series) as this information has been shown to improve the detection of floods [13]. Finally, the high-resolution inundation extent retrieval maps could be post-processed in order to reduce the inherent noise in such high-spatial retrievals. We plan to test random walk techniques for that purpose.

**Acknowledgments:** We would like to thank Toshihiro Sakamoto for proving the methodology to derive the MODIS 500 m surface water maps. We are thankful to Simon Munier for interesting discussions and for sharing some of his tools in the framework of his ESA-funded postdoc (CCI Living Planet Fellowship AO/1-7829/14/I-MB). We also thank Thuy Le Toan, Stéphane Jacquemoud, Thanh Ngo Duc, and Nicolas Delbart for interesting discussions and suggestions for this work. This study was financially supported by a PhD fellowship from the Vietnam International Education Development (911 project). We would like to thank three anonymous reviewers for their helpful comments and suggestions that helped to improve the quality of this manuscript.

**Author Contributions:** Catherine Prigent, Filipe Aires and Binh Pham-Duc conceived and designed the experiments; Binh Pham-Duc performed the experiments; Catherine Prigent, Binh Pham-Duc and Filipe Aires analyzed the data; Filipe Aires, Catherine Prigent and Binh Pham-Duc contributed reagents/materials/analysis tools; Binh Pham-Duc, Catherine Prigent and Filipe Aires wrote the paper. All authors contributed to the discussion of results and the preparation of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432.
2. Bryant, R.G.; Rainey, M.P. Investigation of flood inundation on playas within the Zone of Chotts, using a time-series of AVHRR. *Remote Sens. Environ.* **2002**, *82*, 360–375.
3. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033.
4. Cretaux, J.F.; Berge-Nguyen, M.; Leblanc, M.; Abarca Del Rio, R.; Delclaux, F.; Mognard, N.; Lion, C.; Pandey, R.K.; Tweed, S.; Calmant, S.; et al. Flood mapping inferred from remote sensing data. *Int. Water Technol. J.* **2011**, *1*, 48–62.
5. Brisco, B.; Touzi, R.; Sanden, J.J.V.D.; Charbonneau, F.; Pultz, T.J.; D'Iorio, M. Water resource applications with RADARSAT-2: A preview. *Int. J. Digit. Earth* **2008**, *1*, 130–147.
6. Wang, Y. Seasonal change in the extent of inundation on floodplains detected by JERS-1 Synthetic Aperture Radar data. *Int. J. Remote Sens.* **2004**, *25*, 2497–2508.
7. Pierdicca, N.; Pulvirenti, L.; Chini, M.; Guerriero, L.; Candela, L. Observing floods from space: Experience gained from COSMO-SkyMed observations. *Acta Astronaut.* **2013**, *84*, 122–133.
8. Voormansik, K.; Praks, J.; Antropov, O.; Jagomagi, J.; Zalite, K. Flood Mapping With TerraSAR-X in Forested Regions in Estonia. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 562–577.
9. Martinis, S.; Kuenzer, C.; Wendleder, A.; Huth, J.; Twele, A.; Roth, A.; Dech, S. Comparing four operational SAR-based water and flood detection approaches. *Int. J. Remote Sens.* **2015**, *36*, 3519–3543.
10. Bartsch, A.; Pathe, C.; Wagner, W.; Scipal, K. Detection of permanent open water surfaces in central Siberia with ENVISAT ASAR wide swath data with special emphasis on the estimation of methane fluxes from tundra wetlands. *Hydrol. Res.* **2008**, *39*, 89–100.
11. Brisco, B.; Short, N.; van der Sanden, J.; Landry, R.; Raymond, D. A semi-automated tool for surface water mapping with RADARSAT-1. *Can. J. Remote Sens.* **2009**, *35*, 336–344.
12. Reschke, J.; Bartsch, A.; Schlaffer, S.; Schepaschenko, D. Capability of C-Band SAR for Operational Wetland Monitoring at High Latitudes. *Remote Sens.* **2012**, *4*, 2923–2943.

13. Santoro, M.; Wegmuller, U.; Lamarche, C.; Bontemps, S.; Defourny, P.; Arino, O. Strengths and weaknesses of multi-year Envisat ASAR backscatter measurements to map permanent open water bodies at global scale. *Remote Sens. Environ.* **2015**, *171*, 185–201.
14. Sakamoto, T.; Van Nguyen, N.; Kotera, A.; Ohno, H.; Ishitsuka, N.; Yokozawa, M. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sens. Environ.* **2007**, *109*, 295–313.
15. Leinenkugel, P.; Kuenzer, C.; Dech, S. Comparison and optimisation of MODIS cloud mask products for South East Asia. *Int. J. Remote Sens.* **2012**, *34*, 2730–2748.
16. Nguyen, L.; Bui, T. Flood Monitoring of Mekong River Delta, Vietnam using ERS SAR Data. In Proceedings of the 22nd Asian Conference on Remote Sensing, Singapore, 5–9 November 2001. Available online: <http://www.crisp.nus.edu.sg/~acrs2001/pdf/147nguye.pdf> (accessed on 22 May 2017).
17. Kuenzer, C.; Guo, H.; Huth, J.; Leinenkugel, P.; Li, X.; Dech, S. Flood Mapping and Flood Dynamics of the Mekong Delta: ENVISAT-ASAR-WSM Based Time Series Analyses. *Remote Sens.* **2013**, *5*, 687–715.
18. Amitrano, D.; Martino, G.D.; Iodice, A.; Mitidieri, F.; Papa, M.N.; Riccio, D.; Ruello, G. Sentinel-1 for Monitoring Reservoirs: A Performance Analysis. *Remote Sens.* **2014**, *6*, 10676–10693.
19. Santoro, M.; Wegmuller, U.; Wiesmann, A.; Lamarche, C.; Bontemps, S.; Defourny, P.; Arino, O. Assessing Envisat ASAR and Sentinel-1 multi-temporal observations to map open water bodies. In Proceedings of the 2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Marina Bay Sands, Singapore, 1–4 September 2015; pp. 614–619.
20. ESA. Sentinel-1 Technical Guides. Available online: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar> (accessed on 22 May 2017).
21. Sentinel Scientific Data Hub. Available online: <https://scihub.copernicus.eu/> (accessed on 22 May 2017).
22. SAR Basics with the Sentinel-1 Toolbox in SNAP Tutorial. Available online: <http://step.esa.int/main/doc/tutorials/> (accessed on 22 May 2017).
23. Liu, C. Analysis of Sentinel-1 SAR Data for Mapping Standing Water in the Twente Region. Master Thesis on Science in Geo-information Science and Earth Observation, University of Twente, Twente, The Netherlands, February 2016. Available online: [http://www.itc.nl/library/papers\\_2016/msc/wrem/cliu.pdf](http://www.itc.nl/library/papers_2016/msc/wrem/cliu.pdf) (accessed on 22 May 2017).
24. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *351*, 309.
25. Rokni, K.; Ahmad, A.; Selamat, A.; Hazini, S. Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery. *Remote Sens.* **2014**, *6*, 4173–4189.
26. Hess, L.L.; Melack, J.M.; Simonett, D.S. Radar detection of flooding beneath the forest canopy: A review. *Int. J. Remote Sens.* **1990**, *11*, 1313–1325.
27. Kasischke, E.S.; Bourgeau-Chavez, L.L. Monitoring South Florida Wetlands Using ERS-1 SAR Imagery. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 281–291.
28. Pope, K.O.; Rejmankova, E.; Paris, J.F.; Woodruff, R. Detecting seasonal flooding cycles in marshes of the Yucatan Peninsula with SIR-C polarimetric radar imagery. *Remote Sens. Environ.* **1997**, *59*, 157–166.
29. Liu, Y.; Chen, K.S.; Xu, P.; Li, Z.L. Modeling and Characteristics of Microwave Backscattering From Rice Canopy Over Growth Stages. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6757–6770.
30. Gstaiger, V.; Huth, J.; Gebhardt, S.; Wehrmann, T.; Kuenzer, C. Multi-sensoral and automated derivation of inundated areas using TerraSAR-X and ENVISAT ASAR data. *Int. J. Remote Sens.* **2012**, *33*, 7291–7304.
31. Prigent, C.; Aires, F.; Jimenez, C.; Papa, F.; Roger, J. Multiangle Backscattering Observations of Continental Surfaces in Ku-Band (13 GHz) From Satellites: Understanding the Signals, Particularly in Arid Regions. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1364–1373.
32. Henry, J.B.; Chastanet, P.; Fellah, K.; Desnos, Y.L. Envisat multi-polarized ASAR data for flood mapping. *Int. J. Remote Sens.* **2006**, *27*, 1921–1929.

33. Lehner, B.; Verdin, K.; Jarvis, A. HydroSHEDS Technical Documentation, Version 1.0; World Wildlife Fund US: Washington, DC, USA. Available online: [https://hydrosheds.cr.usgs.gov/webappcontent/HydroSHEDS\\_TechDoc\\_v10.pdf](https://hydrosheds.cr.usgs.gov/webappcontent/HydroSHEDS_TechDoc_v10.pdf) (accessed on 22 May 2017).
34. Aires, F.; Miolane, L.; Prigent, C.; Pham-Duc, B.; Fluet-Chouinard, E.; Lerner, B.; Papa, F. A Global Dynamic Long-Term Inundation Extent Dataset at High Spatial Resolution Derived through Downscaling of Satellite Observations. *J. Hydrometeorol.* **2017**, doi:10.1175/JHM-D-16-0155.1.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



# **Appendix B**

## **Paper B: Comparisons of Global Terrestrial Surface Water Datasets over 15 Years**

**Authors:**

B. Pham-Duc, C. Prigent, F. Aires, F. Papa

**Paper published in:**

Journal of Hydrometeorology, 2017



## Comparisons of Global Terrestrial Surface Water Datasets over 15 Years

BINH PHAM-DUC

*Laboratoire d'Études du Rayonnement et de la Matière en Astrophysique et Atmosphères, CNRS, Observatoire de Paris, Paris, France, and Space and Aeronautics Department, University of Science and Technology of Hanoi, Hanoi, Vietnam*

CATHERINE PRIGENT AND FILIPE AIRES

*Laboratoire d'Études du Rayonnement et de la Matière en Astrophysique et Atmosphères, CNRS, Observatoire de Paris, Paris, France*

FABRICE PAPA

*LEGOS, Université de Toulouse, IRD, CNES, CNRS, UPS, Toulouse, France, and Indo-French Cell for Water Sciences, IRD-IISc-NIO-IITM, Indian Institute of Science, Bangalore, India*

(Manuscript received 23 August 2016, in final form 5 January 2017)


### ABSTRACT

Continental surface water extents and dynamics are key information to model Earth's hydrological and biochemical cycles. This study presents global and regional comparisons between two multisatellite surface water extent datasets, the Global Inundation Extent from Multi-Satellites (GIEMS) and the Surface Water Microwave Product Series (SWAMPS), for the 1993–2007 period, along with two widely used static inundation datasets, the Global Lakes and Wetlands Database (GLWD) and the Matthews and Fung wetland estimates. Maximum surface water extents derived from these datasets are largely different:  $\sim 13 \times 10^6 \text{ km}^2$  from GLWD,  $\sim 5.3 \times 10^6 \text{ km}^2$  from Matthews and Fung,  $\sim 6.2 \times 10^6 \text{ km}^2$  from GIEMS, and  $\sim 10.3 \times 10^6 \text{ km}^2$  from SWAMPS. SWAMPS global maximum surface extent reduces by nearly 51% (to  $\sim 5 \times 10^6 \text{ km}^2$ ) when applying a coastal filter, showing a strong contamination in this retrieval over the coastal regions. Anomalous surface waters are also detected with SWAMPS over desert areas. The seasonal amplitude of the GIEMS surface waters is much larger than the SWAMPS estimates, and GIEMS dynamics is more consistent with other hydrological variables such as the river discharge. Over the Amazon basin, GIEMS and SWAMPS show a very high time series correlation (95%), but with SWAMPS maximum extent half the size of that from GIEMS and from previous synthetic aperture radar estimates. Over the Niger basin, SWAMPS seasonal cycle is out of phase with both GIEMS and MODIS-derived water extent estimates, as well as with river discharge data.

### 1. Introduction

Continental surface waters only cover a few percent of the land surface (Lehner and Döll 2004; Downing et al. 2006; Prigent et al. 2007), but they have a strong impact on the environment, as well as on human life (Vorosmarty et al. 2010). Surface freshwaters comprise

wetlands, rice paddies, rivers, lakes, reservoirs, and episodically inundated areas. Note that the definition of wetlands varies with applications and there is not an overall consensus on the subject (e.g., Reichhardt 1995). Surface waters play a key role in the biogeochemical and hydrological cycles, in biodiversity, and in climate variability. They show very diverse natures, from wetlands to inundated urban areas or rice paddies, associated with different dynamics from the tropics to the boreal regions. Wetlands are considered one of the most biologically diverse of all ecosystems since they support plant and wild animal species during important states of their life cycles. They are the world's largest natural source of methane ( $\text{CH}_4$ ), and they provide about one-third of the total

 Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JHM-D-16-0206.s1>.

Corresponding author e-mail: Binh Pham-Duc, [pham.binh@obspm.fr](mailto:pham.binh@obspm.fr)

DOI: 10.1175/JHM-D-16-0206.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](#).



global emission ( $\sim 165 \text{ Tgyr}^{-1}$ ; Bousquet et al. 2006; Bridgham et al. 2013; Wania et al. 2013). Consequently, monitoring surface freshwater extent and dynamics is a high priority in water management and climate research (e.g., Alsdorf et al. 2007).

Nevertheless, our understanding about the global distribution of the surface waters and their dynamics is still limited, with only a few datasets providing information at the global scale. Efforts have been made to collect information on water surfaces to produce static maps of surface waters, for example, the Global Lakes and Wetlands Database (GLWD) from Lehner and Döll (2004) or the Matthews and Fung (1987) wetland dataset. These maps are representative of the maximum surface water extent, and they do not provide any information on the temporal dynamics. Satellite observations in the visible, infrared, or microwave domains have the potential to detect surface water extent and their variations (Verpoorter et al. 2014; Yamazaki et al. 2015; Mueller et al. 2016; Feng et al. 2016), with different degrees of success depending on the environments. Optical and near-infrared satellite measurements provide good spatial resolutions but are limited by their inability to penetrate clouds and dense vegetation. Microwaves, passive or active, have the ability to penetrate clouds and vegetation, to a certain extent. With synthetic aperture radar (SAR) data, high spatial resolution is obtained, but global products describing the surface water dynamics are not available yet. Passive microwave observations have long been used to detect surface water extents, but used alone, it is difficult to disentangle the vegetation contribution from the measured signal. Prigent et al. (2016) propose a review on the use of different satellite techniques to monitor surface water, discussing in detail their advantages and limitations regarding the diverse applications.

These considerations lead to the conclusion that there is currently not a unique satellite technique for detecting surface water dynamics globally from tropical to boreal regions. In the following, surface waters will include all surface water types (wetlands, rice paddies, rivers, lakes, reservoirs, and episodically inundated areas), as the satellite observations cannot distinguish between the different natures of the surface water. A multisatellite methodology has been developed to derive surface water extent and dynamics at the global scale, benefiting from complementary strengths of satellite observations in the visible, passive, and active microwaves (Prigent et al. 2001, 2007, 2012; Papa et al. 2010): the Global Inundation Extent from Multi-Satellites (GIEMS). More recently, the Surface Water Microwave Product Series (SWAMPS) has been produced, also based on the merging of passive and active microwave satellite

observations (Schroeder et al. 2015). Long time series of global surface water estimates are necessary today to analyze the changes in the wetland-related methane emission (e.g., Melton et al. 2013; Wania et al. 2013), and climate modelers are in strong need of wetland extent information to understand the methane variability over past decades (Ringeval et al. 2010; Pison et al. 2013) for a better prediction of its evolution in the upcoming decades. GIEMS has been extensively evaluated and is limited to 1993–2007 (time extension is underway). SWAMPS is a recent dataset that extends from 1992 to 2013. They have a similar spatial resolution ( $\sim 0.25^\circ \times 0.25^\circ$ ). As these two datasets are currently the only global surface water datasets with monthly time series, a thorough comparison of these estimates is needed for the user community.

Here, global surface water datasets are systematically and objectively compared, including the two multisatellite databases, GIEMS and SWAMPS, along with two static datasets. The analysis covers the common period of the two satellite-derived products (1993–2007), and both the spatial and temporal variabilities of the databases are studied. The surface water datasets are described in section 2. Global and regional comparisons are described and discussed in section 3. A discussion is presented in section 4, and section 5 concludes this study.

## 2. Data

### a. GIEMS

GIEMS was the first global surface water dataset that provided monthly distribution of wetland and surface water extent (including lakes, rivers, and irrigated agriculture). GIEMS data cover the period 1993–2007 and are mapped on an equal-area grid of  $0.25^\circ \times 0.25^\circ$  spatial resolution at the equator (pixels of  $773 \text{ km}^2$ ). Inundated surfaces were detected and their extent was estimated by the method developed by Prigent et al. (2001, 2007, 2012). Global monthly inundation maps are derived from daily data of the following satellite observations: 1) passive microwave emissivity from the Special Sensor Microwave Imager (SSM/I), 2) active microwave backscatter coefficients from the scatterometer on board the European Remote-Sensing (ERS) satellite, and 3) the normalized difference vegetation index (NDVI) derived from visible and near-IR reflectances of the Advanced Very High Resolution Radiometer (AVHRR). Instead of directly using the brightness temperatures from the passive microwave instruments, surface emissivities are calculated to avoid modulation of the signal by atmospheric effects and surface temperature variations (Prigent et al. 2006). Passive microwave emissivities

from SSM/I are primarily used to detect inundation of the land surface while active microwave backscatter is used to assess the vegetation contribution to the passive microwave signal. NDVI information is necessary to distinguish between bare surfaces and inundated surfaces in semiarid regions where they can produce similar passive microwave signatures. Because of the lack of continuity and consistency in the ERS and AVHRR products, the current version of GIEMS uses monthly-mean climatology of ERS and AVHRR, calculated over 1993–2000. Other solutions were carefully tested, such as using other instruments like QuikSCAT instead of ERS, but this was not satisfactory (Papa et al. 2010). The snow-covered areas are filtered out using the National Snow and Ice Data Center (NSIDC) dataset (Brodzik and Armstrong 2013; <http://nsidc.org/data/NSIDC-0046>). Inland seas (Caspian Sea and Aral Sea), big lakes (e.g., Great Lakes in North America), and coastal pixels (possibly contaminated by radiation from the ocean because of the large microwave fields of view) were also carefully suppressed from the GIEMS data. GIEMS has been thoroughly evaluated by comparison with static surface water databases, and its consistency with other hydrological information (e.g., precipitation and river height) has been assessed (Papa et al. 2006, 2007, 2008a, 2010). Recent works have been performed to downscale GIEMS using ancillary high-resolution data. GIEMS-D15 has a 15-arc-s resolution (nearly 500 m at the equator) for three different temporal states of the inundation extent (Fluet-Chouinard et al. 2015). First, the original 12-yr time series of GIEMS (1993–2004) were aggregated to get mean annual minimum, mean annual maximum, and long-term maximum at the pixel level. Second, GLWD data (Lehner and Döll 2004) were added to supplement missing data or to correct GIEMS underestimation for low water fraction. Finally, a global inundation probability map derived from the HydroSHEDS dataset (Lehner et al. 2008) was used to downscale GIEMS from the original resolution to 15 arc s (Fluet-Chouinard et al. 2015). In this project, we averaged GIEMS-D15 to the GIEMS grid ( $0.25^\circ$  at the equator) for comparison with the other datasets. More recently, a new version (GIEMS-D3) at 3-arc-s resolution (Aires et al. 2017) has been produced with an improved downscaling scheme.

#### b. SWAMPS

SWAMPS is a daily global surface water dataset, mapping open water areas and water under low-density vegetation, for the period 1992–2013 (Schroeder et al. 2015), on an equal-area grid of 25 km resolution at the equator (similar to GIEMS). It is derived from combined passive and active microwave observations: SSM/I

and the Special Sensor Microwave Imager/Sounder (SSM/I/S) for the passive microwave and the backscatter coefficients from ERS, QuikSCAT, and the Advanced Scatterometer (ASCAT). The microwave polarization difference index (MPDI), the ratio of the difference between the brightness temperatures in the two orthogonal polarizations over their sum, is the main indicator to detect the water bodies in this methodology. Backscatter coefficients from the ERS (5 GHz), QuikSCAT (10 GHz), and ASCAT (5 GHz) satellites are used to reduce the vegetation effect on the MPDI. The three instruments do not observe with the same angles and frequencies, and ad hoc corrections are performed to limit the effects of the changes on the time series (Schroeder et al. 2015). Snow cover and frozen ground are filtered using the method of Grody and Basist (1996) and Chang et al. (1987). Daily global SWAMPS data were averaged to obtain monthly global SWAMPS estimates at  $0.25^\circ \times 0.25^\circ$  for comparison with the other datasets.

#### c. GLWD

GLWD is a global open water and wetland dataset developed by Lehner and Döll (2004). The dataset is derived from the combination of a variety of existing maps and information. Among these, the Digital Chart of the World (DCW) of the Environmental System Research Institute (ESRI) was the main source map to identify lakes and reservoirs. GLWD is not a satellite product, but a static map with three data levels: GLWD-1, GLWD-2, and GLWD-3. The GLWD-3 data used here include three main types of open water (rivers, lakes, and reservoirs) and nine different natural wetland classes in the form of a global raster map at 30-s resolution. For each wetland category (50%–100%, 25%–50%, and 0%–25%) the average value (i.e., 75%, 37.5%, and 12.5%) is used. GLWD is designed to represent the maximum level of surface water extent regionally and globally. GLWD has been used extensively and compared favorably with different satellite-based land-cover datasets (Nakaegawa 2012). For comparison with the other datasets, GLWD is aggregated to the GIEMS grid.

#### d. Other ancillary datasets

Matthews and Fung (1987) wetland fractions come from aeronautical charts, the information for which is more likely acquired during warm seasons of maximum flooding (Matthews and Fung 1987). It is a static dataset on a  $1^\circ$  regular grid. It has been extensively used in the past by climate groups to estimate the methane emission from wetlands.

Satellite products such as GIEMS and SWAMPS detect all surface water, including inundated areas

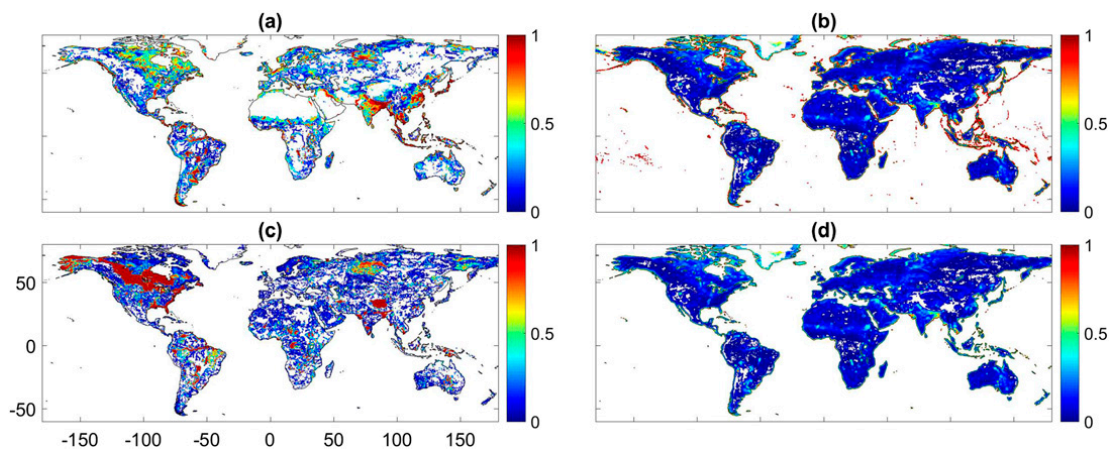


FIG. 1. Global maps of fractional surface water for different datasets on the equal-area grid at  $0.25^\circ \times 0.25^\circ$  at the equator ( $773 \text{ km}^2$  pixels). (a) GIEMS long-term monthly-mean max over 1993–2007, (b) original SWAMPS long-term monthly-mean max over 1993–2007, (c) GLWD, and (d) SWAMPS long-term monthly-mean max over 1993–2007 after coastal filtering.

associated with cultivation (such as the rice paddies in Asia) that are not accounted for in GLWD or in the [Matthews and Fung \(1987\)](#) wetland dataset. The global monthly irrigated and rain-fed crop areas (MIRCA) around the year 2000 ([Portmann et al. 2010](#)) provide information about irrigated and rain-fed agriculture globally. Irrigated rice is inundated during most of its growing season, and the MIRCA dataset provides an estimate of these inundated surfaces that are especially important in Asia ([Adam et al. 2010](#)). MIRCA is averaged from the 5-arc-min resolution to the common  $0.25^\circ$ -resolution grid.

River discharge data in some important basins of the world (e.g., Amazon, Orinoco, Niger, Mississippi, Congo, and Ob) were collected as an ancillary source of information to compare with time series of the surface water datasets. In the present study, we show the comparisons with river discharge for three basins we focused on (Amazon, Niger, and Ob Rivers). For the Amazon River, we use the in situ monthly discharges observed at Obidos, Brazil, which is the closest gauge to the mouth of the river ( $\sim 800 \text{ km}$  from the ocean), and for which data are available for 1993–2007 at the Observation Service for the Geodynamical, Hydrological and Biogeochemical Control of Erosion/Alteration and Material Transport in the Amazon, Orinoco, and Congo basins (SO HYBAM) website (<http://www.ore-hybam.org/>). Daily river discharge at the Lokoja gauge (1998–2005) is used to get the monthly river discharge data for the Niger basin. The data are collected from the Global Runoff Data Centre (GRDC; [http://www.bafg.de/GRDC/EN/Home/homepage\\_node.html](http://www.bafg.de/GRDC/EN/Home/homepage_node.html)). Over the Ob River, we use the in situ records from the Russian Hydrometeorological Service that are available on

a monthly basis until 2004 in the archives of the R-Arctic project (<http://www.r-arcticnet.sr.unh.edu/v4.0/index.html>).

### 3. Comparisons of the surface water datasets

The satellite-derived surface water datasets are compared over their common period 1993–2007, first globally and then at a basin scale. The 23 largest river basins in the world have been analyzed. However, the results are presented only for three basins located in contrasted types of environments: the Amazon basin in the tropics, the Niger basin in a semiarid area, and the Ob basin in the boreal region. The comparison is also systematically performed with the two static datasets previously described: GLWD and the [Matthews and Fung \(1987\)](#) dataset.

#### a. Global comparisons

[Figure 1a](#) shows the GIEMS long-term monthly-mean maximum inundation for each pixel over 1993–2007, along with the SWAMPS equivalent information ([Fig. 1b](#)), for comparison with GLWD ([Fig. 1c](#)). Even at this scale, large differences are evident between the three datasets. GIEMS and GLWD show much larger inland water fractions than SWAMPS. GLWD has particularly large inundation extent in Canada, where many small lakes are located. The major large river floodings (e.g., Amazon, Orinoco, and Ganges–Brahmaputra) appear clearly on both GIEMS and GLWD maps. The large water fraction in SWAMPS is concentrated on the coastal region (see, e.g., Indonesia). That is very likely related to the contamination of the retrieval by the ocean. Close to the coast, part of the energy observed by the microwave instruments can

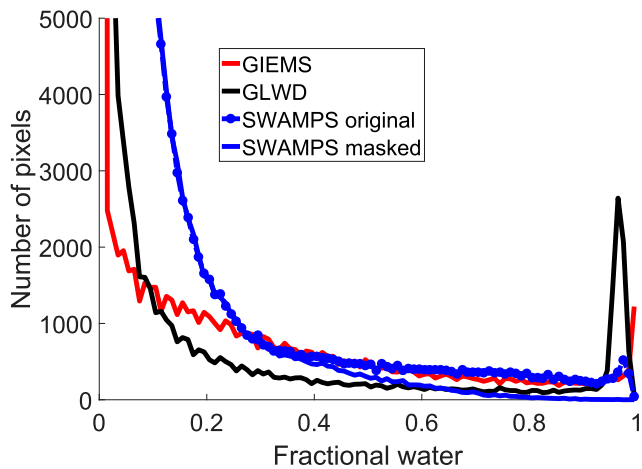


FIG. 2. Histograms of long-term max fractional water surfaces for the four datasets in Fig. 1.

come from the ocean, and the signal can be misinterpreted as coming from terrestrial inundated surfaces, if a careful filtering is not applied. Here we propose to filter the SWAMPS data to eliminate the ocean contamination close to the coast. Figure 1d represents the SWAMPS data where the contaminated coastal pixels are masked. SWAMPS also detects water almost everywhere on the globe, even in the North African desert. Histograms of the maximum fractional water surface are presented in Fig. 2 for the four datasets in Fig. 1. GLWD shows a large number of highly inundated pixels (>90%), mostly located in Canada (see Fig. 1c). SWAMPS has a very large number of fractional water surfaces below 0.2, much

more than the two other datasets. However, it has much less large water fractions, especially after filtering of the coasts.

For each pixel and each satellite-derived dataset, the mean fractional inundation at annual maximum and minimum has been calculated, along with the mean yearly amplitude of the fractional inundation (Fig. 3; note that for SWAMPS the coastal filtering is applied). Maps of the differences of these values between GIEMS and SWAMPS data are also presented. Compared to GIEMS, SWAMPS shows very limited amplitude in the annual cycle of the inundation, even in regions where large seasonal variations are expected.

Time series of the surface water from GIEMS and SWAMPS are compared globally and for three latitude bands [tropical (30°S–30°N), midlatitude (30°–55°N), and boreal (55°–70°N)] in Fig. 4, along with the corresponding values from GLWD, from the Matthews and Fung (1987) wetland dataset, and from the irrigated fields from MIRCA. GLWD (black) shows the maximum level of surface water both globally ( $\sim 13 \times 10^6 \text{ km}^2$ ) and regionally. This is expected as it is representative of the maximum inundation and it has a better spatial resolution than the Matthews and Fung (1987) dataset (that is also expected to present the maximum inundation, but with a much lower spatial resolution and thus likely to miss the small water surfaces). The Matthews and Fung (1987) dataset (cyan) has much less water extent in the boreal and the midlatitude regions, as it is representative of the wetlands only and does not include the lakes and rivers ( $\sim 5.3 \times 10^6 \text{ km}^2$ ). The maximum global surface water extent derived from

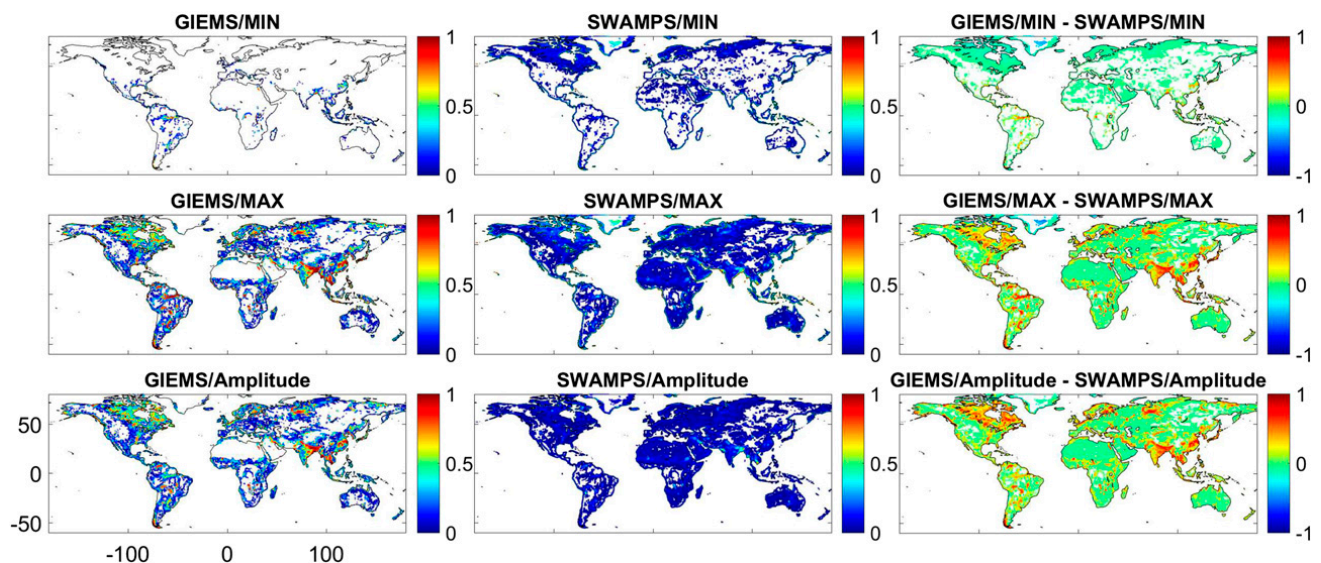


FIG. 3. Global-mean annual (top) min and (middle) max of the fractional inundation and (bottom) amplitude for (left) GIEMS, (center) SWAMPS, and (right) their differences. The information is presented on the  $773 \text{ km}^2$  equal-area grid.

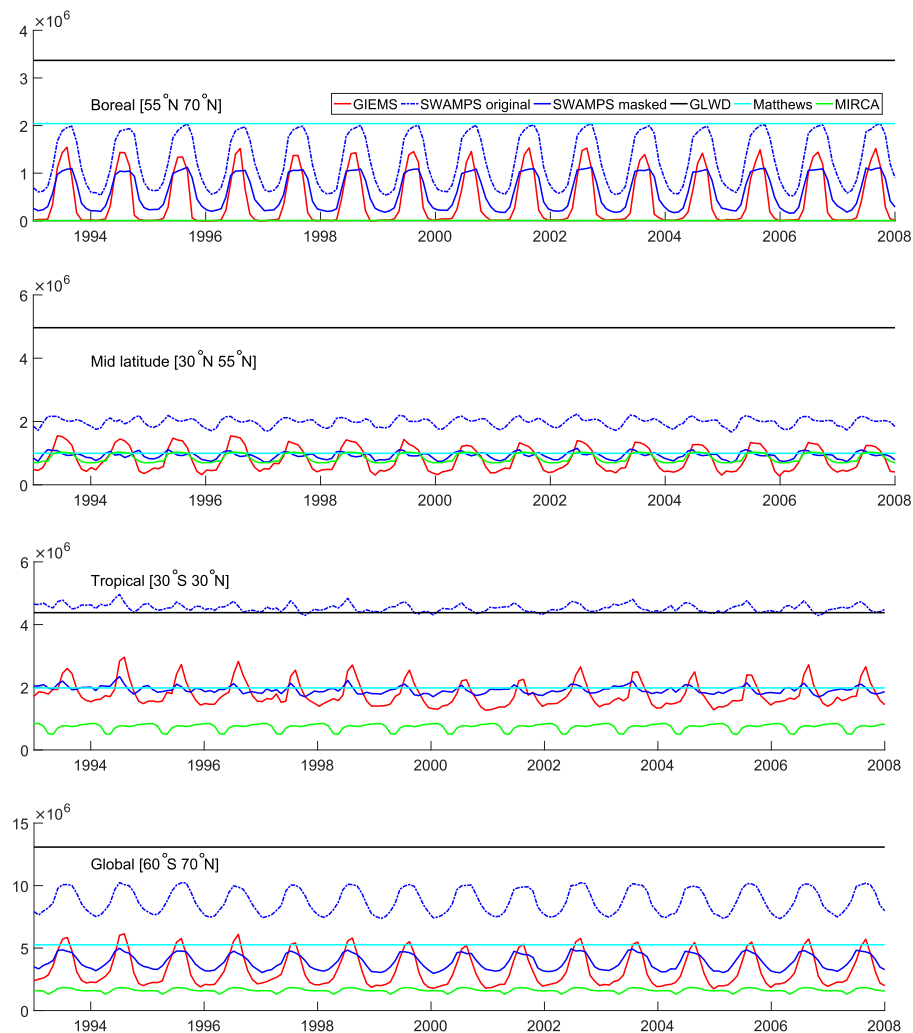


FIG. 4. Time series of surface water extent derived from GIEMS (red), original SWAMPS (dashed blue), SWAMPS with coastal filtering (blue), GLWD (black), [Matthews and Fung \(1987\)](#) dataset (cyan), and MIRCA irrigated fields (green), at three latitude bands and globally, for the period 1993–2007.

GIEMS is  $\sim 6.2 \times 10^6 \text{ km}^2$ , smaller than that derived from GLWD. From comparison with high-resolution SAR estimates over the Amazon basin, the accuracy of the GIEMS product has been estimated at 10% of the  $773 \text{ km}^2$  equal-area grid: because of the low-spatial-resolution satellite sensors, GIEMS tends to miss the small water fraction below 10% of the pixels or to overestimate the large inundation fraction above 90% ([Prigent et al. 2007](#)). However, this problem is partly corrected in GIEMS-D15 ([Fluet-Chouinard et al. 2015](#)), which merges GIEMS low-inundated pixels with GLWD. The maximum of the original SWAMPS inundation extent at global scale ( $\sim 10.3 \times 10^6 \text{ km}^2$ ) is close to the GLWD extent, as mentioned in [Schroeder et al. \(2015\)](#). Over the tropical region, the maximum surface water derived from the original SWAMPS (dashed blue) is very close to that of GWLD (black),

$\sim 4.9 \times 10^6$  and  $\sim 4.4 \times 10^6 \text{ km}^2$ , respectively. However, it is surprising to observe that the global maximum value is reduced to  $\sim 5 \times 10^6 \text{ km}^2$  (i.e., a reduction of nearly 51%) when the coasts are filtered out. Over the tropical region, SWAMPS decreases by nearly 53% after coastal masking to reach  $\sim 2.3 \times 10^6 \text{ km}^2$ . The original SWAMPS dataset detects more water surfaces than GIEMS for all three latitude bands. After coastal filtering, the yearly-mean temporal SWAMPS water surface extent is similar to the GIEMS one, but the seasonal amplitude of SWAMPS is much lower than the GIEMS one. [Table 1](#) shows monthly long-term minimum, long-term maximum, and long-term mean, as well as seasonal amplitude between monthly long-term maximum and minimum of surface water extent derived from GIEMS and SWAMPS after coastal filtering globally, and at three latitude bands, for the studied period (1993–2007).

TABLE 1. Monthly long-term min, long-term max, long-term mean, and seasonal amplitude between monthly long-term max and min of surface water extent derived from GIEMS and SWAMPS after coastal filtering at three latitude bands and globally, for the period 1993–2007. Unit is  $\times 10^6 \text{ km}^2$ .

	Min	Max	Mean	Amplitude
Boreal				
GIEMS	0.0	1.5	0.4	1.5
SWAMPS	0.2	1.2	0.6	1.0
Midlatitude				
GIEMS	0.3	1.6	0.8	1.3
SWAMPS	0.7	1.2	0.9	0.5
Tropical				
GIEMS	1.3	3.0	1.8	1.7
SWAMPS	1.7	2.3	1.9	0.6
Global				
GIEMS	1.7	6.2	3.2	4.5
SWAMPS	3.0	5.0	3.9	2.0

Long-term mean values between GIEMS and SWAMPS are close, but GIEMS shows larger amplitude than SWAMPS regionally and globally. From Fig. 4 we can see that wetland surfaces derived from GIEMS (red) and SWAMPS (dashed blue) have strong temporal correlations globally (92%) and over the boreal region (88%), but lower temporal correlations over mid-latitudes (58%) and the tropics (48%). Furthermore, in

these two regions, GIEMS has a much stronger seasonal cycle than SWAMPS.

In the rest of the study, the coastal filter is applied to the SWAMPS data.

*b. Basin-scale comparisons*

1) COMPARISONS OVER THE AMAZON BASIN

The Amazon basin is the largest drainage basin in the world with the largest discharge, and it is mostly located in the tropical rain forest. Figure 5 shows the long-term maximum inundation maps (1993–2007) over the Amazon basin, derived from the different wetland datasets. Spatial distributions of surface water datasets are similar. Although the spatial correlation between long-term maximum inundated maps of GIEMS and SWAMPS is nearly 90%, SWAMPS fractional surface water is much lower than the GIEMS and GLWD ones. In Schroeder et al. (2015), it is noted that SWAMPS has problems detecting water underneath dense forest canopy. The ability of passive microwaves to detect surface water below dense forest was demonstrated early by Giddings and Choudhury (1989) or Sippel et al. (1994) in their pioneer works. It is rather surprising that SWAMPS cannot detect these surface waters. GIEMS-D15 corrects GIEMS by adding the small surface water fractions that are likely

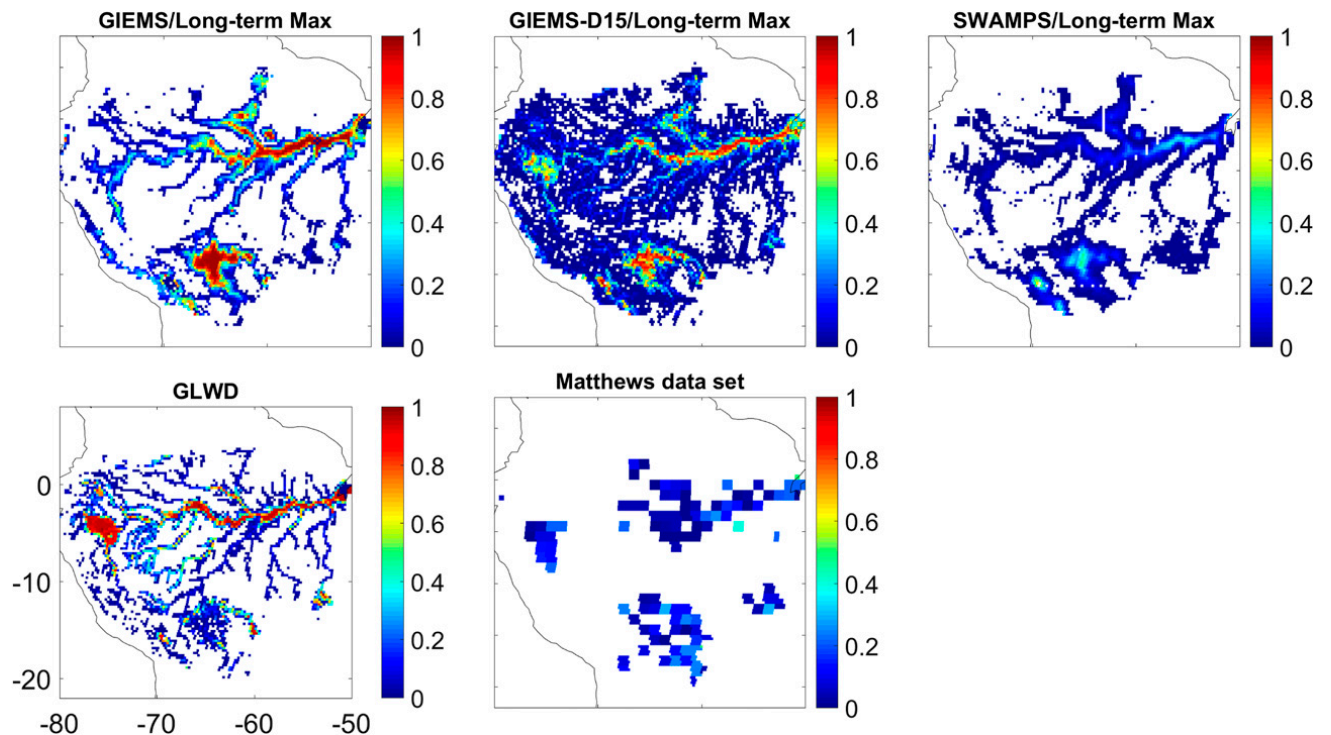


FIG. 5. Long-term max inundation maps averaged over 1993–2007 from GIEMS and SWAMPS, and max inundation maps from the different surface water datasets over the Amazon basin. Shown are GIEMS, GIEMS-D15, SWAMPS, GLWD, and Matthews and Fung (1987) datasets.

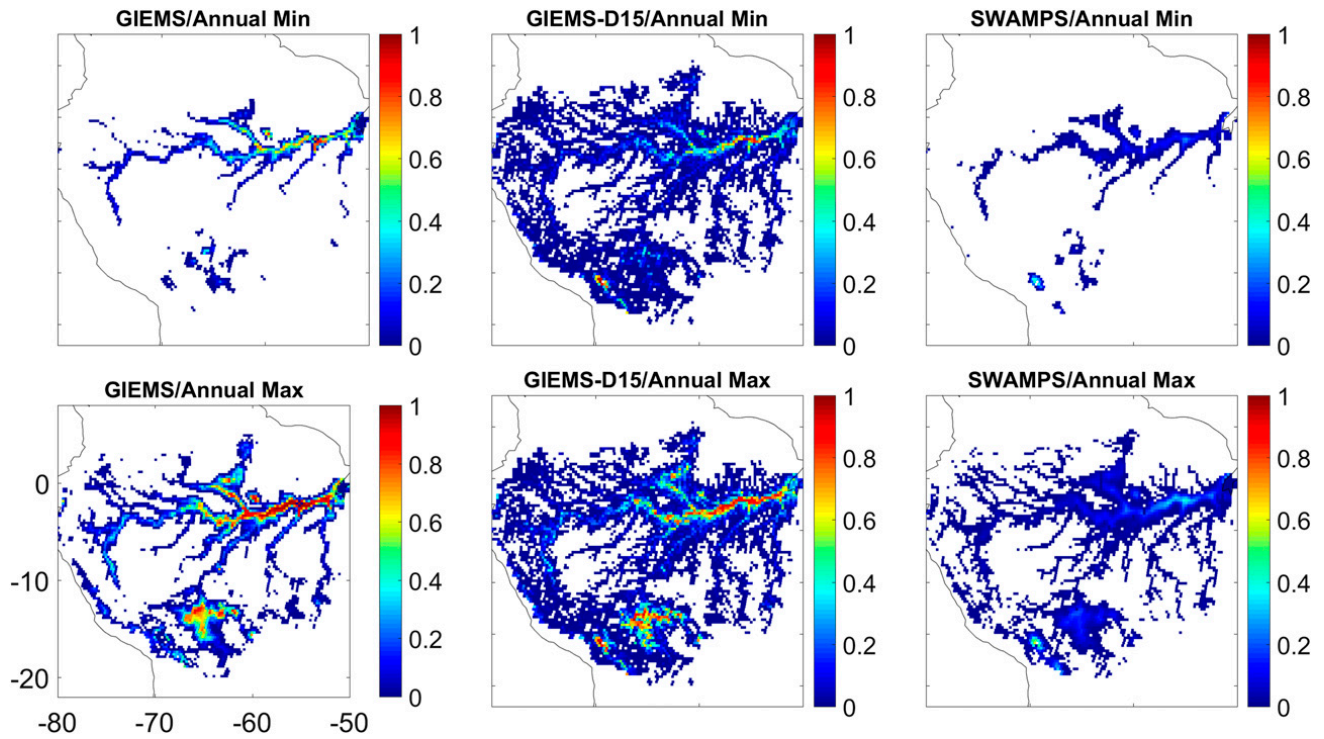


FIG. 6. Mean annual (top) min and (bottom) max of the inundation derived from different surface water datasets over the Amazon basin for (left) GIEMS, (center) GIEMS-D15, and (right) SWAMPS.

missed by GIEMS, and this is clearly seen when comparing GIEMS and GIEMS-D15 inundation maps in Fig. 5.

Mean annual minimum and maximum inundation maps of GIEMS, GIEMS-D15, and SWAMPS are shown in Fig. 6. Similar to the long-term maximum inundated maps, the spatial distributions of the GIEMS and SWAMPS datasets are similar (spatial correlation

of 80% and 90% for the minimum and maximum, respectively), but SWAMPS detects much less surface water than the two GIEMS versions.

Monthly time series of the surface water extents have been calculated over the basin, along with the river discharge at the mouth of the river (Fig. 7, top). The monthly-mean annual cycle from January to December

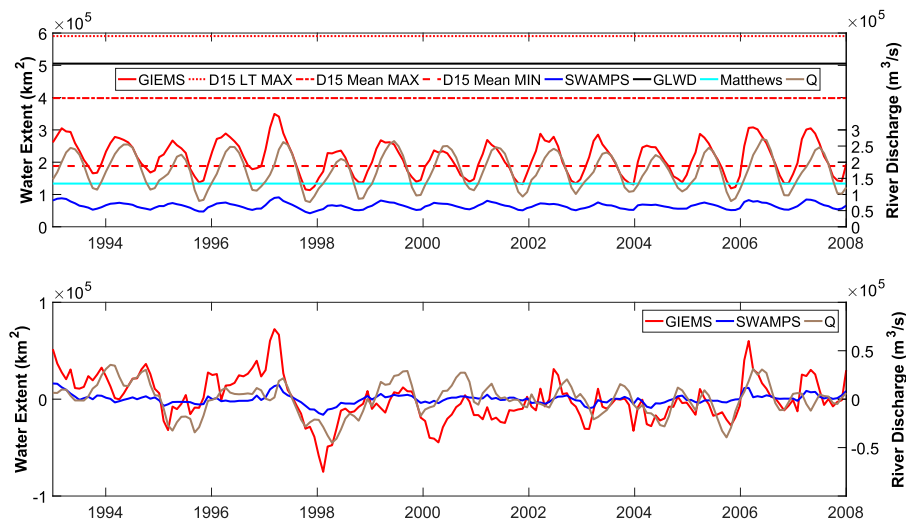


FIG. 7. (top) Time series and (bottom) anomaly of surface water extent derived from the different wetland datasets, along with the river discharge over the Amazon basin from 1993 to 2007. Anomaly is calculated by removing the monthly-mean annual cycle from time series.

TABLE 2. Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge  $Q$  over the Amazon basin for the period 1993–2007. Numbers in parentheses are calculated with 1-month lag between GIEMS/SWAMPS and  $Q$ .

	Time series correlation	Anomaly correlation
GIEMS/SWAMPS	95%	77%
GIEMS/ $Q$	78% (91%)	54% (58%)
SWAMPS/ $Q$	74% (88%)	57% (61%)

is computed by averaging values from all Januaries to Decembers between 1993 and 2007. Then these values are subtracted from the time series for each given month, to obtain the anomaly time series (Fig. 7, bottom). Time series and anomaly correlations are shown in Table 2. The behaviors of time series of GIEMS (red) and SWAMPS (blue) are very similar to that of the river discharge (brown), with very high correlations. Time series correlation between GIEMS and SWAMPS is 95%. It is 78% and 74% with the river discharge for GIEMS and SWAMPS, respectively. Time series correlations with the river discharge increase when calculated with 1-month lag, as expected, reaching 91% and 88%, respectively. The anomaly correlation between GIEMS and SWAMPS is surprisingly high (77%). The

same analysis has been carried out over the Orinoco basin, showing a correlation of 99% between time series of GIEMS and SWAMPS and a correlation of 97% between their anomalies. More details can be found in the supplemental material.

## 2) COMPARISONS OVER THE NIGER BASIN

The Niger basin is characterized by a large inner delta that results in a region of braided streams and has marked seasonal floods. Long-term maximum inundation maps are shown over the Niger basin in Fig. 8, while their minimum and maximum are presented in Fig. 9, and Fig. 10 shows their time series and anomalies, as well as that derived from the river discharge data. Time series and anomaly correlations between GIEMS/SWAMPS and other ancillary datasets are shown in Table 3. Time series of GIEMS and SWAMPS are in opposite phase (Fig. 10), making the time series correlation negative (−40%). Again, GIEMS shows a much stronger seasonal cycle than SWAMPS over this basin. GIEMS and the river discharge (brown) show similar behavior with a time series correlation of nearly 81% (for the common period 1998–2005). In contrast, SWAMPS does not show the same seasonal cycle, making its time series correlation with the river discharge negative (−60%). Note that other studies

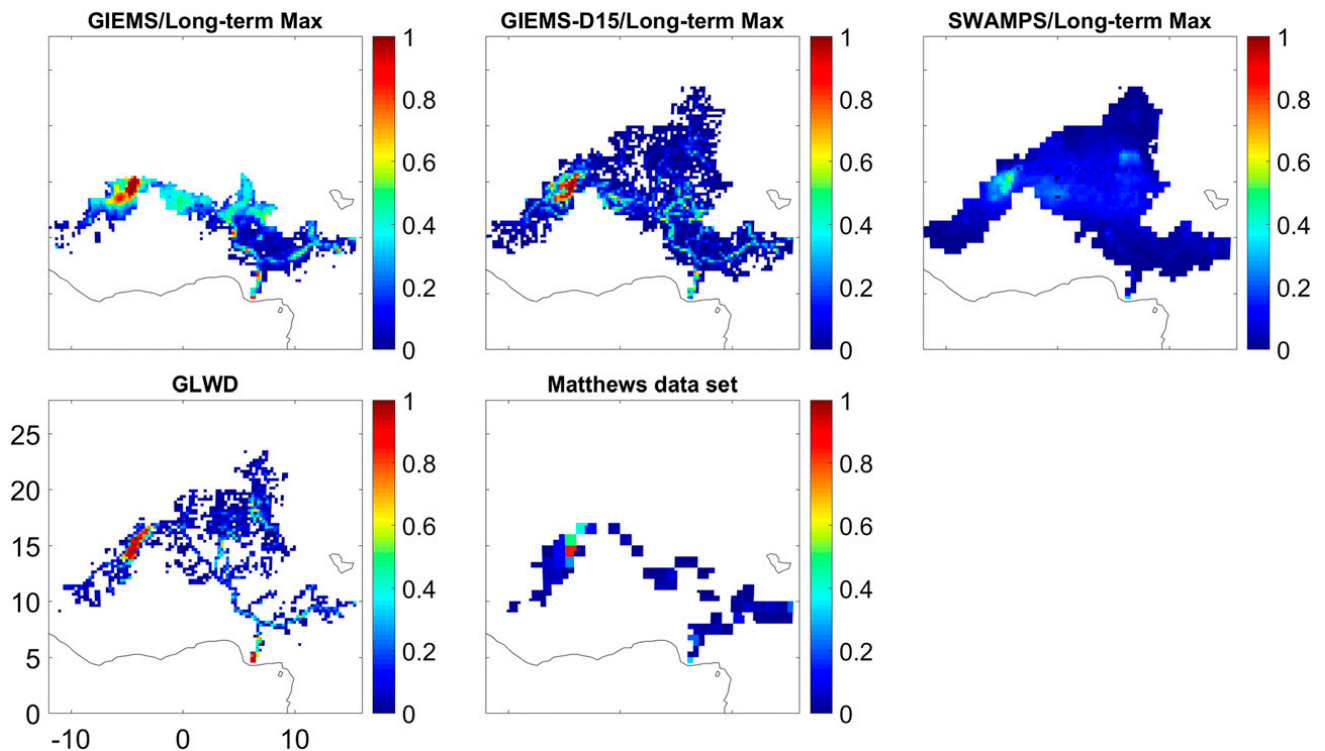


FIG. 8. Long-term max inundation maps averaged over 1993–2007 from GIEMS and SWAMPS, and max inundation maps from the different surface water datasets over the Niger basin. Shown are the GIEMS, GIEMS-D15, SWAMPS, GLWD, and Matthews and Fung (1987) datasets.



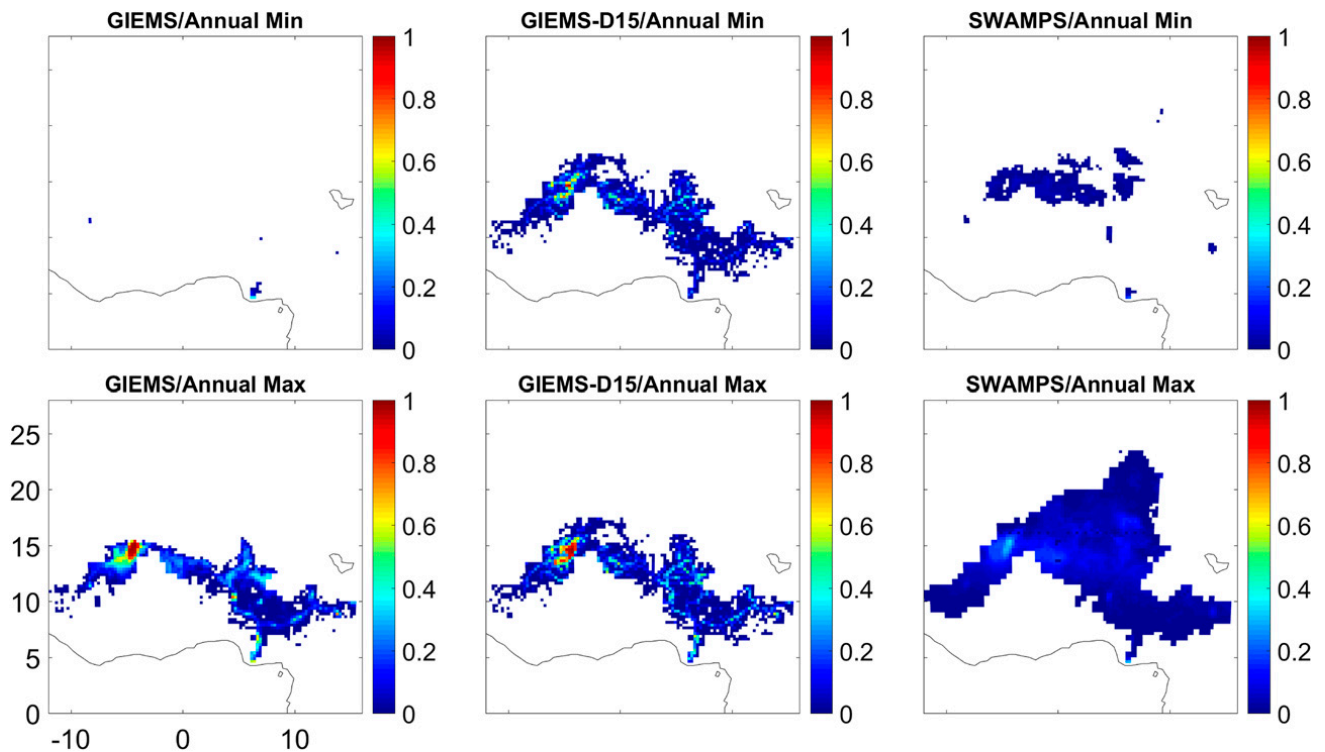


FIG. 9. Mean annual (top) min and (bottom) max of the inundation derived from different surface water datasets over the Niger basin for (left) GIEMS, (center) GIEMS-D15, and (right) SWAMPS.

analyzed GIEMS over the Niger region, for example, [Pedinotti et al. \(2012\)](#) or [Aires et al. \(2014\)](#). [Pedinotti et al. \(2012\)](#) evaluate the ability of the ISBA–Total Runoff Integrating Pathways (TRIP) continental hydrologic system to represent key processes (surface water, rivers and floodplain dynamics, and water storage) related to the hydrological cycle of the Niger basin.

To this end, GIEMS is used to evaluate the long-term simulations that showed that the flooding scheme leads to a nonnegligible increase of evaporation over large flooded areas, which in turns improved the Niger River discharge estimates at several locations. The objective of [Aires et al. \(2014\)](#) is to develop downscaling methodologies to obtain a long time record of inundation extent over

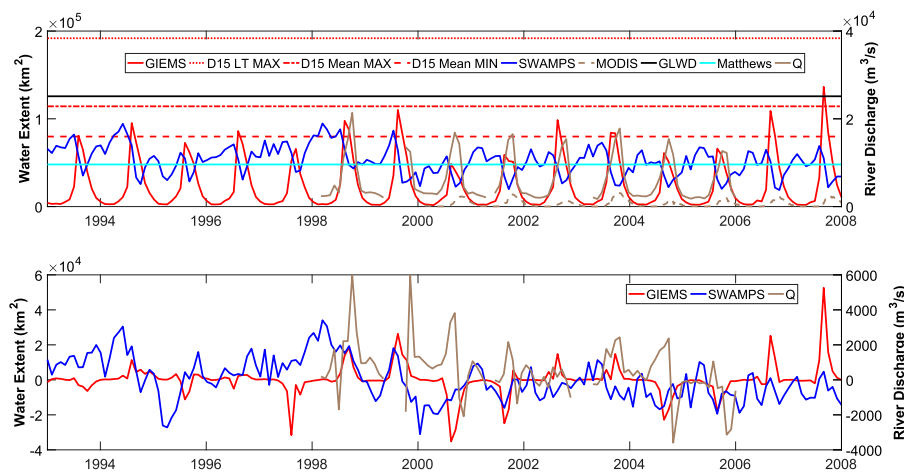


FIG. 10. (top) Time series and (bottom) anomaly of surface water extent derived from the different wetland datasets (1993–2007), along with the river discharge (1998–2005) over the whole Niger basin. MODIS surface water extent time series for the inner delta only (dashed brown) is available from 2000.

TABLE 3. Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge  $Q$  over the Niger basin for the period 1993–2007. Numbers in parentheses are calculated with 2-month lag between GIEMS/SWAMPS and  $Q$ .

	Time series correlation	Anomaly correlation
GIEMS/SWAMPS	–40%	20%
GIEMS/ $Q$	81% (51%)	–3% (44%)
SWAMPS/ $Q$	–60% (10%)	–2% (12%)

the inner Niger delta at high spatial resolution (500 m) based on the existing low-spatial-resolution results of the GIEMS dataset and observations from MODIS.

Time series of the surface water derived from MODIS visible images over the Niger basin for the period 2000–07 (Bergé-Nguyen and Crétaux 2015) were also compared to the behavior of GIEMS and SWAMPS over this region. From Fig. 10 (top) and Fig. 11, it is clear that GIEMS and MODIS surface water time series have similar seasonal dynamics over the common period (2000–07). However, GIEMS has a higher maximum value than MODIS, which could suggest an overestimation from GIEMS over this region. In addition, the interannual variability is not totally similar between GIEMS and MODIS. Similar passive microwave signatures can be observed over arid regions and over water: these two surface types have low emissivities with rather large emissivity polarization differences. As a consequence, reliable and accurate detection of surface water in arid and semiarid regions is not trivial. In GIEMS, NDVI information is used in the process to help solve these ambiguities. SWAMPS obviously encounters difficulties in this type of environment, with false detection of water in deserts and underestimation

of water surfaces in inundated deltas. This is in agreement with Schroeder et al. (2015). As a result, over the Niger delta, SWAMPS does not capture at all the water surface dynamics.

### 3) COMPARISONS OVER THE OB BASIN

The Ob basin in western Siberia is selected to represent the boreal environments. SWAMPS surface waters are again much less extended than the other estimates (see Figs. 12 and 13). SWAMPS surface water peaks generally in May, one month earlier than GIEMS. Time series correlation between GIEMS (SWAMPS) and the river discharge for the studied period is 91% (62%). When calculated with 1-month lag, time series correlation decreases for GIEMS to 80%, while it increases for SWAMPS to 91% (Table 4). The same conclusions can be found for the anomaly correlations between GIEMS, SWAMPS, and river discharge over the Ob basin (also Table 4). The lag between GIEMS and SWAMPS could be partly related to differences in the snow filtering performed monthly with GIEMS and daily with SWAMPS. GIEMS estimates are flagged too quickly by the NSIDC snow mask, so we are missing the end of the high-water-stage season (September–October). The river discharge from the Ob River and GIEMS have been compared in Papa et al. (2008b), and the use of the snow mask was well discussed. The snow flag in GIEMS is under analysis and will be refined for the next version of the dataset.

## 4. Discussion

As already observed in Fig. 1, the maximum surface water extent estimated by SWAMPS for the major basins is limited compared to the other estimates. The annual maximum SWAMPS surface extent (including

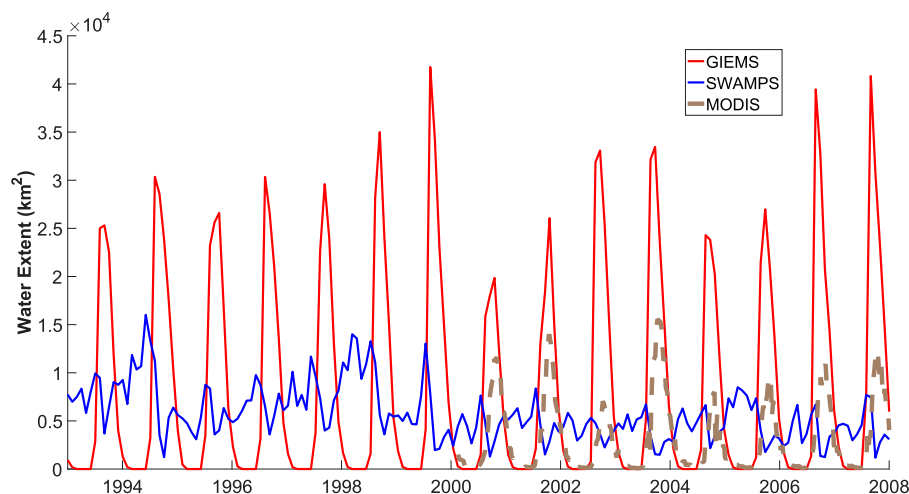


FIG. 11. Time series of surface water extent derived from GIEMS, SWAMPS, and MODIS for the inner Niger delta.

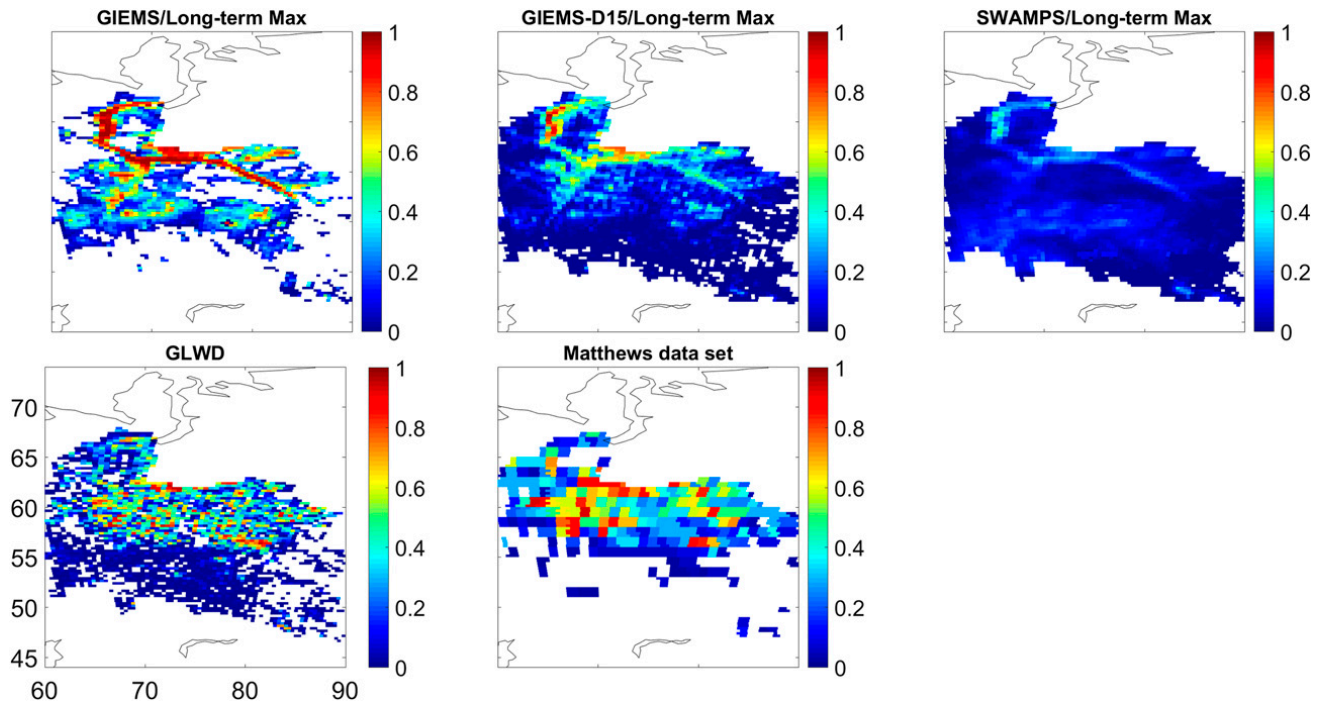


FIG. 12. Long-term max inundation maps averaged over 1993–2007 from GIEMS and SWAMPS, and max inundation maps from the different surface water datasets over the Ob basin. Shown are the GIEMS, GIEMS-D15, SWAMPS, GLWD, and [Matthews and Fung \(1987\)](#) datasets.

the coasts) and the GLWD surfaces are similar, and this feature is advertised in [Schroeder et al. \(2015\)](#). However, ~50% of the surface water with SWAMPS is located along the coasts. This clearly relates to contamination by the ocean in the observation fields of view of the passive microwave observations and to the lack of adequate filtering in the retrieval algorithm. The SWAMPS algorithm fits the global maximum water extent of GLWD, but as the coastal waters were

mistakenly included in the tuning, the inland water extent is therefore strongly underestimated. The range of seasonal variability of SWAMPS is also strongly reduced because coastal regions do not evolve in time.

The underestimation of the SWAMPS extent under dense vegetation is particularly significant, as observed in the Amazon basin. The well-established SAR estimate from [Hess et al. \(2003\)](#) is 243 000 km<sup>2</sup> at the high stage, very close to the GIEMS values, and very

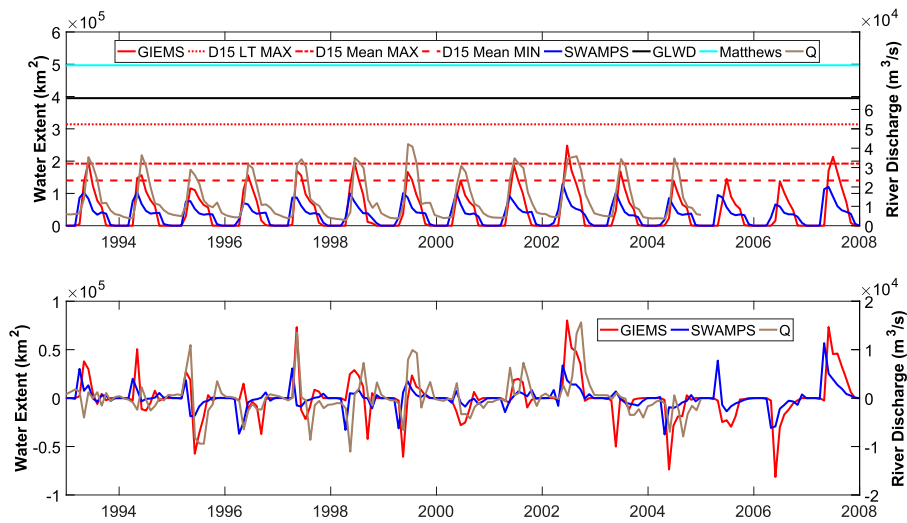


FIG. 13. As in [Fig. 7](#), but for the Ob basin. River discharge (brown) is available until 2004.

TABLE 4. Time series and anomaly correlations between GIEMS, SWAMPS, and river discharge  $Q$  over the Ob basin for the period 1993–2007. Numbers in parentheses are calculated with 1-month lag between GIEMS/SWAMPS and  $Q$ .

	Time series correlation	Anomaly correlation
GIEMS/SWAMPS	70%	38%
GIEMS/ $Q$	91% (80%)	49% (40%)
SWAMPS/ $Q$	62% (91%)	28% (59%)

different from the Phased Array type L-band Synthetic Aperture Radar (PALSAR) estimates (40 000 km<sup>2</sup>) provided by Schroeder et al. (2015) in their paper.

In the microwaves, the surface emissivities of water and desert surfaces are both rather low, with large polarization differences. As a consequence, there can be confusion between deserts and surface waters from passive microwave observations. This is typically what happens over deserts with SWAMPS, with anomalous detection of surface water over arid regions. In GIEMS, the systematic use of visible and near-infrared observations helps suppress these ambiguities.

Figure 14 (top) shows the time correlation between the two datasets and Fig. 14 (bottom) shows the time correlation between their anomalies, for the major 23 river basins in the world. The correlation is important for most basins, for the time series as well as for their anomalies. It

is even very high for some tropical basins (Orinoco and Mekong). This tends to confirm the seasonal variations of the surface water estimates, despite their different seasonal amplitudes. Note that correlations on time series or anomalies can be high despite large bias errors.

Over Asia (i.e., northeastern India and Bangladesh), GIEMS estimates large surface water extents that are related partly to rice paddies (see the MIRCA estimation of the rice paddies extent over the tropical region in Fig. 4). It is also suspected that GIEMS is very sensitive to saturated soil in this region, and as a consequence might overestimate the surface water extent in these regions.

### 5. Conclusions

Two global satellite-derived surface water datasets are compared on a monthly-mean basis from 1993 to 2007 (GIEMS and SWAMPS), along with two widely used static maps of the surface water. The 23 largest basins in the world have been studied, and three basins representative of different environments (the Amazon basin in the tropics, the Niger basin in a semiarid environment, and the Ob River in the boreal region) have been presented. Although they are based on similar observations, mainly passive and active microwaves, the satellite-derived datasets show large differences, globally and regionally, in terms of surface extents both at minimum and maximum inundation in the year. The

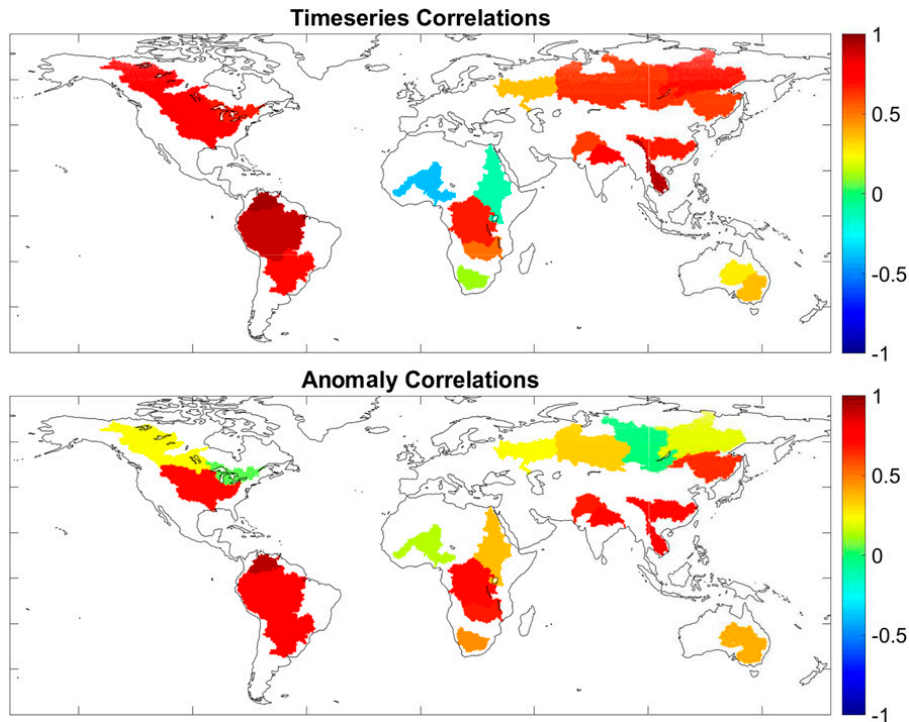


FIG. 14. Basinwide global (top) time series and (bottom) anomaly correlations for the 23 largest basins in the world between GIEMS and SWAMPS datasets.

global maximum inundation extent over 1993–2007 is  $\sim 6.2 \times 10^6 \text{ km}^2$  for GIEMS and  $\sim 10.3 \times 10^6 \text{ km}^2$  for SWAMPS, compared to  $\sim 13 \times 10^6 \text{ km}^2$  for GLWD and  $\sim 5.3 \times 10^6 \text{ km}^2$  for the Matthews and Fung (1987) wetland dataset. Approximately 50% of the SWAMPS inundated surfaces are located along the coast at the maximum annual inundation. This is clearly related to contamination by the ocean in the observation fields of view. Once this problem is filtered out, the long-term maximum surface water from SWAMPS is reduced to  $\sim 5 \times 10^6 \text{ km}^2$ . Globally and for the studied basins, the annual amplitude of the inundation extent is very limited in SWAMPS compared to GIEMS (47% lower). Despite their large difference in the seasonal amplitude, GIEMS and SWAMPS have similar temporal dynamics for most parts of the globe. Over the Amazon basin, GIEMS and SWAMPS show a very high temporal correlation (95%), but with SWAMPS maximum extent is half the size of that observed with GIEMS and with previous SAR estimates. Over the Niger basin, SWAMPS seasonal cycle is out of phase with both GIEMS and MODIS-derived estimates, as well as with river discharge data. This confirms the fact that SWAMPS fails to capture the seasonal dynamic of wetlands here. GIEMS and MODIS surface water time series agree in the seasonal variability, but GIEMS water extent is significantly larger than the MODIS one. In the Ob region, the different snow detection method could explain part of the difference in the seasonal cycle.

A clear advantage of the SWAMPS dataset today is its longer time period, up to 2013. The current version of the GIEMS algorithm requires a large quantity of ancillary satellite products to run, including outputs from the GEWEX International Satellite Cloud Climatology Project (Rossow and Schiffer 1999) that stopped in 2008. Efforts are underway to extend the GIEMS time series to current times, to provide the community with a long time record of carefully evaluated surface water extent all over the globe, using a reduced number of ancillary parameters for more robustness.

*Acknowledgments.* We thank our colleagues for providing ancillary datasets for this work. We thank two anonymous reviewers for their suggestions and comments. This study was financially supported by a PhD fellowship from the Vietnam International Education Development (911 project).

#### REFERENCES

- Adam, L., P. Döll, C. Prigent, and F. Papa, 2010: Global-scale analysis of satellite-derived time series of naturally inundated areas as a basis for floodplain modeling. *Adv. Geosci.*, **27**, 45–50, doi:10.5194/adgeo-27-45-2010.
- Aires, F., F. Papa, C. Prigent, J.-F. Crétaux, and M. Bergé-Nguyen, 2014: Characterization and space–time downscaling of the inundation extent over the inner Niger delta using GIEMS and MODIS data. *J. Hydrometeorol.*, **15**, 171–192, doi:10.1175/JHM-D-13-032.1.
- , L. Miolane, C. Prigent, E. Fluet-Chouinard, B. Lerner, and F. Aires, 2017: A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. *J. Hydrometeorol.*, doi:10.1175/JHM-D-16-0155.1, in press.
- Als Dorf, D. E., E. Rodriguez, and D. P. Lettenmaier, 2007: Measuring surface water from space. *Rev. Geophys.*, **45**, RG2002, doi:10.1029/2006RG000197.
- Bergé-Nguyen, M., and J.-F. Crétaux, 2015: Inundations in the inner Niger delta: Monitoring and analysis using MODIS and global precipitation datasets. *Remote Sens.*, **7**, 2127–2151, doi:10.3390/rs70202127.
- Bousquet, P., and Coauthors, 2006: Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, **443**, 439–443, doi:10.1038/nature05132.
- Bridgman, S. D., H. Cadillo-Quiroz, J. K. Keller, and Q. Zhuang, 2013: Methane emissions from wetlands: Biogeochemical, microbial, and modeling perspectives from local to global scales. *Global Change Biol.*, **19**, 1325–1346, doi:10.1111/gcb.12131.
- Brodzik, M., and R. Armstrong, 2013: Northern Hemisphere EASE-Grid 2.0 weekly snow cover and sea ice extent, version 4. National Snow and Ice Data Center, Boulder, CO, digital media. [Available online at [https://nsidc.org/data/docs/daac/nsidc0046\\_nh\\_ease\\_snow\\_seaice.gd.html](https://nsidc.org/data/docs/daac/nsidc0046_nh_ease_snow_seaice.gd.html).]
- Chang, A., J. Foster, and D. Hall, 1987: Nimbus-7 SMMR derived global snow cover parameters. *Ann. Glaciol.*, **9**, 39–44.
- Downing, J. A., and Coauthors, 2006: The global abundance and size distribution of lakes, ponds, and impoundments. *Limnol. Oceanogr.*, **51**, 2388–2397, doi:10.4319/lo.2006.51.5.2388.
- Feng, M., J. O. Sexton, S. Channan, and J. R. Townshend, 2016: A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic-spectral classification algorithm. *Int. J. Digital Earth*, **9**, 113–133, doi:10.1080/17538947.2015.1026420.
- Fluet-Chouinard, E., B. Lehner, L.-M. Rebelo, F. Papa, and S. K. Hamilton, 2015: Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sens. Environ.*, **158**, 348–361, doi:10.1016/j.rse.2014.10.015.
- Giddings, L., and B. J. Choudhury, 1989: Observation of hydrological features with Nimbus-7 37 GHz data, applied to South America. *Int. J. Remote Sens.*, **10**, 1673–1686, doi:10.1080/01431168908903998.
- Grody, N. C., and A. N. Basist, 1996: Global identification of snowcover using SSM/I measurements. *IEEE Trans. Geosci. Remote Sens.*, **34**, 237–249, doi:10.1109/36.481908.
- Hess, L. L., J. M. Melack, E. M. Novo, C. C. Barbosa, and M. Gastil, 2003: Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. *Remote Sens. Environ.*, **87**, 404–428, doi:10.1016/j.rse.2003.04.001.
- Lehner, B., and P. Döll, 2004: Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.*, **296**, 1–22, doi:10.1016/j.jhydrol.2004.03.028.
- , K. Verdin, and A. Jarvis, 2008: New global hydrography derived from spaceborne elevation data. *Eos, Trans. Amer. Geophys. Union*, **89**, 93–94, doi:10.1029/2008EO100001.

- Matthews, E., and I. Fung, 1987: Methane emission from natural wetlands: Global distribution, area, and environmental characteristics of sources. *Global Biogeochem. Cycles*, **1**, 61–86, doi:10.1029/GB001i001p00061.
- Melton, J. R., and Coauthors, 2013: Present state of global wetland extent and wetland methane modelling: Conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences*, **10**, 753–788, doi:10.5194/bg-10-753-2013.
- Mueller, N., and Coauthors, 2016: Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sens. Environ.*, **174**, 341–352, doi:10.1016/j.rse.2015.11.003.
- Nakaegawa, T., 2012: Comparison of water-related land cover types in six 1-km global land cover datasets. *J. Hydrometeorol.*, **13**, 649–664, doi:10.1175/JHM-D-10-05036.1.
- Papa, F., C. Prigent, F. Durand, and W. B. Rossow, 2006: Wetland dynamics using a suite of satellite observations: A case study of application and evaluation for the Indian Subcontinent. *Geophys. Res. Lett.*, **33**, L08401, doi:10.1029/2006GL025767.
- , —, and W. B. Rossow, 2007: Ob' River flood inundations from satellite observations: A relationship with winter snow parameters and river runoff. *J. Geophys. Res.*, **112**, D18103, doi:10.1029/2007JD008451.
- , A. Güntner, F. Frappart, C. Prigent, and W. B. Rossow, 2008a: Variations of surface water extent and water storage in large river basins: A comparison of different global data sources. *Geophys. Res. Lett.*, **35**, L11401, doi:10.1029/2008GL033857.
- , C. Prigent, and W. B. Rossow, 2008b: Monitoring flood and discharge variations in the large Siberian rivers from a multi-satellite technique. *Surv. Geophys.*, **29**, 297–317, doi:10.1007/s10712-008-9036-0.
- , —, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews, 2010: Interannual variability of surface water extent at the global scale, 1993–2004. *J. Geophys. Res.*, **115**, D12111, doi:10.1029/2009JD012674.
- Pedinotti, V., A. Boone, B. Decharme, J. F. Crétaux, N. Mognard, G. Panthou, F. Papa, and B. A. Tanimoun, 2012: Evaluation of the ISBA-TRIP continental hydrologic system over the Niger basin using in situ and satellite derived datasets. *Hydrol. Earth Syst. Sci.*, **16**, 1745–1773, doi:10.5194/hess-16-1745-2012.
- Pison, I., B. Ringeval, P. Bousquet, C. Prigent, and F. Papa, 2013: Stable atmospheric methane in the 2000s: Key-role of emissions from natural wetlands. *Atmos. Chem. Phys.*, **13**, 11 609–11 623, doi:10.5194/acp-13-11609-2013.
- Portmann, F. T., S. Siebert, and P. Döll, 2010: MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochem. Cycles*, **24**, GB1011, doi:10.1029/2008GB003435.
- Prigent, C., E. Matthews, F. Aires, and W. B. Rossow, 2001: Remote sensing of global wetland dynamics with multiple satellite data sets. *Geophys. Res. Lett.*, **28**, 4631–4634, doi:10.1029/2001GL013263.
- , F. Aires, and W. B. Rossow, 2006: Land surface microwave emissivities over the globe for a decade. *Bull. Amer. Meteor. Soc.*, **87**, 1573–1584, doi:10.1175/BAMS-87-11-1573.
- , F. Papa, F. Aires, W. B. Rossow, and E. Matthews, 2007: Global inundation dynamics inferred from multiple satellite observations, 1993–2000. *J. Geophys. Res.*, **112**, D12107, doi:10.1029/2006JD007847.
- , —, —, C. Jimenez, W. B. Rossow, and E. Matthews, 2012: Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophys. Res. Lett.*, **39**, L08403, doi:10.1029/2012GL051276.
- , D. P. Lettenmaier, F. Aires, and F. Papa, 2016: Toward a high-resolution monitoring of continental surface water extent and dynamics, at global scale: From GIEMS (Global Inundation Extent from Multi-Satellites) to SWOT (Surface Water Ocean Topography). *Surv. Geophys.*, **37**, 339–355, doi:10.1007/s10712-015-9339-x.
- Reichhardt, T., 1995: Academy under fire on wetlands definition. *Nature*, **375**, 171, doi:10.1038/375171a0.
- Ringeval, B., N. de Noblet-Ducoudre, P. Ciais, P. Bousquet, C. Prigent, F. Papa, and W. B. Rossow, 2010: An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales. *Global Biogeochem. Cycles*, **24**, GB2003, doi:10.1029/2008GB003354.
- Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261–2287, doi:10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2.
- Schroeder, R., K. C. McDonald, B. D. Chapman, K. Jensen, E. Podest, Z. D. Tessler, T. J. Bohn, and R. Zimmermann, 2015: Development and evaluation of a multi-year fractional surface water data set derived from active/passive microwave remote sensing data. *Remote Sens.*, **7**, 16 688–16 732, doi:10.3390/rs71215843.
- Sippel, S. J., S. K. Hamilton, J. M. Melack, and B. J. Choudhury, 1994: Determination of inundation area in the Amazon River floodplain using the SMMR 37 GHz polarization difference. *Remote Sens. Environ.*, **48**, 70–76, doi:10.1016/0034-4257(94)90115-5.
- Verpoorter, C., T. Kutser, D. A. Seekell, and L. J. Tranvik, 2014: A global inventory of lakes based on high-resolution satellite imagery. *Geophys. Res. Lett.*, **41**, 6396–6402, doi:10.1002/2014GL060641.
- Vorosmarty, C. J., and Coauthors, 2010: Global threats to human water security and river biodiversity. *Nature*, **467**, 555–561, doi:10.1038/nature09440.
- Wania, R., and Coauthors, 2013: Present state of global wetland extent and wetland methane modelling: Methodology of a model inter-comparison project (WETCHIMP). *Geosci. Model Dev.*, **6**, 617–641, doi:10.5194/gmd-6-617-2013.
- Yamazaki, D., M. A. Trigg, and D. Ikeshima, 2015: Development of a global ~90 m water body map using multi-temporal Landsat images. *Remote Sens. Environ.*, **171**, 337–351, doi:10.1016/j.rse.2015.10.014.



# Bibliography

Acharya, T. D., D. H. Lee, I. T. Yang, and J. K. Lee

2016. Identification of Water Bodies in a Landsat 8 OLI Image Using a J48 Decision Tree. *Sensors*, 16(7):1–16.

Adam, L., P. Döll, C. Prigent, and F. Papa

2010. Global-scale analysis of satellite-derived time series of naturally inundated areas as a basis for floodplain modeling. *Advances in Geosciences*, 27:45–50.

Aires, F., L. Miolane, C. Prigent, B. Pham, E. Fluet-Chouinard, B. Lehner, and F. Papa

2017. A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. *Journal of Hydrometeorology*, 18(5):1305–1325.

Aires, F., F. Papa, C. Prigent, J.-F. Crétaux, and M. Bergé-Nguyen

2014. Characterization and Space-Time Downscaling of the Inundation Extent over the Inner Niger Delta Using GIEMS and MODIS Data. *Journal of Hydrometeorology*, 15(1):171–192.

Alsdorf, D. E., E. Rodríguez, and D. P. Lettenmaier

2007. Measuring surface water from space. *Reviews of Geophysics*, 45(2). RG2002.

Amitrano, D., G. D. Martino, A. Iodice, F. Mitidieri, M. N. Papa, D. Riccio, and G. Ruello

2014. Sentinel-1 for Monitoring Reservoirs: A Performance Analysis. *Remote Sensing*, 6(11):10676–10693.



- Bartsch, A., C. Pathe, W. Wagner, and K. Scipal  
2008. Detection of permanent open water surfaces in central Siberia with ENVISAT ASAR wide swath data with special emphasis on the estimation of methane fluxes from tundra wetlands. *Hydrology Research*, 39(2):89–100.
- Beget, M. E., V. A. Bettachini, C. M. Di Bella, and F. Baret  
2013. SAILHFlood: A radiative transfer model for flooded vegetation. *Ecological Modelling*, 257:25–35.
- Bergé-Nguyen, M. and J.-F. Crétaux  
2015. Inundations in the Inner Niger Delta: Monitoring and Analysis Using MODIS and Global Precipitation Datasets. *Remote Sensing*, 7(2):2127.
- Birkett, C. M.  
1998. Contribution of the topex nasa radar altimeter to the global monitoring of large rivers and wetlands. *Water Resources Research*, 34(5):1223–1239.
- Birkett, C. M., L. A. K. Mertes, T. Dunne, M. H. Costa, and M. J. Jasinski  
2002. Surface water dynamics in the amazon basin: Application of satellite radar altimetry. *Journal of Geophysical Research: Atmospheres*, 107(D20):LBA 26–1–LBA 26–21. 8059.
- Birkett, C. M. B.  
1995. The contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes. *Journal of Geophysical Research*, 100204(15):179–25.
- Bousquet, P., P. Ciais, J. B. Miller, E. J. Dlugokencky, D. A. Hauglustaine, C. Prigent, G. R. V. der Werf, P. Peylin, E. G. Brunke, C. Carouge, R. L. Langenfelds, J. Lathière, F. Papa, M. Ramonet, M. Schmidt, L. P. Steele, S. C. Tyler, and J. White  
2006. Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, 443:439–443.
- Brakenridge, G. and S. V. Nghiem  
2017. Merged AMSR-2 and GPM Passive Microwave Radiometry for Measuring River Discharge and Runoff. *Applied Earth Observations and Remote Sensing*, Pp. 1–19.

- Brakenridge, G. R., S. V. Nghiem, E. Anderson, and S. Chien  
2005. Space-based measurement of river runoff. *Eos, Transactions American Geophysical Union*, 86(19):185–188.
- Bridgman, S. D., H. Cadillo-Quiroz, J. K. Keller, and Q. Zhuang  
2013. Methane emissions from wetlands: biogeochemical, microbial, and modeling perspectives from local to global scales. *Global Change Biology*, 19(5):1325–1346.
- Brisco, B., N. Short, J. van der Sanden, R. Landry, and D. Raymond  
2009. A semi-automated tool for surface water mapping with RADARSAT-1. *Canadian Journal of Remote Sensing*, 35(4):336–344.
- Brodzik, M. J. and R. L. Armstrong  
2013. Northern Hemisphere EASE-Grid 2.0 weekly snow cover and sea ice extent version 4. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center.
- Bryant, R. G. and M. P. Rainey  
2002. Investigation of flood inundation on playas within the Zone of Chotts, using a time-series of AVHRR. *Remote Sensing of Environment*, 82(2 and 3):360–375.
- Bullock, A. and M. Acreman  
2003. The role of wetlands in the hydrological cycle. *Hydrology and Earth System Sciences Discussions*, 7(3):358–389.
- CCMEO, C. C. f. M. . E. O.  
2013. *Fundamentals of Remote Sensing*. Canada Centre for Mapping & Earth Observation.
- Chang, A. T. C., J. L. Foster, and D. K. Hall  
1987. NIMBUS-7 SMMR DERIVED GLOBAL SNOW COVER PARAMETERS. *International Glaciological Society*, 9:39–44.
- Cochrane, T. A., M. E. Arias, and T. Piman  
2014. Historical impact of water infrastructure on water levels of the Mekong River and the Tonle Sap system. *Hydrology and Earth System Sciences*, 18(11):4529–4541.

- Coe, M. T. and C. M. Birkett  
2004. Calculation of river discharge and prediction of lake height from satellite radar altimetry: Example for the lake chad basin. *Water Resources Research*, 40(10). W10205.
- Coe, M. T., M. H. Costa, A. Botta, and C. Birkett  
2002. Long-term simulations of discharge and floods in the amazon basin. *Journal of Geophysical Research: Atmospheres*, 107(D20):LBA 11-1-LBA 11-17.
- Crétaux, J.-F., R. Abarca-del Río, M. Bergé-Nguyen, A. Arsen, V. Drolon, G. Clos, and P. Maisongrande  
2016. Lake volume monitoring from space. *Surveys in Geophysics*, 37(2):269-305.
- Crétaux, J.-F., M. Berge-Nguyen, M. Leblanc, R. Abarca Del Rio, F. Delclaux, N. Mognard, C. Lion, R.-K. Pandey, S. Tweed, S. Calmant, and P. Maisongrande  
2011. Flood mapping inferred from remote sensing data. *International Water Technology Journal*, 1(1):48-62.
- Crétaux, J.-F., S. Biancamaria, A. Arsen, M. Bérgé-Nguyen, and M. Becker  
2015. Global surveys of reservoirs and lakes from satellites and regional application to the syrdarya river basin. *Environmental Research Letters*, 10(1):015002.
- Crétaux, J. F. and C. Birkett  
2006. Lake studies from satellite radar altimetry. *Comptes Rendus - Geoscience*, 338(14-15):1098-1112.
- Domenikiotis, C., A. Loukas, and N. R. Dalezios  
2003. The use of noaa/avhrr satellite data for monitoring and assessment of forest fires and floods. *Natural Hazards and Earth System Sciences*, 3(1/2):115-128.
- Downing, J. A., Y. T. Prairie, J. J. Cole, C. M. Duarte, L. J. Tranvik, R. G. Striegl, W. H. McDowell, P. Kortelainen, N. F. Caraco, J. M. Melack, and J. J. Middelburg  
2006. The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*, 51(5):2388-2397.

- Du, Y., Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li  
2016. Water Bodies ' Mapping from Sentinel-2 Imagery with Modified Normalized Difference Water Index at 10-m Spatial Resolution Produced by Sharpening the SWIR Band. *Remote Sensing*, 8(4).
- Du, Z.  
2012. Estimating surface water area changes using time-series Landsat data in the Qingjiang River Basin , China. *Journal of Applied Remote Sensing*.
- Du, Z., W. Li, D. Zhou, L. Tian, and F. Ling  
2014. Analysis of Landsat-8 OLI imagery for land surface water mapping. *Remote Sensing Letter*, 5(7):672–681.
- Elachi, C. and J. van Zyl  
2006. *Introduction to the Physics and Techniques of Remote Sensing*. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- ESA  
2015. Sentinel-1 Data Hub.
- Famiglietti, J. S., M. Lo, S. L. Ho, J. Bethune, K. J. Anderson, T. H. Syed, S. C. Swenson, C. R. de Linage, and M. Rodell  
2011. Satellites measure recent rates of groundwater depletion in california's central valley. *Geophysical Research Letters*, 38(3). L03403.
- Feng, M., J. O. Sexton, S. Channan, and J. R. Townshend  
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(2):113–133.
- Feng, W., M. Zhong, J.-M. Lemoine, R. Biancale, H.-T. Hsu, and J. Xia  
2013. Evaluation of groundwater depletion in north china using the gravity recovery and climate experiment (grace) data and ground-based measurements. *Water Resources Research*, 49(4):2110–2118.
- Feyisa, G. L., H. Meilby, R. Fensholt, and S. R. Proud  
2014. Automated Water Extraction Index : A New Technique for Surface Water Mapping Using Landsat Imagery. *Remote Sensing of Environment*, 140(July):23–35.

- Fjrtoft, R., J. M. Gaudin, N. Pourthié, J. C. Lalaurie, A. Mallet, J. F. Nouvel, J. Martinot-Lagarde, H. Oriot, P. Borderies, C. Ruiz, and S. Daniel  
2014. Karin on swot: Characteristics of near-nadir ka-band interferometric sar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(4):2172–2185.
- Fluet-Chouinard, E., B. Lehner, L. M. Rebelo, F. Papa, and S. K. Hamilton  
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.
- Frappart, F., D. Blumstein, A. Cazenave, G. Ramillien, F. Birol, R. Morrow, and F. Remy  
2017. Satellite Altimetry: Principles and Applications in Earth Sciences. *Wiley Encyclopedia of Electrical and Electronics Engineering*, (3):1–25.
- Frappart, F., S. Calmant, M. Cauhopé, F. Seyler, and A. Cazenave  
2006a. Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin. *Remote Sensing of Environment*, 100(2):252–264.
- Frappart, F., K. Do Minh, J. L'Hermitte, A. Cazenave, G. Ramillien, T. Le Toan, and N. Mognard-Campbell  
2006b. Water volume change in the lower Mekong from satellite altimetry and imagery data. *Geophysical Journal International*, 167(2):570–584.
- Frappart, F., F. Papa, J. S. Famiglietti, C. Prigent, and W. B. Rossow  
2008. Interannual variations of river water storage from a multiple satellite approach : A case study for the Rio Negro River basin. *Journal of Geophysical Research*, 113:1–12.
- Frappart, F., F. Papa, A. Güntner, S. Werth, G. Ramillien, C. Prigent, W. B. Rossow, and M.-P. Bonnet  
2010. Interannual variations of the terrestrial water storage in the lower ob basin from a multisatellite approach. *Hydrology and Earth System Sciences*, 14(12):2443–2453.
- Frappart, F., F. Papa, A. Güntner, S. Werth, J. Santos, J. Tomasella, F. Seyler, C. Prigent, W. B. Rossow, S. Calmant, and M.-p. Bonnet  
2011. Remote Sensing of Environment Satellite-based estimates of groundwater

- storage variations in large drainage basins with extensive floodplains. *Remote Sensing of Environment*, 115(6):1588–1594.
- Frappart, F., F. Seyler, J.-M. Martinez, J. G. León, and A. Cazenave  
2005. Floodplain water storage in the negro river basin estimated from microwave remote sensing of inundation area and water levels. *Remote Sensing of Environment*, 99(4):387 – 399.
- Fu, L.-L. and A. Cazenave  
2001. *Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications*. Academic Press, San Diego.
- Giddings, L. and B. Choudhury  
1989. Observation of hydrological features with Nimbus-7 37 GHz data, applied to South America. *International Journal of Remote Sensing*, 10:1673–1686.
- Goteti, G., J. S. Famiglietti, and K. Asante  
2008. A catchment-based hydrologic and routing modeling system with explicit river channels. *Journal of Geophysical Research: Atmospheres*, 113(D14). D14116.
- Grody, N. C. and A. N. Basist  
1996. Global identification of snowcover using {SSM}/ {I} measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 34(1):237–249.
- Gstaiger, V., J. Huth, S. Gebhardt, T. Wehrmann, and C. Kuenzer  
2012. Multi-sensoral and automated derivation of inundated areas using {TerraSAR}- {X} and {ENVISAT} {ASAR} data. *International Journal of Remote Sensing*, 33(22):7291–7304.
- Hallberg, G. R., B. E. Hoyer, and A. Rango  
1973. Application of ERTS-1 Imagery to Flood Inundation Mapping. *NASA Special Publication. No. 327, Symposium on significant results obtained from the Earth Resources Satellite-1*, 1(Technical presentations, section A):745–753.
- Henry, J.-B., P. Chastanet, K. Fellah, and Y.-L. Desnos  
2006. Envisat multi-polarized {ASAR} data for flood mapping. *International Journal of Remote Sensing*, 27(10):1921–1929.

- Hess, L. L., J. M. Melack, E. M. L. M. Novo, C. C. F. Barbosa, and M. Gastil  
2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. *Remote Sensing of Environment*, 87(4):404–428.
- Hess, L. L., J. M. Melack, and D. S. Simonett  
1990. Radar detection of flooding beneath the forest canopy: a review. *International Journal of Remote Sensing*, 11(7):1313–1325.
- Houweling, S., T. Kaminski, F. Dentener, J. Lelieveld, and M. Heimann  
1999. Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *Journal of Geophysical Research: Atmospheres*, 104(D21):26137–26160.
- Islam, M. M. and K. Sado  
2000. Flood hazard assessment in bangladesh using noaa avhrr data with geographical information system. *Hydrological Processes*, 14(3):605–620.
- Jain, S. K., A. K. Saraf, and A. Goswami  
2006. Flood inundation mapping using NOAA AVHRR data. *Water Resources Management*, Pp. 949–959.
- Jensen, J. R.  
1995. *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd edition. Upper Saddle River, NJ, USA: Prentice Hall PTR.
- Ji, L., X. Geng, K. Sun, Y. Zhao, and P. Gong  
2015. Target Detection Method for Water Mapping Using Landsat 8 OLI/TIRS Imagery. *Water*, Pp. 794–817.
- Kasischke, E. S. and L. L. Bourgeau-Chavez  
1997. Monitoring {South} {Florida} {Wetlands} {Using} {ERS}-1 {SAR} {Imagery}. *Photogrammetric Engineering & Remote Sensing*, 63(3):281–291.
- Kuenzer, C., H. Guo, J. Huth, P. Leinenkugel, X. Li, and S. Dech  
2013. Flood Mapping and Flood Dynamics of the Mekong Delta: ENVISAT-ASAR-WSM Based Time Series Analyses. *Remote Sensing*, 5(2):687–715.

- Leblanc, M. J., P. Tregoning, G. Ramillien, S. O. Tweed, and A. Fakes  
2009. Basin-scale, integrated observations of the early 21st century multiyear drought in southeast australia. *Water Resources Research*, 45(4). W04408.
- Lehner, B. and P. Döll  
2004. Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296(1-4):1–22.
- Lehner, B., K. Verdin, and A. Jarvis  
2006. Technical Documentation Version 1.0.
- Lehner, B., K. Verdin, and A. Jarvis  
2008. New global hydrography derived from spaceborne elevation data. *Eos*, 89(10):93–94.
- Leinenkugel, P., C. Kuenzer, and S. Dech  
2013. Comparison and enhancement of modis cloud mask products for south-east asia. *International Journal of Remote Sensing*, 34(8):2730–2748.
- Levizzani, V., P. Bauer, and F. J. Turk  
2007. *Measuring Precipitation From Space*. Springer, Dordrecht.
- Liu, C.  
2016. *Analysis of Sentinel-1 SAR data for mapping standing water in the Twente region*. Master thesis, Twente.
- Liu, K.-t., K.-h. Tseng, C. K. Shum, C.-y. Liu, C.-y. Kuo, G. Liu, Y. Jia, and K. Shang  
2016a. Assessment of the Impact of Reservoirs in the Upper Mekong River Using Satellite Radar Altimetry and Remote Sensing Imageries in Southeast. *Remote Sensing*, 8.
- Liu, Y., K. S. Chen, P. Xu, and Z. L. Li  
2016b. Modeling and Characteristics of Microwave Backscattering From Rice Canopy Over Growth Stages. *IEEE Transactions on Geoscience and Remote Sensing*, 54(11):6757–6770.
- Martinez, J. M., J. L. Guyot, and G. Cochonneau  
2008. Monitoring of surface water quality in large rivers with satellite imagery



- Application to the Amazon basin. *13th IWRA World Water Congress 2008, 1-4 September, Montpellier, Fran*, (September 2008).
- Martinis, S., C. Kuenzer, A. Wendleder, J. Huth, A. Twele, A. Roth, and S. Dech  
2015. Comparing four operational SAR-based water and flood detection approaches. *International Journal of Remote Sensing*, 36(13):3519–3543.
- Matthews, E. and I. Fung  
1987. Methane emission from natural wetlands: Global distribution, area, and environmental characteristics of sources. *Global Biogeochemical Cycles*, 1(1):61–86.
- McFeeters, S. K.  
1996. The use of the Normalized Difference Water Index NDWI in the delineation of open water features. *International Journal of Remote Sensing*, 17(7):1425–1432.
- Melton, J. R., R. Wania, E. L. Hodson, B. Poulter, B. Ringeval, R. Spahni, T. Bohn, C. A. Avis, D. J. Beerling, G. Chen, A. V. Eliseev, S. N. Denisov, P. O. Hopcroft, D. P. Lettenmaier, W. J. Riley, J. S. Singarayer, Z. M. Subin, H. Tian, S. Zürcher, V. Brovkin, P. M. van Bodegom, T. Kleinen, Z. C. Yu, and J. O. Kaplan  
2013. Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences*, 10(2):753–788.
- Morrison, R. B. and M. E. Cooley  
1973. Assessment of food damage in Arizona by means of ERTS-1 imagery. *Proceedings, Symposium on Significant Results Obtained from the Earth Resources Satellite-1*, 1:755–760.
- MRC  
2011. Planning atlas of the lower mekong river basin, basic development plan programme for sustainable development. *Mekong River Commission, Cambodia, Lao PDR, Thailand, Vietnam*.

- Mueller, N., A. Lewis, D. Roberts, S. Ring, R. Melrose, J. Sixsmith, L. Lymburner, A. McIntyre, P. Tan, S. Curnow, and A. Ip  
2016. Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174:341–352.
- Nakaegawa, T.  
2011. Comparison of Water-Related Land Cover Types in Six 1-km Global Land Cover Datasets. *Journal of Hydrometeorology*, 13(2):649–664.
- Nguyen, L. and T. Bui  
2001. Flood Monitoring of Mekong River Delta, Vietnam using ERS SAR Data. *Proc. 22nd Asian Conference on Remote Sensing, November 5-9 2001, Singapore*.
- Nyborg, L. and I. Sandholt  
2001. NOAA-AVHRR based Flood Monitoring. *Geoscience and Remote Sensing Symposium*, 00(C):1696–1698.
- Owe, M., K. Brubaker, J. Ritchie, and R. Albert  
2001. *Remote Sensing and Hydrology, 2000*. IAHS, Wallingford, OX, UK.
- Papa, F., F. Frappart, A. Guntner, C. Prigent, F. Aires, A. C. V. Getirana, and R. Maurer  
2013. Surface freshwater storage and variability in the amazon basin from multi-satellite observations, 1993–2007. *Journal of Geophysical Research: Atmospheres*, 118(21):11,951–11,965. 2013JD020500.
- Papa, F., F. Frappart, Y. Malbeteau, M. Shamsudduha, V. Vuruputur, M. Sekhar, G. Ramillien, C. Prigent, F. Aires, R. K. Pandey, S. Bala, and S. Calmant  
2015. Satellite-derived surface and sub-surface water storage in the Ganges-Brahmaputra River Basin. *Journal of Hydrology: Regional Studies*, 4:15–35.
- Papa, F., A. Guntner, F. Frappart, C. Prigent, and W. B. Rossow  
2008a. Variations of surface water extent and water storage in large river basins: A comparison of different global data sources. *Geophysical Research Letters*, 35(11):L11401.
- Papa, F., C. Prigent, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews  
2010. Interannual variability of surface water extent at the global scale, 1993–2004. *Journal of Geophysical Research: Atmospheres*, 115(D12):D12111.

- Papa, F., C. Prigent, F. Durand, and W. B. Rossow  
2006. Wetland dynamics using a suite of satellite observations: A case study of application and evaluation for the Indian Subcontinent. *Geophysical Research Letters*, 33(8):L08401.
- Papa, F., C. Prigent, and W. B. Rossow  
2007. Ob' River flood inundations from satellite observations: A relationship with winter snow parameters and river runoff. *Journal of Geophysical Research: Atmospheres*, 112(D18):D18103.
- Papa, F., C. Prigent, and W. B. Rossow  
2008b. Monitoring flood and discharge variations in the large siberian rivers from a multi-satellite technique. *Surveys in Geophysics*, 29(4-5):297–317.
- Pedinotti, V., A. Boone, B. Decharme, J. F. Crétaux, N. Mognard, G. Panthou, F. Papa, and B. A. Tanimoun  
2012. Evaluation of the ISBA-TRIP continental hydrologic system over the Niger basin using in situ and satellite derived datasets. *Hydrology and Earth System Sciences*, 16(6):1745–1773.
- Pekel, J.-F., A. Cottam, N. Gorelick, and A. S. Belward  
2016. High-resolution mapping of global surface water and its long-term changes. *Nature*, Pp. 1–19.
- Peng, D., L. Xiong, S. Guo, and N. Shu  
2009. Study of dongting lake area variation and its influence on water level using modis data / etude de la variation de la surface du lac dongting et de son influence sur le niveau d'eau , grâce à des données modis. *Hydrological Sciences Journal*, 50(1).
- Pierdicca, N., L. Pulvirenti, M. Chini, L. Guerriero, and L. Candela  
2013. Observing floods from space: Experience gained from COSMO-SkyMed observations. *Acta Astronautica*, 84:122–133.
- Pietroniro, A. and T. D. Prowse  
2002. Applications of remote sensing in hydrology. *Hydrological Processes*, 16(8):1537–1541.

- Pison, I., B. Ringeval, P. Bousquet, C. Prigent, and F. Papa  
2013. Stable atmospheric methane in the 2000s: key-role of emissions from natural wetlands. *Atmospheric Chemistry and Physics*, 13(23):11609–11623.
- Pope, K. O., E. Rejmankova, J. F. Paris, and R. Woodruff  
1997. Detecting seasonal flooding cycles in marshes of the Yucatan Peninsula with SAR-C polarimetric radar imagery. *Remote Sensing of Environment*, 59(2):157–166.
- Portmann, F. T., S. Siebert, and P. Döll  
2010. Mirca2000—global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, 24(1). GB1011.
- Prigent, C., F. Aires, C. Jimenez, F. Papa, and J. Roger  
2015. Multiangle backscattering observations of continental surfaces in Ku-band (13 GHz) from satellites: Understanding the signals, particularly in arid regions. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3):1364–1373.
- Prigent, C., F. Aires, and W. B. Rossow  
2006. Land Surface Microwave Emissivities over the Globe for a Decade. *Bulletin of the American Meteorological Society*, 87(11):1573–1584.
- Prigent, C., D. P. Lettenmaier, F. Aires, and F. Papa  
2016. Toward a High-Resolution Monitoring of Continental Surface Water Extent and Dynamics, at Global Scale: from GIEMS (Global Inundation Extent from Multi-Satellites) to SWOT (Surface Water Ocean Topography). *Surveys in Geophysics*, 37(2):339–355.
- Prigent, C., E. Matthews, F. Aires, and W. B. Rossow  
2001. Remote sensing of global wetland dynamics with multiple satellite data sets. *Geophysical Research Letters*, 28(24):4631–4634.
- Prigent, C., F. Papa, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews  
2012. Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophysical Research Letters*, 39(8):L08403.

- Prigent, C., F. Papa, F. Aires, W. B. Rossow, and E. Matthews  
2007. Global inundation dynamics inferred from multiple satellite observations, 1993-2000. *Journal of Geophysical Research Atmospheres*, 112(12):1–13.
- Rango, A. and V. V. Salomonson  
1974. Regional flood mapping from space. *Water Resources Research*, 10(3):473–484.
- Reichhardt, T.  
1995. Academy under fire on 'wetlands' definition. *Nature*, 375(6528):171–171.
- Reschke, J., A. Bartsch, S. Schlaffer, and D. Schepaschenko  
2012. Capability of c-band sar for operational wetland monitoring at high latitudes. *Remote Sensing*, 4(10):2923–2943.
- Ringeval, B., N. de Noblet-Ducoudre, P. Ciais, P. Bousquet, C. Prigent, F. Papa, and W. B. Rossow  
2010. An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales. *Global Biogeochemical Cycles*, 24(2). GB2003.
- Robinove, C. J.  
1978. Interpretation of a Landsat image of an unusual flood phenomenon in Australia. *Remote Sensing of Environment*, 7(3):219–225.
- Rokni, K., A. Ahmad, A. Selamat, and S. Hazini  
2014. Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sensing*, 6(5):4173–4189.
- Rossow, W. B. and R. A. Schiffer  
1999. Advances in Understanding Clouds from ISCCP. *Bulletin of the American Meteorological Society*, 80(11):2261–2287.
- Rouse Jr., J., R. Haas, J. Schell, and D. Deering  
1974. Monitoring Vegetation Systems in the Great Plains with Ert's. *NASA Special Publication*, 351:309.

- Roux, E., J. Santos, A. Cesar, V. Getirana, M.-p. Bonnet, S. Calmant, J.-m. Martinez, E. Roux, J. Santos, A. Cesar, V. Getirana, M.-p. Bonnet, S. Calmant, J.-m. Martinez, and F. Seyler  
2010. Producing time series of river water height by means of satellite radar altimetry - a comparative study. *Hydrological Sciences Journal*, 55(1):104–120.
- Sakamoto, T., N. Van Nguyen, A. Kotera, H. Ohno, N. Ishitsuka, and M. Yokozawa  
2007. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sensing of Environment*, 109(3):295–313.
- Sakumura, C., S. Bettadpur, and S. Bruinsma  
2014. Ensemble prediction and intercomparison analysis of grace time-variable gravity field models. *Geophysical Research Letters*, 41(5):1389–1397.
- Santoro, M., U. Wegmuller, C. Lamarche, S. Bontemps, P. Defourny, and O. Arino  
2015a. Strengths and weaknesses of multi-year Envisat ASAR backscatter measurements to map permanent open water bodies at global scale. *Remote Sensing of Environment*, 171:185–201.
- Santoro, M., U. Wegmuller, A. Wiesmann, C. Lamarche, S. Bontemps, P. Defourny, and O. Arino  
2015b. Assessing Envisat ASAR and Sentinel-1 multi-temporal observations to map open water bodies. *2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Pp. 614–619.
- Schroeder, R., K. C. McDonald, B. D. Chapman, K. Jensen, E. Podest, Z. D. Tessler, T. J. Bohn, and R. Zimmermann  
2015. Development and evaluation of a multi-year fractional surface water data set derived from active/passive microwave remote sensing data. *Remote Sensing*, 7(12):16688–16732.
- Sippel, S. J., S. K. Hamilton, J. M. Melack, and B. J. Choudhury  
1994. Determination of inundation area in the Amazon River floodplain using the SMMR 37 GHz polarization difference. *Remote Sensing of Environment*, 48(1):70–76.

Smith, L. C.

1997. Satellite remote sensing of river inundation area, stage, and discharge: a review. *Hydrological Processes*, 11(10):1427–1439.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins

2004. Grace measurements of mass variability in the earth system. *Science*, 305(5683):503–505.

Tran, D. T.

2003. Researches in estuarine environment and ecosystem of red river: An overview on activities and result. *Science and Technics Publishing House, Hanoi*, X(Collection of works on Marine Environment and Resources):34–53.

Trenberth, K. E., L. Smith, T. Qian, A. Dai, and J. Fasullo

2007. Estimates of the global water budget and its annual cycle using observational and model data. *Journal of Hydrometeorology*, 8(4):758–769.

Tucker, C. J.

1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2):127 – 150.

Vermote, E.

2015. MOD09A1 MODIS Surface Reflectance 8-Day L3 Global 500m SIN Grid V006. *NASA EOSDIS Land Processes DAAC*.

Verpoorter, C., T. Kutser, D. A. Seekell, and L. J. Tranvik

2014. A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18):6396–6402.

Voormansik, K., J. Praks, O. Antropov, J. Jagomagi, and K. Zalite

2014. Flood Mapping With TerraSAR-X in Forested Regions in Estonia. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(2):562–577.

Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S. E. Bunn, C. a. Sullivan, C. R. Liermann, and P. M. Davies

2010. Global threats to human water security and river biodiversity. *Nature*, 467(7315):555–561.

Wahr, J.

2015. Time-variable gravity from satellites. *Treatise on Geophysics (Second Edition)*, Pp. 193 – 213.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna

2004. Time-variable gravity from grace: First results. *Geophysical Research Letters*, 31(11). L11501.

Wang, Y.

2004. Seasonal change in the extent of inundation on floodplains detected by JERS-1 Synthetic Aperture Radar data. *International Journal of Remote Sensing*, 25(13):2497–2508.

Wania, R., J. R. Melton, E. L. Hodson, B. Poulter, B. Ringeval, R. Spahni, T. Bohn, C. A. Avis, G. Chen, A. V. Eliseev, P. O. Hopcroft, W. J. Riley, Z. M. Subin, H. Tian, P. M. van Bodegom, T. Kleinen, Z. C. Yu, J. S. Singarayer, S. Zürcher, D. P. Lettenmaier, D. J. Beerling, S. N. Denisov, C. Prigent, F. Papa, and J. O. Kaplan

2013. Present state of global wetland extent and wetland methane modelling: methodology of a model inter-comparison project (WETCHIMP). *Geoscientific Model Development*, 6(3):617–641.

Wolter, K. and M. Timlin

1993. Monitoring enso in coads with a seasonally adjusted principal component index. *17th Climate Diagnostics Workshop*, Pp. 52–57.

Wolter, K. and M. S. Timlin

1998. Measuring the strength of enso events: How does 1997/98 rank? *Weather*, 53(9):315–324.

Wouters, B., J. A. Bonin, D. P. Chambers, R. E. M. Riva, I. Sasgen, and J. Wahr

2014. Grace, time-varying gravity, earth system dynamics and climate change. *Reports on Progress in Physics*, 77(11):116801.

Xu, H.

2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14):3025–3033.



- Yamazaki, D., M. A. Trigg, and D. Ikeshima  
2015. Development of a global 90m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, 171:337–351.
- Yang, H., Z. Wang, H. Zhao, and Y. Guo  
2011. Water body Extraction Methods Study Based on RS and GIS. *Procedia Environmental Sciences*, 10:2619–2624.
- Yang, X., S. Zhao, X. Qin, N. Zhao, and L. Liang  
2017. Mapping of urban surface water bodies from sentinel-2 msi imagery at 10 m resolution via ndwi-based image sharpening. *Remote Sensing*, 9(6).
- Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kito  
2012. Aphrodite: Constructing a long-term daily gridded precipitation dataset for asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93(9):1401–1415.
- Zhai, K., X. Wu, Y. Qin, and P. Du  
2015. Geo-spatial Information Science Comparison of surface water extraction performances of different classic water indices using OLI and TM imageries in different situations. *Geo-spatial Information Science*, 18(1):32–42.

