SATELLITE-BASED MONITORING OF SURFACE WATER DYNAMICS

Linlin Li

SATELLITE-BASED MONITORING OF SURFACE WATER DYNAMICS

DISSERTATION

to obtain the degree of doctor at the University of Twente, on the authority of the rector magnificus, prof.dr. T.T.M. Palstra, on account of the decision of the Doctorate Board, to be publicly defended on Wednesday 22 January, 2020 at 14.45 hrs

by

Linlin Li born on 30 April 1986 in Hebei, China This thesis is approved by:

Prof.dr. A.K. Skidmore, supervisor **A/Prof.dr. A. Vrieling,** co-supervisor **A/Prof.dr. T. Wang**, co-supervisor

ITC dissertation number 373 ITC, P.O. Box 217, 7500 AE Enschede, The Netherlands

ISBN 978-90-365-4941-7 DOI 10.3990/1.9789036549417

Cover designed by Linlin Li. The basemap in the front cover was from Esri World Imagery Printed by ITC Printing Department Copyright © 2020 by Linlin Li



Graduation committee:

Chairman/Secretary Prof.dr.ir. A. Veldkamp University of Twente Supervisor Prof.dr. A.K. Skidmore University of Twente Co-supervisors A/Prof.dr. A. Vrieling University of Twente A/Prof.dr. T. Wang University of Twente Members

Prof.dr. Z. SuUniversity of TwenteProf.dr. D. van der WalUniversity of TwenteProf.dr. W.G.M. BastiaanssenTechnical University DelftA/Prof.dr. A.R. MuñozUniversity of Málaga

To my family

Acknowledgements

A PhD journey is truly a marathon event full of challenges. Alongside the challenges, there is also the joy of the journey itself and the many inspiring people encountered along the way. I would like to thank everyone who was involved in this journey.

First and foremost, I would like to express my sincere thanks to my promoter, Prof. dr. Andrew K. Skidmore, who provided me the opportunity to undertake this PhD research at ITC and led me to the scientific world in the first place. It has been an honour to be his PhD student. He always encouraged me to keep learning new things and seeking answers to "why". He has taught me, both consciously and unconsciously, how to think critically and defend my views, and how to work smarter instead of harder. I am very grateful for his continuous guidance and support.

This thesis would not have been possible without the support and guidance of my daily supervisor A/Prof. dr. Anton Vrieling. I am so lucky to have him as my supervisor for this PhD research. During all these years, he guided me with infinite patience and full support, providing useful discussions, constructive criticisms, enlightened ideas, and excellent English editing. He was always the first person helping me out whenever I got stuck, and encouraged me in those tough times in my PhD pursuit. I would like to send my sincere gratitude to him for his significant contributions, efforts and trust to help me passing the finish line.

I am in particular thankful to my supervisor A/Prof. dr. Tiejun Wang, who introduced this interesting topic to me and offered clear guidance in writing the research proposal at the beginning of my PhD. His constant availability for discussion, his ability to explain many research topics in an interesting way, and his enthusiasm for research was contagious and motivational for me.

My special thanks go to A/Prof. dr. Antonio-Román Muñoz from University of Málaga, who helped me collecting water level data, arranged everything for the field work in Málaga, and provided precious advices to the first article of this research. I enjoyed very much the collaboration with Dr. Eren Turak from the Office of Environment and Heritage (NSW) on the first two articles of this research. He provided many constructive suggestions via emails discussions and Skype meetings.

I extend my special gratitude to Esther who was always ready to help. I appreciate more her importantly friendship and companionship throughout my PhD marathon. My sincere thanks also go to Willem Nieuwenhuis for always being so kind and for his technical assistance in data processing. I am thankful to Loes, Theresa, Marga, and Benno for their assistance and support throughout my stay and my research. I would especially like to thank Andy, Yousif, Roshanak, and Thomas, for giving interesting lectures and presentations.

My time at ITC was enjoyable in large part due to the many friends and groups that became a part of my life. I am grateful for the time spent with roommates, friends and PhD fellows at ITC. My special thanks go to Zhihui, a lifetime friend and a funny roommate. I will never forget our memorable trips in Europe, and countless happy hours together cooking and watching TV. I am grateful for Fangyuan's comfort and support, and being my sunshine in my tough times. Many thanks to Hong and Wen Bai who always made me feel like home when I visited theirs house. Linlin Pei was like a big sister to me, gave me a lot of support in life and took care of me when I was sick. I had a great time with the three visiting scholars: Haiting, Yongsheng and Changlin, during their shortterm visit at ITC. I thank them for their hospitality when I was in Wuhan. Thank all my office mates Mitra, Abebe, jing, Haili, Sam and Marcelle. I enjoyed our discussions and talks during coffee breaks. I also appreciate Lucas' help with my job search and interview. I would like to thank the following people for sharing research experiences, exchanging ideas, giving presentations and practicing debating skills during our monthly PhD tutorials: Mitra, Abebe, Parinaz, Sonia, Anna, Festus, Elnaz, Maria, Fangyuan, Yiwen, Zhihui, Jing, Haidi, and Xi. My time at ITC was also enriched by joining the Yoga group and badminton team, which helped me keep active. It is an unforgettable moment when we won the badminton match on the international sports day.

I would also like to convey my thanks to the Chinese community at ITC. They are: Donghai & Xiaojing, Yiwen, Ying, Xi, Peiqi, Mengmeng, Yifang, Shaoning, Bingbing, Xu & Xiaolong, Junping, Ruosha, Xiaoling, Yifei, Xin, Yijian, Tina, Xuelong, and Lichun. I am also grateful to Bob Su who is so kind and supportive. I still remember him helping us get to know ITC and celebrating the Mid-Autumn Festival with us because we are far away from home.

Enormous thanks go to my parents and brother, for all their unconditional love and support in all my pursuits. And most of all to my loving and supportive partner Zhao, thank you for being my rock.

Table of Contents

Ackn	owledge	ementsi
Table	e of Con	itentsiii
Chap	ter 1 G	eneral Introduction1
1	.1	The need for monitoring surface water2
1	.2	Remote sensing for surface water monitoring
1	.3	An overview of optical remote sensing methods for surface water
d	letection	n and existing datasets7
1	.4	Research challenges
1	.5	Research Objectives
1	.6	Thesis structure
Chap	ter 2 Ev	valuation of MODIS Spectral Indices for Monitoring Hydrological
Dyna	mics of	a Small, Seasonally-Flooded Wetland in Southern Spain21
A	Abstract	
2	2.1	Introduction
2	2.2	Study area
2	2.3	Data
2	2.4	Methods
2	2.5	Results and discussion
2	2.6	Conclusions
Chap	ter 3 M	onitoring the Dynamics of Surface Water Fraction from MODIS
Time	Series i	in a Mediterranean Environment47
A	Abstract	
3	3.1	Introduction
3	3.2	Study area
3	3.3	Data
3	3.4	Methods
3	3.5	Results
3	3.6	Discussion
3	3.7	Conclusions

Chapter 4	A New Dense 18-Year Time Series of Surface Water Fraction	
Estimates	from MODIS for the Mediterranean Region	77
Abstr	act	78
4.1	Introduction	79
4.2	Study area	81
4.3	Data	83
4.4	Methods	88
4.5	Results	97
4.6	Discussion	106
4.7	Conclusions	109
Chapter 5	5 Evaluation of a New 18-year MODIS-Derived Surface Water Fra	ction
Dataset fo	or Constructing Mediterranean Wetland Surface Water Dynamics	111
Abstr	act	112
5.1	Introduction	113
5.2	Data	115
5.3	Methods	116
5.4	Results	121
5.5	Discussion	130
5.6	Conclusion	137
Appe	ndix	139
Chapter 6	ó Synthesis	147
6.1	Introduction	148
6.2	The potential and limitations of spectral information for monitoring	z
	hydrological dynamics	149
6.3	Advantages of sub-pixel surface water fraction mapping techniques	5
	for the quantification of small water bodies	151
6.4	Robustness of the machine learning approach for application to lar	ge
	regions	152
65	A new long and dense time series of water extent for monitoring	
0.5	surface water dynamics	153
6.6	Potential applications of the new surface water dataset	155
6.6	Potential applications of the new surface water dataset	154

6.7	Outlook	
Bibliograph	y	
Summary		
Samenvatti	ng	
Biography		

v

Chapter 1

General Introduction

1.1 The need for monitoring surface water

Terrestrial surface water comprises rivers, streams, lakes, ponds, reservoirs and other inland water bodies, which together cover approximately 3% of the global land mass (Pekel *et al.* 2016). Despite their limited global extent, wetlands are essential for both humans and ecosystem health. They provide water resources for various human uses, support high levels of biodiversity, and provide important and diverse habitat and ecosystem services (Dudgeon *et al.* 2006; Zedler and Kercher 2005). They also play a crucial role in the global hydrological cycle and climate system (Chahine 1992; Tranvik *et al.* 2009). Terrestrial surface water affects the climate system via land-atmosphere interaction processes such as methane (CH₄) and carbon dioxide (CO₂) exchange (Holgerson and Raymond 2016; Raymond *et al.* 2013; Tranvik *et al.* 2009), as well as other biogeochemical processes.

In spite of their fundamental importance, water related ecosystems are fragile and vulnerable to climate change and anthropogenic disturbance (Nath and K Deb 2010; Vörösmarty *et al.* 2000). Especially with the increasing human population and accelerated economic development, the exerted pressure on water resources will continue to increase in the coming years (Prigent *et al.* 2012; Vörösmarty *et al.* 2000). Natural factors affecting water bodies include anomalous high-rainfall-driven flood events (Cian *et al.* 2018), drought events due to rainfall deficits (van Dijk *et al.* 2013), seasonal thawing and snowmelt in spring (Watts *et al.* 2012), and longer-term environmental changes (Lutz *et al.* 2014; Street and Grove 1976). Many human activities directly affect the availability of water resources. Examples are groundwater pumping, drainage of wetlands, irrigation schemes, and construction of new dams. Anthropogenic changes in land surfaces such as urbanization, agriculture and deforestation also lead to changes in surface water.

These changes strongly affect ecosystem functioning, which further results in shifting species distributions and composition (Koning 2005; Robledano *et al.* 2010), especially for species that are sensitive to hydroperiod variability (Baldwin *et al.* 2006; Roshier *et al.* 2002). It may also affect other ecosystem functions including ground water recharge and nutrient cycling (Leibowitz 2003). Globally, the biodiversity of water-related ecosystems continues to decline at an alarming rate (Collen *et al.* 2014). In addition to these direct threats, the changes of surface water further influence climate change (Degu *et al.* 2011; Ekhtiari *et al.* 2017; Foley *et al.* 2003; Hossain *et al.* 2009; Kabat *et al.* 2004).

A range of global initiatives and policy frameworks, including the Sustainable Development Goals (SDGs) and the Aichi Biodiversity Targets under the Convention on

Biological Diversity (CBD), have aimed to ensure sustainable development of water resources, to reduce its changes, and to prevent the loss of biodiversity (CBD 2010; Griggs *et al.* 2013). Specifically, the SDG Target 6.6 highlights the need to measure 'Change in the extent of water-related ecosystems over time' (Dickens *et al.* 2017). Globally-consistent maps of surface water extent at high spatial and temporal resolution are much needed for assessing progress towards the Aichi targets for 2020 (Turak *et al.* 2017). The Global Climate Observing System (GCOS) includes the area of water bodies and water level as Essential Climate Variables (ECV), in support of climate change assessment and policy development (GCOS 2011).

Recognizing the importance of surface water, and to assist in monitoring whether targets are attained, it is crucial and urgent to accurately and efficiently monitor the location and temporal dynamics of surface water.

1.2 Remote sensing for surface water monitoring

Traditionally, *in situ* gauge measurements are the main data source for the understanding of hydrological dynamics. Gauge stations collect a variety of hydrological data, including water stage, discharge and streamflow, but provide little information about the spatial dynamics of surface water extent (Alsdorf *et al.* 2007). Gauge stations are typically located on large rivers, lakes and canals, and their distribution is non-uniform throughout the world. For more than two decades, gauge stations have declined dramatically in both developed and developing countries (Shiklomanov *et al.* 2002). However, even in places where gauges exist, legal and institutional restrictions often make the data unavailable for scientific purposes. Therefore, assessing changes in water resources at global scale is exceedingly difficult using *in situ* observations alone, owing to the restricted spatial coverage and limited availability.

Satellite remote sensing provides unique capabilities for mapping the location, extent, and changes of surface water bodies across a wide range of spatial and temporal scales. Compared to traditional *in situ* measurements, remote sensing is more efficient because of its geospatial consistency, accessibility, repeatability, and global coverage.

Two types of remote sensing instruments are suitable for monitoring earth surface water at multiple spatial scales, *i.e.*, microwave and optical sensors. Microwave sensors are able to function day and night under any weather condition and have the ability to penetrate clouds and partially also vegetation. Schumann and Moller (2015) conducted a detailed review of microwave remote sensing for flood inundation and found synthetic aperture radar (SAR) to be the most suitable microwave sensor type for monitoring flood inundation. However, the high costs associated with obtaining SAR datasets for large areas and in a timely fashion has until recently limited their usefulness in monitoring the global surface water dynamics.

Optical satellite data have commonly been employed and are the preferred source due to their straightforward interpretability of water features (Bioresita *et al.*, 2018), high availability of data, records of multiple decades, as well as suitable spatial and temporal resolutions (Huang *et al.* 2018a). Table 1.1 shows an overview of the most commonly used satellite systems for surface water detection as well as their features. Spatial resolution and temporal resolution are important characteristics of optical remote sensors, and relevant when deciding what sensor to use for detecting and monitoring surface water. Spatial resolution determines the level of spatial detail that is captured by the sensor. The temporal resolution describes the time it takes for a satellite sensor to revisit a specific area. Thus, fine spatial resolution sensors can accurately estimate the location and extent of surface water while fine temporal resolution imagery is effective for intensive monitoring and analysis of the dynamics of surface water. Generally, there is trade-off between spatial and temporal resolution, even though recent satellite systems like Sentinel-2 achieve since 2017 a combination of fine spatial (10 m) and temporal (5-day revisit) resolution.

Category	Satellite	Sensor(s)	Revisit	Spatial	Number	Operation
			time	resolution	Of	period
			(day)	(m)	bands	
Coarse spatial	NOAA/TIROS	AVHRR	0.5	1,100	5	1978–
resolution	SPOT	Vegetation	1	1,150	4	1998–2014
(≥250 m)	Aqua/Terra	MODIS	0.5	250-1,000	36	1999–
	PROBA-V	Vegetation	1	333-1,000		2013–
	Suomi NPP	VIIRS	0.5	375–750	22	2011-
	ENVISAT	MERIS	3	300	15	2002-2012
	Sentinel-3	OLCI	2	300	21	2016–
Medium spatial	Landsat	MSS/TM/	16	15-80	4–9	1972–
resolution		ETM+/OLI				
(10–250 m)	Terra	ASTER	16	15-90	14	1999–
Fine spatial	SPOT	HRV/HRVIR	26	2.2-20	4–5	1986–
resolution	Sentinel-2	MSI	5	10-60	13	2015-
(≤ 10 m)	IKONOS	Panchromatic	1.5-3	1–4	5	1999–
		Multispectral				
	QuickBird	Panchromatic	2.7	0.61-2.24	5	2001-
		Multispectral				
	WorldView	Panchromatic	1-4	0.31–2.40	4–17	2007-
	RapidEye		1-5.5	5	5	2008-
	ZY-3		5	2.1-5.8	4	2012-
	GF-1/GF-2		4-5	4-5	5	2013-

Table 1.1. An overview of optical satellite sensors frequently used to map and monitor surface water.

Coarse spatial resolution sensors

Coarse spatial resolution remote sensors (≥ 250 m) offer multispectral measurements using a wide swath and consequently a high temporal resolution. A typical example is the Advanced Very High Resolution Radiometer (AVHRR) on board the National Oceanic and Atmospheric Administration (NOAA) satellites. This sensor was originally designed to monitor the ocean and atmosphere but was later found to be effective in detecting largescale flood events (Wiesnet *et al.* 1974). Since then, many studies have examined the ability of NOAA/AVHRR to monitor flood inundation at regional to global scale, taking advantage of its high frequency of global coverage, wide swath and low cost (*e.g.*, Barton and Bathols 1989; Dietz *et al.* 2017; Klein *et al.* 2014; Sheng *et al.* 2001).

The Moderate-Resolution Imaging Spectroradiometer (MODIS), flown on two NASA satellites: Terra and Aqua, has been extensively used in many land surface applications including surface water due to its global coverage, short repeat time, broad coverage and free availability. Since 2000, MODIS has accumulated an almost two-decade-long data, which makes it perfect for tracking changes in the surface water over long time. A number of studies have used MODIS for monitoring surface water at large scales (e.g., Kaptue et al. 2013; Khandelwal et al.; Ovakoglou et al. 2016; Pekel et al. 2014; Sharma et al. 2015). The utility of MODIS for monitoring flood has been repeatedly demonstrated by Dartmouth Flood maps disseminated by the Observatory (http://floodobservatory.colorado.edu).

The Visible Infrared Imaging Radiometer Suite (VIIRS) sensor aboard the Suomi National Polar-orbiting Partnership (Suomi-NPP) satellite, launched in 2011, is a new generation of operational coarse-resolution (375–750 m) sensor. It is considered to be the upgrade and replacement of AVHRR and MODIS. Several studies have shown the potential of VIIRS in detecting and monitoring surface water (Huang *et al.* 2015; Huang *et al.* 2017), and flood (Lacava *et al.* 2019; Li *et al.* 2018b) at local scales.

Another coarse-resolution sensor is the newly launched Ocean and Land Color Instrument (OLCI) onboard Sentinel-3 (Sentinel-3A launched in 2016 and Sentinel-3B launched in 2018). It provides 21 visible and infrared bands at 300 m resolution allowing global coverage in every two days. Only few studies have explored the potential of Sentinel-3 OLCI image for water body mapping (*e.g.*, Wang *et al.* 2019), and more are expected in the near future.

Medium spatial resolution sensors

Landsat imagery is the most popular data source for surface water mapping because of its suitable spectral bands, medium spatial resolution (30 m), as well as long term continuous record. The sensors on the early Landsat missions are the Multispectral Scanner (MSS), which was later upgraded to Thematic Mapper (TM) on Landsat-4 and Landsat-5, the Enhanced Thematic Mapper Plus (ETM+) on Landsat-7, and the Operational Land Imager (OLI) on Landsat-8. The opening of access to the Landsat mission data by NASA in 2008 greatly expanded its applications for long-term mapping of surface water (Díaz-Delgado *et al.* 2016; Dong *et al.* 2015; Feyisa *et al.* 2014; Jin *et al.* 2017; Schaffer-Smith

et al. 2017). Launched in 2013, Landsat-8 is the most recent Landsat satellite, but its OLI data have already been widely used in detecting surface water (Wang *et al.* 2018a; Xia *et al.* 2017; Yang *et al.* 2015). Recently the open access in combination with cloud computing have boosted planetary-scale monitoring of land surfaces such as tree cover (Hansen *et al.* 2013) and surface water (Donchyts *et al.* 2016; Pekel *et al.* 2016).

Fine spatial resolution sensors

Fine spatial resolution imagery (≤ 10 m) provided among others by SPOT, IKONOS, QuickBird, RapidEye, Worldview, and ZY-3 have also been proved effective in mapping surface water bodies or flood inundation (Davranche *et al.* 2010; Fisher and Danaher 2013; Xu *et al.* 2004). The fine spatial resolution allows small water bodies being accurately detected. However, due to the limited extent and availability of fine resolution imagery, the studies were usually focused on rather small, image-footprint limited regions.

The Sentinel-2 MultiSpectral Instrument (MSI), with Sentinel-2A launched in 2015 and Sentinel-2B launched in 2017, offers an unprecedented combination of fine spatial resolution (10–60 m), frequent revisit (5-day repeat), systematic global coverage, and a wide field of view (295 km) (Drusch *et al.* 2012; Gascon *et al.* 2017). It provides new opportunities for surface water monitoring at both fine spatial and fine temporal resolution (Du *et al.* 2016; Kaplan and Avdan 2017b; Ogilvie *et al.* 2018a; Yang *et al.* 2018b).

1.3 An overview of optical remote sensing methods for surface water detection and existing datasets

1.3.1 Methods for surface water detection

The main principle of surface water detection from multispectral satellite images is the significantly lower reflectance of water in infrared channels, compared to that of other land cover types (Figure 1.1). Based on this, various methods have been developed for detecting surface water from optical remote sensing imagery.



Figure 1.1. Reflectance of several typical land cover objects (*https://speclab.cr.usgs.gov/spectral-lib.html*).

Hard binary classification approaches, which map pixels as either water or non-water, are widely used. The commonly used binary classification method applies thresholding to a single band (*e.g.*, Frazier and Page 2000; Jain *et al.* 2005; Klein *et al.* 2014; Ryu *et al.* 2002), or to spectral indices (*e.g.*, Gao 1996; McFeeters 1996; Wang *et al.* 2015b; Wang *et al.* 2018a; Wang *et al.* 2018b; Xiao *et al.* 2002a; Xu 2006). A number of water indices have been developed, and their performances for water detection were also examined and compared (*e.g.*, Boschetti *et al.* 2014; Fisher *et al.* 2016; Li *et al.* 2015; Rokni *et al.* 2014; Zhou *et al.* 2017). Pixel-based classification techniques, either supervised or unsupervised (Manavalan *et al.*, 1993; Ozesmi & Bauer, 2002), can be used to generate land cover maps from which water maps could be extracted. Decision trees were also built using multispectral bands to separate water coverage from other land cover classes (Acharya *et al.*, 2016; Olthof, 2017; Sun *et al.*, 2011).

Soft classification methods do not assign a pixel to one class, but instead estimate fractions of different covers within the pixel. As such, these methods can compensate for the limitations of coarse resolution images for mapping water bodies that have a similar size or a smaller than a pixel. Sub-pixel fraction maps may be obtained through the use of regression modelling. For example, Guerschmann *et al.* (2011) developed a logistic regression model using spectral information and a DEM to predict water fraction on the Australian continent. Weiss and Crabtree (2011) developed multi-linear

regression models to estimate surface water fraction from MODIS based on spectral indices. Muster *et al.* (2013) used simple linear regression model to estimate surface water fraction in three Arctic tundra wetlands. Other studies built linear multivariate regression model to predict sub-pixel surface water (Frohn *et al.* 2012; Gómez-Rodríguez *et al.* 2010; Huang *et al.* 2014b; Reschke and Huttich 2014; Rover *et al.* 2010). These efforts were empirical and were developed for specific study regions, thus limiting their applicability to other regions.

Spectral unmixing is a widely used soft classification method, which is based on the premise that a pixel's observed reflectance can be modelled as a linear combination of all end-member spectra of the features within the pixel, weighted by their respective fractional abundance (Adams *et al.* 1995). It has been used to increase mapping precision by estimating sub-pixel water fractions from coarse resolution data such as AVHRR (Hope *et al.* 1999) and MODIS (Li *et al.* 2013b; Schroeder *et al.* 2015), and Landsat (Li *et al.* 2013b; Olthof *et al.* 2015). A substantial challenge in spectral unmixing is to determine the spectra and number of endmembers. Most studies that used spectral unmixing used between two and four endmembers. Nonetheless, this number may be inadequate to spectrally characterize a complex and heterogeneous landscape. Moreover, endmembers are considered pure surface components, but they often show important spectral diversity themselves. For example in the case of water, the spectral signature varies according to water composition (*e.g.*, algae, sediment and dissolved organic matter), submerged aquatic vegetation and bottom reflection, which also depends on water depth (Hommersom *et al.* 2011; Jensen 2009).

An alternative approach for surface water fraction estimation is the use of machine learning techniques such as support vector regression, multivariate adaptive regression splines, artificial neural networks and regression-tree algorithms (*e.g.*, Drzewiecki 2016; Rover *et al.* 2010; Xia *et al.* 2017). These methods can achieve higher perdition accuracy, but they require a large amount of training data from field data or higher resolution imagery from the same time period.

1.3.2 Existing water-related datasets

Many satellite-derived surface water-related datasets have been developed during the last decade. These datasets vary in geographic scope, temporal extent of the record, spatial resolution an in frequency of surface water estimates.

At the global scale, a number of static water body maps exist, with a spatial resolution ranging from 14.25 m to 1 km (see Table 1.2). Coarse resolution (\geq 250 m) maps include the 30 arc-second (~1 km) Global Lakes and Wetlands Dataset (GLWD: Lehner and Doll 2004), the 1 km land-water mask (Salomon *et al.* 2004), the global inundation extent from Multi-Satellite at 15 arc-second (~500 m) (GIEMS-D15: Fluet-Chouinard *et al.* 2015), the 300 m ESA Climate Change Initiative Water Bodies Product (ESA CCI-WB: Santoro *et al.* 2013), and the MODIS 250 m land/water mask (MOD44W: Carroll *et al.* 2009). Finer spatial resolution maps include the 1 arc-second (~30 m) Shuttle Radar Topography Mission (SRTM) Water Body Dataset (SWBD 2005), which covers the globe between 60°N and 56°S, and the 30 m Global Inland Water (GIW) dataset (Feng *et al.* 2015). The finest-resolution global water bodies map currently available is the GLObal WAter BOdies database (GLOWABO) with 14.25 m resolution (Verpoorter *et al.* 2014). It is derived from observations of the Enhanced Thematic Mapper Plus (ETM+) sensor onboard the Landsat 7 satellite collected in year 2000 ± 3 years.

Besides these dedicated water products, global land-cover maps also contain a water class. They have been developed at five spatial resolutions including 1 km (Bartholomé and Belward 2005; Hansen *et al.* 2000; Loveland *et al.* 2000), 500 m (Friedl *et al.* 2002; Friedl *et al.* 2010), 300 m (Arino *et al.* 2008; Arino *et al.* 2007), 250 m (Wang *et al.* 2015a), and 30 m (Chen *et al.* 2015; Gong *et al.* 2013; Yu *et al.* 2013).

An overview of global water-related datasets can be found in Hu *et al.* (2017). Several studies (*e.g.*, Nakaegawa 2012; Pham-Duc *et al.* 2017) performed a comparison of these global water maps. These static maps represent a snapshot of water extent for a particular time but do not seek to provide an understanding of the variability in water body extent over time.

In recent years, needs for dynamic and long-term mapping of surface water are growing. Especially with the increasing availability of freely available satellite time series data, improved computational capacities and the development of novel water detection techniques, dynamic mapping and monitoring of surface water over multi-decadal time periods and at different spatial scales has become feasible.

At local to regional scales, numerous applications have performed long-term surface water analysis for representative large lakes, floodplains, deltas, large wetland complexes, or large river basins. This has been achieved with for example Landsat (Jin *et al.* 2017; Schaffer-Smith *et al.* 2017), but also with coarser-resolution satellite data and shorter time

intervals including 16-day (Ordoyne and Friedl 2008; Weiss and Crabtree 2011), 10-day (McCarthy *et al.* 2003), 8-day (Yang *et al.* 2011) and daily (Chen *et al.* 2013) time steps.

At continental scale, a few datasets have been developed. For example, a Small Water Bodies (SWB) product for the African continent was developed from 10-day composites of 1 km SPOT VEGETATION (SPOT-VGT) data from 1999 to 2014 (Bartholomé 2007). It is based on a contextual algorithm exploiting the local contrast of the water surface with respect to the surrounding region (Gond et al. 2004). After the termination of the SPOT-VGT mission in May 2014, the algorithm was extended to the PROBA-V 1 km datasets to ensure the continuity of the service (Bertels et al. 2016). The SWB is available from the Copernicus GIO Global land portal (http://land.copernicus.eu/global/products/wb). In addition, a near real-time water surface dataset providing dynamic information about the water surfaces at 8-day temporal resolution and 250 m spatial resolution for the African continent over a 7-year period (2004 to 2010) was developed by Pekel et al. (2014). For the North American continent, the U.S. Geological Survey (USGS) has developed a Dynamic Surface Water Extent product (DSWE) (Jones 2015), which provides surface water inundation per-pixel derived from Landsat 4-8 data. This product is available from EarthExplorer. For the Australian continent, Guerschman et al. (2011) generated of a time series of fractional cover of standing water at 500 m and 8-day time step for the period 1999 to 2010, using a empirical statistical approach with the MODIS bands and derived indices. Its strengths and limitations were disscussed by Ticehurst et al. (2014). Recently, Mueller et al. (2016) presented a 25-year surface water product, called Water Observations from Space (WOfS), from analyzing the entire Landsat archive across Australia using a regression tree algrithom. WOfS is publicly accessible through www.ga.gov.au/wofs.

Global efforts to monitoring surface water dynamics have typically focused on either relatively coarse spatial resolution or long-time intervals (Table 1.3). For example, higher spatial resolution dynamic maps have been produced using Landsat satellite imagery at long time intervals (5-year) by Yamazaki *et al.* (2015). Coarse spatial resolution data has been used to produce time series of global water maps at relatively short intervals. For example, the MODIS 250 m land/water mask (MOD44W: Carroll *et al.* 2009) was produced annually for 2000–2015 using a decision tree classification method (MOD44W_Version 6: Carroll *et al.* 2017). A global inundation fraction map was produced at a monthly interval but at a coarse spatial resolution of 25 km, using multiple satellites including Advanced Very High Resolution Radiometer (AVHRR), passive microwave Special Sensor Microwave/Imager (SSM/I), and active microwave

scatterometer on board the European Remote Sensing (ERS) satellite (GIEMS: Papa *et al.* 2010; Prigent *et al.* 2007). This product was later downscaled to a 90 m spatial resolution using topographic and hydrologic information (GIEMS-D3: Aires *et al.* 2017). Recently, much progress has been made with global Landsat-based surface water assessment. The global surface water (GSW) datasets (Pekel *et al.* 2016) quantified changes in global surface water over the past 32 years with a monthly time interval. With the growing need for global near-real time monitoring, a few studies and datasets have advanced to daily time resolution by using MODIS data. Examples include the global near-real-time flood detection maps (<u>https://floodmap.modaps.eosdis.nasa.gov/</u>) produced by the Dartmouth Flood Observatory, 500-m resolution daily global surface water change database (2001–2016) developed by Ji *et al.* (2018), and the Global WaterPack which maps daily global inland water bodies at 250 m resolution for the years 2013–2015 (Klein *et al.* 2017).

Table 1.2. An overview of existing global	static water maps.			
Dataset name	Spatial resolution	Type of acquisition	Temporal scale	Reference
GLWD: Global Lakes and Wetlands Dataset	~ 1 km	Compilation of historical maps	Pre-1990s	Lehner and Doll (2004)
GLWM: Global Land-Water Mask	1 km	MODIS	for circa 2001	Salomon et al. (2004)
GIEMS-D15 : Global Inundation Extent from Multi-Satellite-15 arc-seconds	500 m	ENVISAT-ASAR, MERIS		Fluet-Chouinard <i>et al.</i> (2015)
ESA CCI-WBP: ESA CCI Water Bodies Product	~300 ш	ENVISAT-ASAR, MERIS	for circa 2010	Santoro <i>et al.</i> (2013)
MOD44W: MODIS Land/Water Mask	250 m	MODIS, SRTM	for circa 2000–2001	Carroll et al. (2009)
GIW: Global inland surface water dataset	30 m	Landsat	for circa 2000	Feng et al. (2015)
HydroLAKES: Global database of lakes	Shapefiles of lakes larger than 0.1 km ²	Compilation of historical inventory and datasets		Messager <i>et al.</i> (2016)
GLOWABO: Global Water Bodies	14.25 m	Landsat and SRTM	2000 ± 3 years	Verpoorter et al. (2014)

Chapter 1

Table 1.3. An overview of existing global dyn.	amic water	datas ets.			
Database name	Temporal interval	Spatial resolution	Temporal scale	Observations	Reference
G3WBM: Global 3 arc-second Water Body Map	5-year	~90 m	1990–2001	Landsat	Yamazaki <i>et al.</i> (2015)
MOD44W_Vesion 6: MODIS Land/Water Mask	annually	250 m	2000–2015	MODIS	Carroll <i>et al</i> . (2017)
GIEMS : Global Inundation Extent from Multi-Satellite	monthly	~25 km	1993–2004	AVHRR, SSMI, Scatterometer ERS	Papa <i>et al.</i> (2010); Prigent <i>et al.</i> (2007);
GIEMS-D3 : Global Inundation Extent from Multi Satellite-3 arc-second	- monthly	~90 m	1993–2007	AVHRR, ERS SCAT, SSMI/SSMIS,	Aires <i>et al.</i> (2017)
GSW: Global surface water	monthly	30 m	1982–2015	Landsat	Pekel <i>et al.</i> (2016)
Aqua Monitor	16-day	30 m	1985–2016	Landsat	Donchyts et al. (2016)
GWP: Global WaterPack	daily	250 m	2013–2015	MODIS	Klein <i>et al.</i> (2017)
Daily global surface water change database	daily	500 m	2001–2016	MODIS	Ji <i>et al</i> . (2018)

1.4 Research challenges

1.4.1 Characteristics of surface water bodies

Small in size: Small water bodies are abundant globally. Literature suggests that small to intermediate-size surface water bodies (*e.g.*, $0.01-10 \text{ km}^2$) together account for a large fraction of the total global surface water area (Downing *et al.* 2006; Verpoorter *et al.* 2014). A high resolution (14.25 m) global water map (GLOWABO: Verpoorter *et al.* 2014) derived from circa 2000 Landsat scenes shows that there are ~27 million water bodies larger than 0.01 km² globally, with a total surface area of $4.76 \times 10^6 \text{ km}^2$ excluding the Caspian Sea. Analysis of the distribution of water size and number (Figure 1.2) shows that the largest total water area corresponds to water bodies in the 0.1–1 km² size-range, followed by 1–10 km² and 0.01–0.1 km². These three categories (0.01–10 km²) together make up about 52% of the global inland water by total area, and 99.9% by total amount (Verpoorter *et al.* 2014).

Recent evidence points to the significant role of small water bodies in many natural processes. For example, small ponds tend to have higher concentrations of CO_2 and CH_4 than large lakes and thus are important for global carbon cycle (Holgerson and Raymond 2016). Small water bodies are also ecologically important as they provide habitats for a wide range of species including rare and declining species (Biggs *et al.* 2017; Bolpagni *et al.* 2019; Downing 2008). In spite of their importance, small water bodies are typically missed from present assessments of global surface water area (Downing *et al.* 2006; Ogilvie *et al.* 2018a).



Figure 1.2. The distribution of number (N) and area of global water bodies (excluding the Caspian Sea) detected by Downing et al. 2006 and GLOWABO: Verpoorter et al. (2014). Numbers on y-axis are the lower/upper boundary of decadal size classes (Figure source: Verpoorter et al. 2014).

Highly dynamic: Surface water demonstrates large variability. Water bodies fluctuate in size and location due to natural and anthropogenic processes. A wide range of dynamic patterns can be observed according to the frequency of inundation and duration of standing water (Cowardin *et al.* 1979). The 8-day repeat coverage is considered to be a minimum for effectively capturing water bodies with short hydroperiods while simultaneously accounting for frequent cloud cover (Guerschmann *et al.* 2011; Wulder *et al.* 2016). However, in extreme cases such as flooding, water is only present for a few days. Statistical estimates (Najibi and Devineni 2018) from the global flood database of the Dartmouth Flood Observatory (DFO) suggest that, during last three decades, 59 % of the global total flood events have short duration floods (1–7 days), while 27 % have moderate duration floods (8–20 days). The recent Global Climate Observing System (GCOS) report states that Essential Climate Variables (ECV) need to be established for water extent and lake ice cover products ideally with daily temporal resolution (Belward 2016).

1.4.2 Large variability of water spectral signatures

The spectral signature of water bodies varies greatly over time and space. This is caused by variations in water depth, bottom material, sediment load, chlorophyll concentration, dissolved organic matter, aquatic vegetation, algae, turbidity and any combination of these variables (Hommersom *et al.* 2011; Jensen 2009; Klein *et al.* 2017). Furthermore, the spectral response is affected by changing atmospheric conditions, sun angle, and sensor view angle, which further complicates surface water mapping (Liu 2012; Ticehurst *et al.* 2014). This makes the reliable detection of water difficult particularly at large spatial scales (*e.g.*, national, continental and global) where a large range of water conditions may be expected.

1.4.3 Presence of noise

Optical remote sensing imagery are often plagued with noise and outliers, due to cloud and aerosol contaminations. These can significantly decrease the relevant information of images and affect the performance of a classifier (Karpatne *et al.* 2016). Shadows induced by *e.g.*, clouds, mountains and buildings show similar spectral characteristics as water bodies, making it difficult to distinguish them from the water class (He *et al.* 2016; Li *et al.* 2013a; Lin *et al.* 2019).

1.4.4 Lack of high-quality training data

Supervised learning algorithms generally achieve higher mapping accuracy as compared to unsupervised approaches, which is largely due to the use of discriminative information contained in the labelled training datasets. Training data can be gathered on the basis of ground measurements, visual interpretation (*e.g.*, Jin *et al.* 2017; Pekel *et al.* 2016; Pekel *et al.* 2014; Tulbure *et al.* 2016), or classification of fine resolution satellite imagery and/or aerial photography from Google Earth (http://earth.google.com). Applying supervised learning algorithms at large-scale (*e.g.*, national to global) and for long periods of time is challenging, because this requires training data be collected across various geographic regions and various points in time due to the great variability of spectral signature of water bodies over time and space as described above. This usually requires considerable time and effort. Moreover, the training data contains certain errors, which will affect the accuracy of classifier. Therefore, it is important that the training data have high accuracy.

1.4.5 Spatial and temporal resolution

One of the challenges for long-term monitoring surface water using remote sensing data is the existing trade-off between spatial and temporal resolution of satellite imagery. In recent years, satellite constellations have been launched that combine fine spatial (3–10 m) with fine temporal (daily to 5-day revisit) resolution. Examples of this include Sentinel-2 and PlanetScope. Nevertheless, these systems are recent and consequently cover only a few years of data until present. The combination of fine spatial and fine temporal resolutions is important for accurate monitoring of surface water, especially in arid, semi-arid and Mediterranean regions, where small-sized water bodies are abundant and they exhibit large variability in response to variability in precipitation and evapotranspiration (Ruiz 2008).

Recently, much progress has been made towards long-term global Landsat-based assessment of surface water (*e.g.*, Pekel *et al.* 2016). However, such assessments can have large temporal gaps due to both the limited number of acquisitions during specific time intervals, and location- and time-dependent persistency of cloud cover. They may not always be effective in capturing the dynamics of seasonally inundated water, not to mention rapid changes and short-duration surface water. Coarser-resolution data such as MODIS can generate denser and long time series data, thus can fill the spatial and temporal gaps of Landsat-based dataset. Sub-pixel mapping from coarse resolution imagery is one way to deal with the trade-off between the spatial resolution and temporal resolution of remotely sensed data, and can ensure both small-sized and highly dynamic water bodies being capture.

1.5 Research Objectives

The main objective of this research is to develop improved capabilities for long-term mapping and monitoring surface water dynamics over a large geographical area at fine temporal resolution inclusive of small (<1 km²) water bodies, using freely available optical remote sensing data. To achieve this, four specific objectives are formulated as follows:

- To explore the potential of various spectral indices for monitoring the temporal variability in the hydrology of a small wetland.
- To evaluate machine learning algorithms that incorporate MODIS spectral information and a topographic metric for accurate estimation of surface water

fraction, and to test the transferability of the algorithms to different geographic and climatic zones.

- To develop a global rule-based regression-tree model for the estimation of surface water fraction for the whole Mediterranean region from MODIS data, and to develop a new long-term surface water fraction dataset at 500 m resolution and at 8-day interval for the Mediterranean region.
- To assess if our new MODIS water fraction dataset can effectively capture the temporal changes in surface water extent for a wide range of water bodies with different sizes (*i.e.*, 0.01–3100 km²).

1.6 Thesis structure

This thesis comprises six chapters, a general introduction, four core chapters and a synthesis. Each core chapter has been prepared as a stand-alone research paper that has been published in or submitted to a peer-reviewed ISI journal. The six chapters are organized as follows:

Chapter 1 presents the research background, reviews the relevant literature on existing water-related datasets and methods used to map and monitor surface water from various satellite sensors, presents research challenges and gaps, research objectives, and the structure of the thesis.

Chapter 2 assesses and compares the applicability of several indices for characterizing the temporal variability in the hydrology of a shallow seasonally flooded wetland in southern Spain.

Chapter 3 develops an approach for the estimation of surface water fraction from MODIS data. This is done by exploring the use of MODIS spectral information and a topographic metric as input to machine learning algorithms (*e.g.*, random forest, rule-based regression model) and testing the transferability of the models to different geographic and climatic zones in Spain.

Chapter 4 transfers the approach used to derive surface water fraction developed in Chapter 3 to larger spatial scales (*i.e.*, the Mediterranean region), with considerable improvements regarding input data, training data and commission error processing. The algorithm was applied to 18 years of MODIS data (2000–2017) to generate a time series of surface water fraction maps at 8-day interval for the Mediterranean region.

Chapter 5 provides an in-depth evaluation of the dense 18-year surface water fraction (SWF) dataset developed in **Chapter 4**. It demonstrates the ability of the developed

MODIS dataset for constructing detailed surface water dynamics for hundreds of Mediterranean wetlands, especially for those with abnormal changes and short-duration water inundation. In addition, it shows how the MODIS dataset can accurately capture the dynamics of very small water bodies $(0.01-1 \text{ km}^2)$. The results are compared and validated with a Landsat-based dataset and water level data.

Chapter 6 summarizes the main findings and conclusions, and trends related to the use of Earth Observation (EO) in surface water mapping and monitoring.

Chapter 2

Evaluation of MODIS Spectral Indices for Monitoring Hydrological Dynamics of a Small, Seasonally-Flooded Wetland in Southern Spain^{*}

^{*} This chapter is based on: Li, L., Vrieling, A., Skidmore, A., Wang, T., Muñoz, A.R., & Turak, E. (2015). Evaluation of MODIS spectral indices for monitoring hydrological dynamics of a small, seasonally-flooded wetland in southern Spain. *Wetlands*, 35, 851–864.

Abstract

Monitoring spatio-temporal dynamics of hydrology in seasonally-flooded wetlands is important for water management and biodiversity conservation. Spectral data and derived indices from the Moderate Resolution Imaging Spectroradiometer (MODIS) have been used for hydrological monitoring of large wetlands. However, comparable studies for small wetlands (<25 km²) are lacking. Our aims are to examine whether MODIS-derived indices at 500 m spatial resolution can perform this task for small wetlands, and to compare the performance of various indices. First we evaluated if water levels are a good indicator for wetland inundation extent. A high correlation between water level and Landsat-derived inundation extent was found ($R^2 = 0.957$). Secondly, we compared 10 years of water level fluctuations with seven spectral indices at a 16-day interval. The Tasseled Cap brightness index (TCBI) had the highest correlation with water level for the complete time series including dry and wet years. Thirdly, we analyzed how these indices behave for areas with different inundation characteristics. Again TCBI showed a consistently accurate performance, which was independent of inundation frequency. We therefore conclude that TCBI is the best-suited index for monitoring of hydrological variability in small seasonally-flooded wetlands such as the Fuente de Piedra lake in southern Spain. We recommend testing this index further for other seasonally-flooded wetlands in semi-arid areas.

2.1 Introduction

Seasonally or intermittently flooded wetlands are ecologically important ecosystems in arid, semi-arid, and Mediterranean-type regions (Haas et al. 2009; Roshier et al. 2001; Waterkeyn et al. 2008). Forty-five percent of the Ramsar-listed seasonally-flooded inland wetlands are found in these climate zones. They undergo periodic cycles of inundation and drought, primarily in response to variability in precipitation and evapotranspiration (Ruiz 2008). In arid, semi-arid and Mediterranean environments, about 30 percent of Ramsar-listed seasonal wetlands are small-sized, measuring between 10 and 2,500 hectares. Despite their small size, they often act as critical refuge and breeding areas, offer food sources for wildlife, and harbor many plant and animal species that would otherwise not survive in the surrounding landscape (Gibbs 1993; Roshier et al. 2002; Semlitsch and Bodie 1998; Sim et al. 2013; Zacharias et al. 2007). There is concern that seasonal wetlands are often neglected due to their ephemeral character and small size. The abundance and quality of seasonal wetlands around the world is declining rapidly due to global climate change, expansion of agricultural land and irrigation schemes (Castañeda and Herrero 2008; Roshier et al. 2001; Zacharias and Zamparas 2010). Although the European Union (European Communities 1992) and the Ramsar Convention (Ramsar Convention on Wetlands 2002) include seasonal wetlands in their conservation plans, only a subset of them are considered. The lack of knowledge about the changes in wetland extent of these aquatic systems makes conservation a difficult task for resource managers. Hence, an urgent need exists at national and international levels to report and monitor changes in wetland extent and conditions for large areas, using cost-effective tools, and including the important small-sized wetlands.

Hydrological dynamic processes, mainly expressed by spatial and temporal variation in inundation status, are important determinants of the formation and maintenance of a seasonally flooded wetland. Hydrological modifications may strongly affect ecosystem functioning and normally result in shifting species distributions and composition (Koning 2005; Robledano *et al.* 2010), especially for species that are sensitive to hydroperiod variability (Baldwin *et al.* 2006; Roshier *et al.* 2002). It may also affect other ecosystem functions including ground water recharge and nutrient cycling (Leibowitz 2003). Therefore, it is important to monitor the wetland inundation dynamics for water management, ecosystem assessment and biodiversity conservation.

In situ water level gauges are a main data source for understanding hydrological dynamics and essential for quantifying temporal patterns of water fluctuation with good temporal

resolution (Alsdorf *et al.* 2007). Gauge stations are typically located on large rivers, lakes and canals, but less frequently in seasonally flooded areas. Due to the inaccessibility of certain regions or financial and operational constraints, globally many wetlands lack gauge stations resulting in a limited knowledge and understanding of their hydrological conditions (Alsdorf *et al.* 2003). While gauge measurements provide key data on the wetland hydrology, they may offer little information about spatial patterns of hydrologically-relevant variables like inundation status (Alsdorf *et al.* 2007; Huang *et al.* 2014a).

Remote sensing provides temporally and spatially continuous synoptic observation of ecosystem processes, and these observations may allow for monitoring of spatio-temporal hydrological variability for large areas in a repeatable and cost effective manner (Smith 1997). Two types of remote sensing instruments are suitable for monitoring wetland hydrology at local and regional scales, *i.e.*, microwave and optical sensors. The microwave technique of synthetic aperture radar (SAR) can provide imagery under all weather conditions and thus has been used for monitoring spatial and temporal patterns of flood inundation (e.g., Bourgeau-Chavez et al. 2005; Kim et al. 2014; Marti-Cardona et al. 2010; Richards et al. 1987; Townsend 2001; Wdowinski et al. 2008). A main disadvantage of SAR is that the resulting backscattered signal is a complex combination of effects that depend on incidence angle, vegetation density and orientation, relative water height and wind effects. This could cause opposite backscatter responses for similar conditions, and the effective disentangling of such effects requires additional information (O'Grady and Leblanc 2014; Smith 1997). Radar (or laser) altimeters can monitor water heights in reservoirs and lakes with a higher temporal resolution, but may miss many water bodies due to the spacing between the satellite orbits (Alsdorf et al. 2007). Optical remote sensors, such as those onboard Landsat and SPOT, have been used frequently for small wetlands monitoring. For example, Herrero and Castaneda (2009) used a series of 52 Landsat images to monitor the flooding surface of small saline wetlands in northeast Spain over the past 20 years. While providing accurate spatial information about flood extent, these sensors do not yet allow frequent and continuous monitoring over large regions that suffer from persistent cloud cover. In comparison, the relatively coarse spatial resolution (>100 m) imagery derived from satellite sensors such as AVHRR (Advanced Very High Resolution Radiometer) and MODIS (Moderate Resolution Imaging Spectroradiometer) provides more consistent and frequent observations over long timespans, making such imagery potentially well-suited for spatio-temporal analysis of wetland hydrology.
An often-used approach to study temporal changes from coarse-resolution optical sensors is to summarize their spectral information in multispectral indices and consequently study the spatio-temporal variation of these indices. The best-known multi-spectral index is the Normalized Difference Vegetation Index (NDVI) (Tucker 1979) that combines spectral reflectance measurements from red and near-infrared (NIR) bands. The NDVI provides a measure of the photosynthetic activity of the green vegetation and NDVI time series have been used extensively for monitoring vegetation dynamics (Beck et al. 2006; Pettorelli et al. 2005; Petus et al. 2013; Vrieling et al. 2011; Vrieling et al. 2013). It has also been used for water/land delineation (Borro et al. 2014; Chipman and Lillesand 2007) as water strongly absorbs light in the NIR spectral region (causing low reflectance) while much less absorption occurs over land surfaces. Other indices have been specifically developed for detecting and monitoring surface wetness (e.g., Gao 1996; McFeeters 1996; Xiao et al. 2002a; Xu 2006). These mostly combine shortwave infrared (SWIR) or near infrared (NIR) bands, *i.e.*, the spectral domain containing specific physical water absorption features, and visible (VIS) spectral regions. In general, NIR-SWIR indices are mainly proposed for vegetation water content detection while VIS-NIR (SWIR) combinations are almost all proposed for the detection of open water.

Several studies have explored coarse spatial resolution data for monitoring flood duration, timing and frequency of ephemeral wetlands (Chen *et al.* 2013; Feng *et al.* 2012; Guerschman *et al.* 2011; Huang *et al.* 2014a; Tornos *et al.* 2015; Xiao *et al.* 2005). Most of these studies used MODIS data to monitor flood extent by differentiating inundated/non-inundated or mixed pixels. Until present, there have been few attempts to link water level changes to temporal patterns of spectral indices. One exception is Ordoyne and Friedl (2008) who demonstrated the utility of multi-temporal MODIS data for characterizing the hydrologic regime of the Everglades in South Florida, USA. However, all MODIS studies focused on relative large wetlands covering at least 50 km². Whether MODIS and its derived spectral indices are suitable to describe and monitor variations in water level for small (< 25 km²) seasonal wetlands has not yet been tested.

This paper aims to explore the potential of MODIS-derived spectral indices for characterizing the temporal variability in the hydrology of small shallow wetlands. Specifically, the objectives are:

 to investigate if water level is a good proxy of wetland hydrological variability by establishing a relationship between water level variation and inundated area for a small (~14 km²) wetland in southern Spain;

- (2) to evaluate if time series of MODIS-derived spectral indices can effectively capture the hydrological variability of this wetland in relation to the water-level data; and
- (3) to explain how varying inundation characteristics within the wetland affect the temporal behavior of these indices.

2.2 Study area

The study was carried out in Fuente de Piedra lake (36°06' N, 4°45' W), a shallow and saline (athalassohaline) lake and associated marshland with an area of 1,364 hectares. It occupies the center of a topographically endorheic basin with a catchment area of about 150 km² in southern Spain, situated between the Guadalquivir watershed and the Guadalhorce watershed (Heredia *et al.* 2010) (see Figure 2.1). The study site is one of the most important breeding sites for the Greater Flamingo (Phoenicopterus roseus) in the Mediterranean region, second only to the Camargue, France (Geraci *et al.* 2012). Its natural values were recognized and listed as a Wetland of International Importance (Ramsar) and Special Protection Area for Birds (SPA). It is designated by the environmental council of the Andalusian regional government as a Nature Reserve, and therefore it is a protected site.

The wetland is fed by five small rivers, rainfall, and highly mineralized ground water (Kohfahl *et al.* 2008). Evaporation from the lake surface constitutes the main water output. It has a maximum depth of approximately 70 cm during the 10-year study period and experiences both seasonal and interannual variations of water level and inundation extent that are predominantly linked to variability in precipitation and evaporation. The mean annual rainfall is 460 mm and mean annual evaporation is approximately 1600 mm. Usually the lake is flooded in autumn (September–October), has its highest water levels during spring (February–March), and dries up partially or completely around June and July (García and Niell 1993; Kohfahl *et al.* 2008). During these summer months the excess evaporation causes salt to deposit on the soil (see Figure 2.1c).

Different vegetation communities are found within and outside the wetland system. Dense vegetation containing reeds, saltmarshes and tamarisks (see Figure 2.1e) are present in channels feeding into the lake, and form a natural purification buffer of runoff water entering the lake. On small elevated dikes and islets inside the lake, drought- and saline-tolerant vegetation (*e.g., Sarcocornia, Suaeda* and *Arthrocnemun*) is present (Ministry of Agriculture Fisheries and Environment 2013) (see Figure 2.1d). In drier years with low water levels, surface water does not reach this vegetation, but during wet years it may be

partially inundated (Wang 2008). Planktonic and submerged macrophytes are generally negligible except for extremely wet years such as 1990 and 1998 (Conde-Álvarez *et al.* 2012; García *et al.* 1997). During the period considered in this study (2000–2009) when annual precipitation levels were low, no significant development of aquatic vegetation was observed, which can partly be attributed to the efforts to purify the wastewater from nearby towns (Ministry of Agriculture Fisheries and Environment 2013). Surrounding the wetland area, olive trees and wheat are cultivated. While this is predominantly rainfed agriculture, groundwater extraction from wells is sometimes used as supplementary irrigation.



Figure 2.1. Location of the study area. (a) Color composite image displaying bands 7, 4, 2 as RGB (Landsat 7 ETM+ on May 2, 2000); (b) map of Spain showing the location of Fuente de Piedra; (c), (d) and (e) photos (June 2014) showing exposed soil, salt-resistant vegetation and marshland surrounding the lake. The approximate location of the photos is shown in (a).

2.3 Data

2.3.1 Water level data

The Fuente de Piedra lake water level data have been collected since 1983 using a limnograph. This mechanical recorder draws the curve of water level fluctuations on a paper by registering the movements of a flute floating in a well which connects with the lakebed. The daily mean values of the water level are subsequently calculated and stored in a database. The instrument measures the water surface height from the bottom of the lake (*i.e.*, the 0-cm water level indicates that the lake is dry, even though values below 0 representing groundwater levels are recorded as the well is deeper than the lake). In this study, a 10-year daily mean water level dataset between 2000 and 2009 was obtained from Consejería de Medio Ambiente, Natural Reserve of Fuente de Piedra (Junta de Andalucía).

2.3.2 Remote sensing imagery and pre-processing

We used 78 Landsat TM/ETM+ images of Path/Row 201/34, acquired through the Global Visualization Viewer (GLOVIS; http://glovis.usgs.gov/) of the United States Geological Survey. The TM sensor has a spatial resolution of 30 m for the six reflective bands and 120 m for the thermal band. Landsat ETM+ images consist of eight spectral bands with a spatial resolution of 30 meters for bands 1 to 5 and band 7. Resolution for band 6 (thermal infrared) is 60 meters and resolution for band 8 (panchromatic) is 15 meters. In our study, we used the 30 m visible and near-infrared bands of TM and ETM+. All images are radiometrically- and terrain-corrected products (L1T) and have a scene quality score of 9, which means perfect scenes with no errors detected. Table 2.1 summarizes all the Landsat TM/ETM+ dataset used in this study and corresponds to all cloud free images available for the study area from 2000 to 2009 (10 years). In the case of Landsat 7 ETM +, images acquired after May 2003 have wedge-shaped gaps of missing data on both sides of each scene as Scan Line Corrector (SLC) was damaged. However, as our study area is in the scene center, no gaps occurred over the area. We transformed the raw digital numbers (DN) contained in the images to top of atmosphere (TOA) reflectance according to the Landsat 7 Science Data Users Handbook (Irish 2000).

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
January		1	1		1	1	1			
February	1		1							
March			1					1	1	1
April			1	1	1	1			1	1
May	2		1	1	1			1		1
June	1	1	2		1	1	2	2	1	1
July	1		1	2		1			1	2
August	1	1		2				1	1	2
September		1	1		1	1		1		1
October				1	1				1	1
November		1		1		2	2	1	1	1
December	1				2	1	1	1	1	
Total	7	5	9	8	8	8	6	8	8	11

Table 2.1. Number of Landsat images available per year (2000–2009) and month that are cloud-free over the study area, and consequently used in this study.

The MODIS sensor has 36 spectral bands of which seven are specifically designed for studying the land surface, *i.e.*, blue (459–479 nm), green (545–565 nm), red (620–670 nm), near infrared (841–876 nm), and shortwave infrared (SWIR1: 1230–1250 nm, SWIR2: 1628–1652 nm, SWIR3: 2105–2155 nm). The MODIS Land Science Team provides a suite of standard MODIS data products to the users, including the Nadir BRDF-Adjusted Reflectance (NBAR) 16-day composite product (MCD43A4). This product provides 500-meter resolution reflectance data for each of the MODIS bands (1–7) adjusted using a bidirectional reflectance distribution function (BRDF) to model the values as if they were taken from nadir view (Schaaf *et al.* 2002). The product thus removes view angle effects, and in addition masks cloud cover and reduces atmospheric contamination. For this study MCD43A4 composites were used to evaluate if coarse-resolution index time series can capture the hydrological variability of the Fuente de Piedra lake. We used all 16-day composites between 2000 and 2009, resulting in 227 images.

2.4 Methods

2.4.1 Evaluating water level as a proxy for inundated area

To evaluate if water level gauge measurements provide a good proxy for inundated area and thus allow to effectively describe the wetland's hydrological conditions, we used the Landsat scenes to set a high-resolution baseline. To discriminate water from non-water we used the Normalized Difference Water Index (NDWI) developed by McFeeters (1996). It has been widely used as an index for surface water detection (Bai et al. 2011; Rokni et al. 2014). We acknowledge that NDWI may not be the most effective for inundation detection in case of dominant floating vegetation or submerged vegetation (Rodriguez et al. 2014), but this was not the case for the wetland considered here (section 2.2). The NDWI ranges from -1 to 1, with values above 0 generally representing water bodies. However the threshold values applied to separate water from land may vary significantly from one scene to the next due to aerosol interference and variable solar/viewing geometry (Feng et al. 2012; Ji et al. 2009). Therefore we established threshold values for each individual NDWI image using the Otsu thresholding algorithm. The algorithm assumes that the image contains two classes of pixels following a bi-modal histogram (foreground pixels and background pixels). It then calculates the optimum threshold separating the two classes that minimizes the weighted within-class variance (Otsu 1979).

The inundated area derived from each of the 78 Landsat scenes was linked to the corresponding water levels. We fitted a second-order polynomial regression through the data to describe the inundation status in relation to water level variation. When the recorded water level is 0 cm, the lake is dry, *i.e.*, the inundated area is 0 km². Therefore, the regression analysis is only performed for those dates when the recorded water level was above 0 cm.

In addition, we calculated the per-pixel water occurrence frequency (WOF) by evaluating for each pixel the ratio between the number of images for which the pixel was inundated and the total number of images (*i.e.*, 78). Based on this, we classified the wetland areas into five classes: never inundated (WOF = 0), seldom inundated ($0 < WOF \le 0.15$), occasionally inundated ($0.15 < WOF \le 0.3$), sometimes inundated ($0.3 < WOF \le 0.45$) and often inundated ($0.45 < WOF \le 0.6$).

2.4.2 MODIS-derived spectral indices

A range of spectral indices have been proposed to perform surface wetness detection from multi-spectral imagery in different contexts. These all follow the same logic as the NDVI (Tucker 1979), i.e., a difference between two spectral bands divided by the sum of the two. McFeeters (1996) introduced the Normalized Difference Water Index (NDWI) to delineate open water features using the green and near-infrared (NIR) band. Xu (2006) found that NDWI often does not distinguish between water areas and built-up land, and proposed the Modified Normalized Difference Water Index (MNDWI) by substituting the SWIR band for the NIR band. Several spectral indices combining the NIR and SWIR bands have been proposed using different portions of the SWIR region (Ji et al. 2011). These include the Normalized Difference Water Index (NDWI, Gao 1996) (referred as LSWI_{B5} in this paper) and the Land Surface Water Index (LSWI, Xiao et al. 2002b) (referred as LSWI_{B6} in this paper) which use the SWIR-band centered at 1.24 and 1.64 μm, respectively. The combined NIR/SWIR indices are sensitive to leaf water and soil moisture and for this reason widely adopted for studying vegetation phenology, vegetation change and seasonal inundation (Campos et al. 2012; Davranche et al. 2013; Ordoyne and Friedl 2008; Xiao et al. 2006; Xiao et al. 2005; Yan et al. 2010). Table 2.2 summarizes the five most common band-ratio indices adopted for water detection (including open water, vegetation water content, and soil moisture).

Table 2.2. List of spectral i	indices tested in this study	ÿ.			
Spectral index	Original purpose	Equation	MODIS bands	Code in this manuscript	Reference
Normalized difference water index	Open water detection	$NDWI = \frac{\rho_{Green} - \rho_{NIR}}{\rho_{Green} + \rho_{NIR}}$	B4, B2	IMUN	McFeeters (1996)
Modified Normalized Difference Water Index	Open water detection	$MNDWI = \frac{\rho_{Green} - \rho_{SWIR}}{\rho_{Green} + \rho_{SWIR}}$	B4, B6	IMUNM	Xu (2006)
Normalized difference water index	Vegetation liquid water	$NDWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}}$	B2, B5	LSWI _{B5}	Gao (1996)
Land surface water index	Rice flooding mapping	$LSWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}}$	B2, B6	LSWI _{B6}	Xiao <i>et al.</i> (2002b)
Normalized difference vegetation index	vegetation monitoring; Water mapping	$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}}$	B2, B1	IAUN	Tucker (1979) ; Chipman and Lillesand (2007)

32

		<i>s</i> - <i>jj</i>	- J				**=/*
Index	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
Tasseled Cap							
Brightness	0.3956	0.4718	0.3354	0.3834	0.3946	0.3434	0.2964
Index (TCBI)							
Tasseled Cap							
Wetness	0.10839	0.0912	0.5065	0.4040	-0.2410	-0.4658	-0.5306
Index (TCWI)							

Table 2.3. Tasseled Cap coefficients for MODIS NBAR (source: Zhang et al. 2002).

Except for spectral band-ratio indices using two multispectral bands, there are also approaches synthesizing information contained in multiple bands. The tasseled cap transformation, first suggested by Kauth and Thomas (1976) for Landsat MSS, is a useful tool for compressing spectral data into a few bands that can be directly associated with the physical parameters of the land surface (Crist 1985; Crist and Cicone 1984). The first three components of the Tasseled Cap transformation are brightness, greenness and wetness. Brightness, hereinafter referred to as "Tasseled Cap Brightness Index (TCBI)", is a weighted sum of all six bands and correlated to texture and moisture content of soils (Crist *et al.* 1986). Greenness is a contrast between near-infrared and visible reflectance, and is thus a measure of the presence and density of green vegetation. Wetness, hereinafter referred to as "Tasseled Cap Wetness Index (TCWI)", contrasts the sum of the visible/ near infrared with the shortwave infrared bands, providing a measure of soil moisture tension (Crist *et al.* 1986; Jian *et al.* 2012) and plant moisture (Cohen 1991; Jin and Sader 2005; Toomey and Vierling 2005). The coefficients of TCBI and TCWI proposed by Zhang *et al.* (2002) were used for MODIS data (Table 2.3).

For each 16-day MCD43A4 composite, we calculated all the spectral indices that are designed to be correlated to surface wetness. Because the resulting MODIS-derived time-series of $LSWI_{B5}$ and $LSWI_{B6}$ showed a strong similarity, we only present $LSWI_{B6}$ results in this paper.

2.4.3 Evaluation of MODIS-derived spectral indices for monitoring hydrological variability

For each 16-day time step and spectral index we calculated the mean value of the MODIS pixels contained within the lake. Only pixels whose centers were inside the lake boundary were included. The temporal patterns of average MODIS indices for the entire lake were

then analyzed. This resulted in a single average 10-year time series from 2000 to 2009 for the lake for each spectral index with a 16-day temporal resolution.

To analyze which index most accurately describes the temporal hydrological variability, we averaged the water level data to 16-day periods corresponding to the MODIS composite period. Pearson's correlation coefficients were calculated between each water index series and the corresponding water level data.

Besides the lake-average indices, we also performed correlation analysis between water level and index time series for single MODIS pixels. For this analysis we included pixels within 1 km outside the lake as the surrounding vegetation may also reflect the hydrology condition in the lake. In an attempt to group areas of similar behavior, we converted the Landsat-derived inundation frequency map to vector layers. From these layers we calculated the spatial mean index values of the MODIS pixels contained within each inundation occurrence zone and within the 1km buffer area, and then related these to water level data. In addition, we calculated the dynamic range for each index time series (defined here as the absolute difference between the 5th to 95th percentile of all index values) to examine which spectral index responds most strongly to the hydrological fluctuations. To better understand how the behavior of each spectral index within an inundation zone relates to the spectral properties, we also calculated the spatial mean reflectance for each MODIS band and each inundation class for several different water levels ranging from -30 to 62 cm.

2.5 Results and discussion

2.5.1 Relationship between water level and inundated area

Figure 2.2 and Figure 2.3 illustrate the changes in inundated area as a function of water level. The data are most accurately fitted with a second order polynomial that indicates a strong positive relationship ($R^2 = 0.957$) between water level and inundated area. Similar high correlations using a second order polynomial were also achieved by Sippel *et al.* (1998) and Jung *et al.* (2011), but other mathematical relationships are also found between the two parameters, including a linear function (Liu *et al.* 1983), a power function (Hayashi and van der Kamp 2000), and an exponential function (Mahe *et al.* 2013; Mahe *et al.* 2011). The type of relationship is largely dependent on the lake bathymetry (Alsdorf *et al.* 2007; Medina *et al.* 2010) or river morphology (Smith 1997). Other effects also play a role. For example wind effects (Wang 2008) may explain the range of inundated area

between 4.0 and 6.5 km² around the water level of 22 cm (Figure 2.3). Despite the underestimation of inundation area at low water levels (< 5cm) and the scattering around the regression line, Figure 2.3 shows a strong relationship between water level and inundation area. Consequently water level data can be used in this study as a proxy for the important hydrological fluctuations occurring in the wetland.



Figure 2.2. Inundation maps derived from multi-temporal Landsat imagery corresponding to different water levels and corresponding date.



Figure 2.3. Relationship between measured water level and Landsat-derived inundation areas. A second order polynomial was fitted to the data.

Figure 2.4 presents the inundation frequency map obtained from the number of times each pixel was inundated in the 78 Landsat images. Central regions (shown in blue) experience variable water depths during the wet season. Never inundated regions (shown in white) occur in the southwest and north. The southwest areas are known as "Canchones del Suroeste", which are unique natural islets in the lake providing good nesting areas for flamingos only in years with very high water levels when foxes and other predators cannot reach these islets (Rendón-Martos 1996). These areas are mudflats and partly covered by

drought- and salt-resistant vegetation (*e.g., Sarcocornia, Suaeda* and *Arthrocnemun*) (Figure 2.1d). The northern fringe of the lake is covered by dense vegetation which is a combination of marsh vegetation mixed with reeds, rushes and tamarisks (Figure 2.1e).



Figure 2.4. Inundation frequency classes map obtained from the number of Landsat images between 2000 and 2009 that showed inundation. WOF stands for water occurrence frequency.

2.5.2 MODIS-derived indices vs. water level

Figure 2.5 presents a visual comparison of the multi-temporal profiles of six spectral indices throughout a 10 year period (2000–2009) in conjunction with water level. During the 10 years of study, the water level of the lake has varied substantially, both within and between years. All hydrological cycles have a filling phase of strong water level increase during September to November. Water levels remain relatively high until spring (February–March). This is followed by a drying period with dropping water levels, resulting in a complete dry up of the lake around June–July. Based on the inter-annual variations in lake water levels, two categories of hydrological years can be identified: relatively wet years (2000 to 2004 and 2009) and relatively dry years (2005 to 2008). During wet years, the lake water levels start rising from late September to mid-October. Throughout the wet years, water depths are higher than the multi-annual average and water persists on the surface for longer periods (Mid-October to late June). Dry years are characterized by longer dry seasons which can start as early as May and end in October.



Figure 2.5. Comparison between spectral index time series from MODIS and water level data. The solid red line represents in-situ water level measurements; the blue points represent spatial mean MODIS index values of the Fuente de Piedra lake for each MODIS composite; the purple dashed line represents the multi-annual average water level fluctuation; the black dashed line indicates the 0 cm water depth (no open water in the lake).

Table 2.4. Pearson's correlation coefficients between in-situ water level data and MODIS-derived indices for the whole time series (2000–2009), for the wet years only (2000–2004 and 2009), and for the dry years only (2005–2008).

	2000-2009	2005-2008	2000–2004 and 2009
NDWI	0.58	0.29	0.66
MNDWI	0.52	0.74	0.39
LSWI _{B6}	0.28	0.75	0.22
NDVI	0.12	0.35	0.15
TCBI	-0.87	-0.92	-0.88
TCWI	0.67	0.92	0.53

Visual inspection of the temporal graph (Figure 2.5) indicates a relatively stronger correspondence between water level and MNDWI, TCBI and TCWI. All the indices except TCBI were positively related to the water level data. The TCBI values exhibited a smooth, regular seasonal annual pattern with a higher variability in index values. The TCWI and MNDWI curves also showed clear seasonal patterns but with more noise. During the 10-year study period, the TCWI and MNDWI values were mostly below 0, with a smaller dynamic range of TCWI than that of MNDWI. The other indices showed a more random pattern. One exception is that LSWI_{B6} closely follows the water level fluctuation during dry years (2005–2008).

Table 2.4 shows the Pearson's correlation coefficients between each water index and the corresponding water level data. There was a strong negative correlation between TCBI and water level time series, which is consistent with the findings by Ordoyne and Friedl (2008) for the Florida Everglades. TCBI is a linear combination of all spectral bands with positive coefficients (Table 2.3); the higher the overall reflection of the ground surface across the spectrum, the higher the TCBI. Because water is a strong absorber of radiation across the visible and infrared portion of the spectrum, a decrease of water in the system (including open water and soil moisture), will cause an increase in TCBI. When water retreats completely from the ground surface (*i.e.*, no open water), measured water level variation is associated with the relative water content in the soil. A high water content makes soils darker (hence low TCBI) than if these were dry, particularly for soils with low organic matter content (Jensen 2009). Strong correlations were consistent between TCBI and water level for both dry and wet years, suggesting that TCBI is a reliable index for hydrological monitoring in the Fuente de Piedra lake under a wide range of conditions.

In contrast to TCBI, TCWI had the highest positive correlation coefficient with water level (r = 0.67) for the 10-year period, with a much higher correlation coefficient (r = 0.92) for the dry years (2005–2008). This result confirms findings by other studies (Ordoyne and Friedl 2008; Van Trung *et al.* 2013) and can be explained by the fact that water strongly absorbs radiation in the SWIR part of the electro-magnetic spectrum (Campbell 2002). The correlation coefficient for wet years (2000–2004 and 2009) was much lower (r = 0.53) and can be attributed to higher TCWI values in the dry season (May–July) of these wet years as compared to dry years, even if the water level is the same. These high values may be explained by salt deposition on the soil surface during the dry season, which occurs particularly in wet years when more water evaporates. The presence of salt causes the TCWI to increase considerably due to the high spectral reflectance of crusted saline soil surfaces in the visible and near-infrared regions of the spectrum (Howari *et al.* 2002; Koshal 2012). The different findings for the dry and wet years suggest that TCWI can be a proper index for monitoring wetland hydrological variability in non-saline wetlands, but not for saline wetlands with seasonal salt deposits.

While TCBI and TCWI clearly exhibited the highest correlation with water level fluctuations, we found that band-ratio indices also gave reasonable correlation coefficients, but this was largely season-dependent. For dry years, LSWI_{B6} and MNDWI had a stronger relation with water level than NDVI and NDWI. This could be partly attributed to the smaller sensitivity of NIR reflection to variations in soil moisture, vegetation water content (Eitel *et al.* 2006; Olsen *et al.* 2013) and open water (Campbell 2002; Li *et al.* 2003) as compared to SWIR reflection. Low correlation between LSWI_{B6}/MNDWI and water level for wet years may again be explained by the accumulation of salt at the soil surface (as for TCWI). NDWI showed a higher correlation in open water extent for which NDWI is designed. For all years, the NDVI-water level correlation coefficients were very low (r < 0.4).

2.5.3 Temporal behaviour of MODIS-derived indices in response to inundation characteristics

Figure 2.6 shows the correlation coefficient between MODIS-derived indices and water level for individual MODIS pixels. Important spatial variations can be observed. Generally, for NDWI, LSWI_{B6} and NDVI the sign of the correlation coefficient was opposite when comparing areas inside and outside the lake, while for TCBI, TCWI and MNDWI, the sign was the same.

Comparison of Pearson's correlation coefficients for different inundation frequency zones indicated that the sign and strength of correlations between water level and the MODIS-derived indices were highly dependent on inundation frequency (Table 2.5). TCBI was an exception as it showed high correlations with water level in all six inundation frequency classes. This result demonstrated that TCBI was sensitive to hydrological variability as expressed by fluctuations in both open water extent and soil moisture. Because the TCBI is a linear combination of all MODIS bands with positive coefficients (Table 2.3), TCBI's sensitivity to hydrological variability across classes should be explained by an overall increase of reflectance with decreasing water levels. In fact, Figure 2.7 shows that this is the case for all inundation classes. This suggests that TCBI is well suited for monitoring relative wetness under a wide range of hydrological conditions.

Table 2.5. Pearson's correlation coefficients between spectral indices derived from MODIS for different inundation frequency zones and in-situ water level data over 2000–2009.

	NDWI	MNDWI	$LSWI_{B6}$	NDVI	TCBI	TCWI
Upland	-0.69	0.40	0.81	0.79	-0.73	0.80
Never inundated	-0.11	0.85	0.86	0.61	-0.86	0.94
Seldom inundated	0.21	0.47	0.46	0.42	-0.89	0.75
Occasionally inundated	0.50	0.48	0.26	0.14	-0.88	0.72
Sometimes inundated	0.70	0.51	0.09	-0.26	-0.86	0.53
Often inundated	0.84	0.57	-0.08	-0.45	-0.84	0.33



Figure 2.6. Pearson's correlation coefficient between spectral indices derived from MODIS and measured water level for the 2000–2009 period for each MODIS pixel.



Figure 2.7. Average MODIS spectral signatures for the different inundation classes. The X-axis represents the central wavelength of each of the seven MODIS bands. Each line provides a spatial mean for an inundation class for a single image that corresponds to the water levels shown in the legend.

TCWI was positively related to water level in all inundation classes. The highest correlation (r = 0.94) existed in never inundated areas followed by upland (r = 0.80). This finding can be explained by the fact that TCWI is closely associated with both soil moisture (Crist and Cicone 1984; Crist et al. 1986; Jian et al. 2012) and plant water content (Cohen 1991; Jin and Sader 2005; Toomey and Vierling 2005). TCWI takes the difference between a weighted sum of the visible/NIR reflectance and the SWIR reflectance (Table 2.3). TCWI would be most sensitive if the reflectance of both wavelength domains would change in the opposite direction. Figure 2.7 shows that this is not the case for any inundation class, but for the never inundated areas, the visible/NIR domain (469 nm-858 nm) changes least while SWIR (1240 nm-2130 nm) shows strong increases with decreasing water levels, resulting in a high sensitivity for TCWI. The likely explanation for this behavior in the never inundated class, is that TCWI responds here predominantly to changes in water content of soil and vegetation. These results are supported by the study of Ordoyne and Friedl (2008) who concluded that TCWI can quantify hydrological variation when the water table is below the soil surface. With an increase in the frequency of inundation, the TCWI-water level correlation decreased, indicating that TCWI might not appropriate for detecting variability in the presence of open water. It is rather an indicator of the water content of soil and vegetation. $LSWI_{B6}$ exhibited a similar behavior to TCWI. The high LSWI_{B6}-water level correlations for never inundated and upland areas confirm that LSWI_{B6} is also a good indicator of vegetation and soil water content (Chen et al. 2005; Wang et al. 2008; Xiao et al. 2006; Xiao et al. 2005; Zhang et al. 2011).

NDWI had the highest correlation coefficient (r = 0.84) with water level for often inundated areas where the fluctuation in open-water presence is larger but low correlation coefficients in occasionally, seldom, and never inundated areas. This confirms that the index is able to monitor open water fluctuations, but not changes in soil water content. The upland class showed a high negative NDWI-water level correlation coefficient (r = -0.69), which can be explained by Figure 2.7 that shows a decrease in green reflectance (555 nm) and a simultaneous increase in NIR reflectance (858 nm) for upland as water level increases above 0 cm. This likely relates to reduced presence of green vegetation and/or the drying of vegetation around the lake at moments when the lake water level is low. Dry vegetation reflects less NIR radiation than green plants resulting in a higher NDWI. NDVI showed high and positive correlations with water level for upland and never inundated areas but low and negative correlations inside the lake. For a wetland in Australia, Petus *et al.* (2013) also found differential NDVI temporal behavior inside the wetland as compared to the surrounding regions. This can be attributed to the fact that different wetland plant species have different phenological responses in relation to water availability (Baird and Wilby 1999; Van Trung *et al.* 2013). In large areas inside the lake, plants that are able to stand extreme drought and salinity may appear at low densities when water retreats. Instead, in the upland and never inundated areas of the lake, the vegetation is relatively dense and dominated by crops, marsh and scrub that strongly depend on water availability, which is higher with higher water tables. These opposite correlation signs for NDVI may also partially explain the low correction coefficient for NDVI and water level for the whole lake (Table 2.4). While a good index for monitoring green vegetation abundance in response to water level in vegetated wetlands such as marshes (Jiang *et al.* 2015), NDVI proved not be appropriate for monitoring wetland hydrology in this saline lake with sparse vegetation.

The dynamic range of TCBI, expressed by the 5th to 95th percentile range, varied greatly in relation to different inundation characteristics with a higher value of 0.72 for usually inundated areas and a low value of 0.38 for upland areas (Table 2.6). This difference in dynamic range suggests that the TCBI variability could be used to spatially separate seasonally-flooded wetlands from other areas, even if the temporal behavior is similar. Also MNDWI and NDWI showed higher variability for inundated areas as compared to uplands suggesting MNDWI and NDWI would also allow to separate wetlands clearly from its surroundings. An additional factor for NDWI to achieve this separation is the opposite temporal behavior (Table 2.5 and Figure 2.6). This implies that MODIS-derived indicates can, besides monitoring wetland hydrology, also play a role in identification and mapping of small wetlands, and monitoring of their extent.

Table 2.6. The 5th to 95th percentile variability as derived from the area-average time series of each spectral index per inundation class.

	NDWI	MNDWI	LSWI _{B6}	NDVI	TCBI	TCWI
Upland	0.14	0.07	0.20	0.24	0.38	0.17
Never inundated	0.14	0.24	0.30	0.22	0.48	0.21
Seldom inundated	0.18	0.30	0.26	0.21	0.59	0.20
Occasionally inundated	0.18	0.36	0.25	0.15	0.64	0.20
Sometimes inundated	0.28	0.50	0.31	0.15	0.71	0.21
Often inundated	0.36	0.54	0.39	0.16	0.72	0.21

2.5.4 Future outlook for wetland monitoring with coarse spatial resolution multi-spectral data

Although our study was conducted for a single small wetland, it clearly showed the good potential of particularly the TCBI for wetland hydrological monitoring. We expect that this potential can be used for monitoring the hydrology of other wetlands with sparse vegetation cover in arid and semi-arid environments. The soils in and around the Fuente de Piedra lake are mostly mineral, which could have affected the results, because dry soils are bright (high TCBI) and wetter soils and inundated areas are dark (low TCBI) across the part of the electro-magnetic spectrum considered here. Mineral soils, low in organic matter content, are the dominant soils in arid and semi-arid environments. Soils with a higher organic matter content will also experience a decline in reflection when wetted or inundated, but given that such soils are already dark when in dry conditions, TCBI may possibly be less sensitive to this decline. Further studies in other wetlands should confirm the applicability and limitations of TCBI for monitoring wetland hydrology. Besides examining indices for wetlands with different soil types, we also recommend further testing for seasonally-flooded wetlands with other diverging characteristics, e.g., marshes with dense emergent vegetation, and for even smaller wetlands that comprise only a few MODIS pixels.

Closely monitoring hydrological variability is important for understanding how climate change and human actions affect the dynamics of seasonally-flooded wetlands, or may

affect these in the future. Given that seasonally-flooded wetlands host a high diversity of fauna and flora that depend on these dynamics, we need tools to study them. This study demonstrated a promising option to monitor wetlands remotely using freely-available coarse spatial resolution but high temporal resolution data, which could benefit the hydrological monitoring of many seasonally-flooded wetlands globally. This in turn has potential to greatly contribute to the management and conservation of these habitats and species living in them.

Taking into account the importance of the Fuente de Piedra lake for conservation (*e.g.*, for waterbirds), the results of this study may be directly applied for explaining animal species abundance by hydrological variability. Given that many of the waterbird species (*e.g.*, the Greater Flamingo), are not confined to only this wetland, but make breeding and feeding decisions based on wetland conditions in a wider area, the MODIS indices may also be applied for assessing spatio-temporal variation of wetlands across Spain and the Mediterranean Basin (Amat *et al.* 2007). For example, the colony size of greater flamingos at Fuente de Piedra is also affected by water levels in the Guadalquivir marshes, which are located 140 km away. The Fuente de Piedra lake usually dries up during the late breeding season, and flamingos breeding in this locality must move to Guadalquivir marshes to obtain their food during the chick-rearing period (Rendón-Martos 1996). Hence, simultaneous monitoring hydrological dynamics of various wetlands with MODIS may prove an important tool to better explain and predict animal populations.

2.6 Conclusions

Seasonally-flooded wetlands are among the world's most unique and valuable ecosystems, but are under threat worldwide. Methods that allow monitoring these small wetlands from remote sensing are urgently needed by resource managers and ecologists, especially in areas where conflicts arise between the water demand for agriculture and conservation of wetlands in semi-arid environments. Results from this work suggest that the MODIS-derived spectral indices have good potential for characterizing and monitoring temporal variability in the hydrology of small seasonally-flooded wetlands. Particularly TCBI proved useful and gave consistent good results for wet and dry years, and for areas characterized by different inundation frequencies. This is relevant as it could provide opportunity to improve hydrological monitoring particularly for data-poor and ungauged wetlands. We recommend further testing of MODIS indices for hydrological monitoring of seasonally-flooded wetlands with different soil and vegetation characteristics. Besides wetland hydrological monitoring, the differential temporal

behavior of MODIS indices within and outside the lake make these indices a promising tool for mapping and monitoring wetland extent over large areas.

Chapter 3

Monitoring the Dynamics of Surface Water Fraction from MODIS Time Series in a Mediterranean Environment^{*}

^{*} This chapter is based on: Li, L., Vrieling, A., Skidmore, A., Wang, T., & Turak, E. (2018). Monitoring the dynamics of surface water fraction from MODIS time series in a Mediterranean environment. *International Journal of Applied Earth Observation and Geoinformation*, 66, 135–145.

Abstract

Detailed spatial information of changes in surface water extent is needed for water management and biodiversity conservation, particularly in drier parts of the globe where small, temporally-variant wetlands prevail. Although global surface water histories are now generated from 30 m Landsat data, for many locations they contain large temporal gaps particularly for longer periods (>10 years) due to revisit intervals and cloud cover. Daily Moderate Resolution Imaging Spectrometer (MODIS) imagery has potential to fill such gaps, but its relatively coarse spatial resolution may not detect small water bodies, which can be of great ecological importance. To address this problem, we propose and test options for estimating the surface water fraction from MODIS 16-day 500 m Bidirectional Reflectance Distribution Function (BRDF) corrected surface reflectance image composites. The spatial extent of two Landsat tiles over Spain were selected as test areas. We obtained a 500 m reference dataset on surface water fraction by spatially aggregating 30 m binary water masks obtained from the Landsat-derived C-version of Function of Mask (CFmask), which themselves were evaluated against high-resolution Google Earth imagery. Twelve regression tree models were developed with two approaches, Random Forest and Cubist, using spectral metrics derived from MODIS data and topographic parameters generated from a 30 m spatial resolution digital elevation model. Results showed that accuracies were higher when we included annual summary statistics of the spectral metrics as predictor variables. Models trained on a single Landsat tile were ineffective in mapping surface water in the other tile, but global models trained with environmental conditions from both tiles can provide accurate results for both study areas. We achieved the highest accuracy with Cubist global model ($R^2 = 0.91$, RMSE = 11.05%, MAE = 7.67%). Our method was not only effective for mapping permanent water fraction, but also in accurately capturing temporal fluctuations of surface water. Based on this good performance, we produced surface water fraction maps at 16-day interval for the 2000–2015 MODIS archive. Our approach is promising for monitoring surface water fraction at high frequency time intervals over much larger regions provided that training data are collected across the spatial domain for which the model will be applied.

3.1 Introduction

Terrestrial surface water bodies play an important role in the global carbon cycle and climatic processes (Chahine 1992; Tranvik *et al.* 2009). They support a high level of biodiversity and provide a range of ecosystem services (Dudgeon *et al.* 2006; Zedler and Kercher 2005). Globally-consistent maps of surface water extent at high spatial and temporal resolution are a major information need for assessing progress towards the Aichi targets for 2020 of the Convention of Biological Diversity (Turak *et al.* 2017). One of the challenges for producing these maps is that surface water can exhibit a strong seasonal and inter-annual variability. For example, in arid and semi-arid regions, surface water bodies are often small and their size varies in time in response to changes in precipitation and evapotranspiration (Ruiz 2008). Hydrological changes may strongly affect ecosystem functioning and result in shifting species distributions and composition (Koning 2005; Robledano *et al.* 2010). Therefore, monitoring the dynamics of surface water is of critical importance for understanding the health and functioning of wetlands, and the ecosystem services they provide.

One option for monitoring dynamic surface water is using multi-temporal optical imagery of 10-30 m spatial resolution (e.g., the Landsat-series and Sentinel-2). A suite of recent studies have exploited the Landsat archive to assess long-term variability of surface water and flooding extent at the regional (Halabisky et al. 2016; Heimhuber et al. 2016; Tulbure et al. 2016), continental (Mueller et al. 2016) and global scale (Donchyts et al. 2016; Pekel et al. 2016). Although Landsat sensors can provide accurate spatial information on surface water extent and those with small sizes, until recently their revisit interval was too long (16 days and more) to capture rapid changes in water extent due to seasonal hydrological fluctuations or extreme weather events. Furthermore, Landsat archive features some temporal gaps depending on geographical location (Pekel et al. 2016). For areas that suffer from persistent cloud cover, there might be large temporal gaps between two clear observations. In a Landsat-based study of surface water monitoring, Halabisky et al. (2016) could effectively reconstruct surface-water hydrographs (i.e., temporal details on seasonal, intra-annual and long-term changes in surface water extent) for 750 wetlands in Douglas County, Washington, USA from 28 years of Landsat data, they indicated that their approach would not give meaningful results for wetlands with frequent cloud cover or with short hydroperiods. While revisit times are reducing for high resolution sensors (e.g., Sentinel-2 offers a 5-day repeat since March 2017: Feng et al.

2016), it will be some years before it becomes possible to construct long-term (>10 years) and dense time series of surface water over large regions.

Moderate resolution imagery derived from satellite sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) provides observations at a much higher frequency over long time-spans, with obvious advantages for monitoring surface water dynamics. A number of studies have used MODIS for surface water mapping (e.g., Kaptue et al. 2013; Khandelwal et al. 2017; Ovakoglou et al. 2016; Pekel et al. 2014; Sharma et al. 2015). These studies have in common that they provide a binary classification for each grid cell and time period, *i.e.*, water or no water. However, binary classifications cannot represent water-bodies that are smaller than the grid cell itself. For example, Khandelwal et al. (2017) successfully analyzed temporal variations in surface extent of global reservoirs using MODIS time series, but their method was only applicable for large water bodies comprising more than ten MODIS pixels. In areas with many (temporal) small water bodies that are equal or less the size of a grid cell, binary classification could result in large omission errors. Such omission errors may be trivial when tracking global changes in surface area, but can be essential for assessing biodiversity, which is greatly influenced by the density of small wetlands (Deane et al. 2017; Semlitsch and Bodie 1998).

To overcome the problem outlined above, it is possible to estimate surface water fraction; *i.e.*, the percentage of surface water within a single grid cell. Existing studies on surface water fraction mapping have used techniques like linear spectral mixture modeling (LSMM) and machine learning. For instance, Weiss and Crabtree (2011) developed a multi-linear regression model to estimate surface water fraction from MODIS using spectral indices, *i.e.*, normalized difference vegetation index (NDVI), normalized difference water index (NDWI), and tasseled cap indices for the Yukon Flats National Wildlife Refuge (Alaska, USA), but achieved a moderate model accuracy ($R^2 \approx 0.625$) between reference and modeled surface water fraction. LSMM has been widely used for the estimation of fractional cover of land surface components such as vegetation (Gan et al. 2014; Guerschman et al. 2015; Lu et al. 2003; Song 2005; Xiao and Moody 2005), snow cover (Painter et al. 2009; Vikhamar and Solberg 2003), urban area (Yang et al. 2014), and water (Hope et al. 1999; Li et al. 2013b; Olthof et al. 2015; Schroeder et al. 2015; Sheng et al. 2001; Sun et al. 2011). This approach is based on the premise that a pixel's observed reflectance can be modelled as a linear combination of all end-member spectra of the features within the pixel, weighted by their respective fractional abundance (Adams et al. 1995). A substantial challenge in linear unmixing is to determine the spectra and number of endmembers, which can be at a maximum of one endmember less than the number of spectral bands. For example, studies that used LSMM to estimate surface water fraction used between two and four endmembers. Nonetheless, this number may be inadequate to spectrally characterize a complex and heterogeneous landscape. Moreover, endmembers in LSMM are considered pure surface components, but they often show important spectral diversity themselves. For example in the case of water, the spectral signature varies according to water composition (*e.g.*, algae, sediment and dissolved organic matter), submerged aquatic vegetation and bottom reflection, which also depends on water depth (Hommersom *et al.* 2011; Jensen 2009).

An alternative approach for surface water fraction estimation is the use of machine learning techniques such as support vector regression, multivariate adaptive regression splines, artificial neural networks and regression-tree (RT) algorithms. Research comparing algorithms have shown that RT algorithms are often among the top performers across a range of applications on fractional surface cover mapping (e.g., Drzewiecki 2016; Xia et al. 2017). Unlike LSMM that needs to consider and estimate fractional cover for all endmembers within each pixel, RT is a nonlinear algorithm that can be used to derive fractional cover for a single specific land surface component. RT has been used extensively in remote sensing, for example to derive the percentage of tree cover (Hansen et al. 2002; Kobayashi et al. 2014), and to estimate fractional impervious surface area in the National Land Cover Data (NLCD) product by USGS (Xian et al. 2011). To our knowledge, only few researchers have attempted to estimate surface water fraction through time using RT algorithms. Rover et al. (2010) compared various methods to map surface water fraction in the Yukon Flats and showed that the RT method produced the highest accuracy ($R^2 = 0.93$, RMSE = 11%), outperforming the LSMM. In a case study of the 2005 Louisiana floods, Sun et al. (2012) applied a RT algorithm to derive water fraction from MODIS using as predictors MODIS band reflectance and spectral water indices and achieved high accuracy. Hence, RT algorithms seem to be promising for accurate mapping of surface water fraction over large regions. However, several important issues regarding the application of the RT algorithm have not yet been addressed in current literature. First, existing studies merely incorporated reflectance or reflectance-derived predictor variables, whereas the inclusion of ancillary environmental variables (e.g., digital elevation model) may potentially improve the accuracy of the regression tree models. Second, the transferability of trained regression tree models to larger areas is unknown, as existing studies focused on small areas with homogeneous climate characteristics. Third, no studies have examined how different surface water fraction and water-permanence types may affect model performance. In this study we aim to:

- evaluate two rule-based regression-tree methods that incorporate MODIS spectral information and a topographic metric derived from digital elevation model (DEM) for accurately mapping and monitoring surface water fraction;
- (2) assess the transferability of the resulting models to different geographic and climatic zones;
- (3) assess the accuracy of the method outcomes as a function of surface water fraction and of the variability in surface water presence.

3.2 Study area

We selected two study areas in Spain for building and training the models (see Figure 3.1a). The two areas have different characteristics in terms of climate, and each is defined by the spatial extent of a Landsat tile (*i.e.*, P199R031 and P202R034, hereinafter referred to as T_1 and T_2).

The first area is located in the middle and lower Ebro river basin, north-eastern Spain. The area has a wide variety of climatic environments due to the complex topography (Figure 3.1b), with annual precipitation ranging from 348mm to 1020mm, based on the WorldClim database (Hijmans *et al.* 2005). The Ebro is the largest river on the Iberian Peninsula and one of the largest in the Mediterranean region. Multiple dams are built along the Ebro river and its major tributaries, mainly for water storage in reservoirs, hydropower production and for agricultural irrigation (Batalla *et al.* 2004). The northern part of the area belongs to the Pyrenees mountains and contains several small lakes. Multiple endorheic lakes are also scattered across the Oceanic and semi-arid regions such as Sarinena and the brackish lakes of Chiprana. The semi-arid region of Los Monegros (south-central part of study area) contains a dozen small, shallow saline lakes which are regulated by high evapotranspiration, low rainfall, and ground water flow (Lecina *et al.* 2010).



Figure 3.1. (a) The map of Spain showing the two Landsat tiles selected as study areas. They cover different climate zones as mapped by Peel et al. (2007). Detailed view of the Landsat tile and sample locations for P199R031 (b) and P202R034 (c).

The second area is located in Andalusia, southern Spain. This area has a Mediterranean climate (Figure 3.1c), which is characterized by dry, hot summers and mild, moist winters with irregular precipitation, most of it occurring from October to February. The mean annual temperature and precipitation are approximately 17 °C and 600 mm, respectively, based on the WorldClim database. The main river of this region is the Guadalquivir. This area comprises extensive estuaries and wetlands, such as the marshes of Doñana in the floodplain of the lower Guadalquivir River and the Odiel marshes, both areas were declared as Ramsar sites (Matthews 1993) and they are important for the conservation of migrant water birds (Rendón *et al.* 2008).

3.3 Data

3.3.1 Landsat data

To generate training and validation data on the response variable as input for modeling surface water fraction, we used a binary water mask derived from the C-version of Function of Mask (CFmask) band contained in the Surface Reflectance (SR) higher-level data products that are generated by the United States Geological Survey (USGS). We note that in the recently-released version of the SR product (*i.e.*, June 2017), the CFmask results are integrated into "Level-2 Pixel Quality Assurance band", but we used the CFmask band available prior to that release. The CFmask algorithm assesses for each pixel if clouds, cloud shadows, snow, or surface water are present based on atmospherically corrected Landsat 4–7 and Landsat 8 images. Water was identified by

applying fixed threshold values to NIR reflectance and NDVI values (Zhu *et al.* 2015; Zhu and Woodcock 2012). We downloaded all available CFmask bands for the period 2000 to 2015 from the USGS EarthExplorer (https://earthexplorer.usgs.gov/), and retained those scenes for which the cloud cover was less than 10 percent. This resulted in 50 scenes for T_1 and 64 scenes for T_2 (Figure 3.2).

3.3.2 MODIS data

The MODIS Nadir BRDF-Adjusted Reflectance (NBAR) 16-day composite product (MCD43A4, Collection 5) was used to generate the main predictor variables for the estimation of the surface water fraction. This product provides 500 m resolution surface reflectance data for each of the MODIS bands (1–7) corrected to nadir-view geometry at a constant solar angle using a bi-directional reflectance distribution function (BRDF) model (Schaaf *et al.* 2002). We downloaded all MCD43A4 data for the 2000–2015 time period for three MODIS tiles (*i.e.*, H17V04, H17V05 and H18V04) that together cover our study areas from the Land Processing Distributed Active Archive Center (LP DAAC).



Figure 3.2. Number of Landsat images per tile and month between 2000 and 2015.

3.3.3 Google Earth imagery

We visually interpreted very high resolution imagery freely available in Google Earth for validation of the water presence detected by the CFmask algorithm. Specifically, we used multi-date historical imagery from acquisition dates near in time (< 1 month) from the Landsat imagery described in Section 3.3.1. Most of the high resolution imagery in our study area is from DigitalGlobe's QuickBird (61 cm spatial resolution) and WorldView

satellites (31–46 cm spatial resolution) (www.digitalglobe.com). The availability of historical imagery varies from place to place. On average, two images per sample location for T_1 and five images per sample location for T_2 were available from 2000 to 2015.

3.3.4 Reference data

We applied the global 30 m resolution inland water body dataset (GIW) as a reference for selecting sample locations. This binary surface-water dataset was derived from large collections of 30 m cloud-free Landsat data acquired during or near 2000, with commission errors less than 4% and omission errors less than 14% (Feng *et al.* 2015). We downloaded the GIW dataset for two Landsat tiles (*i.e.*, T_1 and T_2) from the GLCF website (http://www.landcover.org), which corresponded to 8 August 2000 for T_1 and 10 April 2001 for T_2 .

3.3.5 Terrain data

We used the Shuttle Radar Topography Mission (SRTM) elevation data acquired from the Global Land Cover Facility (http://www.landcover.org/data/srtm/). This dataset was acquired by the Space Shuttle Endeavour on mission STS-99 during February 2000 and provides near global coverage at 30 m resolution.

3.4 Methods

Our method for deriving surface water fraction consisted of five steps as described in the workflow diagram below (Figure 3.3).



3.4.1 Selection of sample locations

Binary surface-water maps from the GIW dataset were used for selecting sample locations. The binary maps were aggregated to the 500 m MODIS resolution. The aggregated value per cell is its surface water fraction, which was calculated by dividing the 30 m water pixels by the total number of Landsat-scale pixels that fall inside each MODIS grid cell. This resulted in an initial map of surface water fraction indicating the percentage of surface water within each MODIS grid cell. Rather than directly using this fraction map for training and validation, we used this map as a basis to select sample locations as this dataset is based on one single date rather than a dynamic dataset with multi-temporal layers. To achieve this, we divided the map into seven strata (0%, 0%–20%, 20%–40%, 40%–60%, 60%–80%, 80%–100% and 100%), and randomly selected 20 grid cells from each of the seven strata. This resulted in a set of 140 MODIS-scale reference grid cells per Landsat image tile (see Figure 3.1), which were further split into 70 training and 70 validation locations using random sampling from each strata.

3.4.2 Training and validation dataset

We used the time series of the CFmask bands for generating training and validation data as input to Step 2. As a first step, we assessed if the water mask of the CFmask bands are of good quality by comparing them with very-high resolution imagery. Specifically, we randomly selected five Landsat pixels falling within the selected 140 (per tile) MODIS reference cells. The selected Landsat pixels were overlaid on Google Earth and were visually classified into water and non-water. This was repeated for all very-high resolution images available in Google Earth that were acquired within one month of an existing Landsat image. A small number of pixels (31) from T_1 and T_2 were excluded in the accuracy assessment because we had recorded a lower confidence in our interpretation due to complex site conditions or cloudy/hazy conditions of the available imagery in Google Earth. The assessment results (Table 3.1) indicated that the Landsat CFmask resulted in an accurate identification of water presence and could thus serve as an accurate source of reference data.

Table 3.1. Accuracy assessment of the Landsat water/non-water classification data used for model training and validation.

Landsat tiles	Correct number	Omission	Commission
T1	1314 (97.2%)	31 (2.3%)	7 (0.5%)
T ₂	2904 (92.9%)	183 (5.9%)	37 (1.2%)

The CFmask bands were transferred to land/water binary images and were further aggregated to the scale of the MODIS grid cells (using the same method described in Section 3.4.1). This resulted in layers of per-cell surface water fraction estimates. Per-cell 500 m surface water fraction estimates derived from all Landsat acquisitions (2000–2015) for training locations were used as the training dataset, and the entire set of Landsat-derived estimates for validation locations were used as the validation dataset. Furthermore, poor quality observations containing clouds or cloud shadows as detected by CFmask were removed. This resulted in 3463 samples for training and 3245 samples for validation for study area T_1 , and 4352 samples for training and 4165 samples for validation for T_2 .

3.4.3 Predictor variables

Besides the individual MODIS bands, we used various band combinations and also an ancillary non MODIS-based topographic index as predictor variables to be evaluated (Table 3.2). For each Landsat training sample/date, the MODIS temporal composite was selected that contained the date of Landsat acquisition. The surface reflectance for individual MODIS bands, as well as seven MODIS-derived spectral indices have been widely used to characterize surface water and wetlands (Li et al. 2015; Reschke and Huttich 2014). Pekel et al. (2014) found that water and other land-cover types can be better discriminated using the Hue-Saturation-Value (HSV) color space applied on SWIR, NIR, and Red reflectance, notwithstanding the great variability in water spectral signatures. We thus included the HSV transformation results as potential predictors here as well. In addition, a single ancillary variable was considered, i.e., the Topographic Wetness index (TWI). The TWI has higher values for sites lower in the catchment and thus has been used to predict likely locations for water bodies and wetlands (e.g., Drake et al. 2015; Grabs et al. 2009). We calculated TWI from 30 m elevation data derived from the SRTM using the formula in Table 3.2; flow accumulation serves as an input to that formula and was calculated as the number of cells that flow into the downslope cell according to Jenson and Domingue (1988), and slope is derived from the maximum

elevation change between each cell and its eight direct neighbours. We subsequently aggregated the result to the 500 m resolution MODIS cells by calculating the mean value within each cell.

Temporal metrics summarize spectral band or index temporal characteristics over a defined period. Their usefulness for fractional cover mapping has been demonstrated in several studies (Reschke and Huttich 2014; Weiss and Crabtree 2011). From the time series of each MODIS-derived predictor variable, we calculated the annual mean, minimum, maximum, standard deviation, and coefficient of variation (CV) per calendar year. We subsequently used that as input in our models when assessing surface water fraction for any 16-day composite of that same year.

Table 3.2. Overview of the MODIS-derived predictor variables used in SRTM DEM is TWI. MODIS shortwave infrared bands are referred t (2105–2155 nm).	his study. The only variable not derived fron 5 as SWIR1 (1230–1250 nm), SWIR2(1628	n MODIS but from the -1652 nm) and SWIR ₃
Predictor variable	Formula	Reference
MODIS individual bands (Red, NIR, Blue, Green, SWIR ₁ , SWIR ₂ , SWIR ₃)	Ι	
NDVI – Normalized Difference Vegetation Index	(NIR - Red)/(NIR + Red)	Tucker (1979)
NDWI – Normalized Difference Water Index	(Green - NIR)/(Green + NIR)	McFeeters (1996)
MNDWI – Modified Normalized Difference Water Index	$(Green - SWIR_2)/(Green + SWIR_2)$	Xu (2006)
NDWI – Normalized Difference Water Index (referred as LSWI _{B5})	$(NIR - SWIR_1)/(NIR + SWIR_1)$	Gao (1996)
LSWI – Land Surface Water Index (referred as LSWI _{B6})	$(NIR-SWIR_2)/(NIR+SWIR_2)$	Xiao <i>et al.</i> (2002a)
TCWI – Tasseled Cap Wetness Index	$\begin{array}{l} 0.10839 \times Red + 0.0912 \times NIR + 0.5065 \times \\ Blue + 0.404 \times Green + 0.241 \times SWIR_1 - \\ 0.4658 \times SWIR_2 - 0.5306 \times SWIR_3 \end{array}$	Zhang <i>et al.</i> (2002)
TCBI – Tasseled Cap Brightness Index	$\begin{array}{l} 0.3956 \times Red + 0.4718 \times NIR + 0.3354 \times Blu \\ + 0.3834 \times Green + 0.3946 \times SWIR_1 + 0.3434 \\ \times SWIR_2 + 0.2964 \times SWIR_3 \end{array}$	e Zhang <i>et al.</i> (2002) 4

60
(Continued) Table 3.2. Overview	of the MODIS-derived predictor variables used in this study.	
Predictor variable	Formula	Reference
Value (HSV)	max(SWIR2, NIR, Red)	Pekel <i>et al.</i> (2014)
Saturation (HSV)	$1-\min(\mathrm{SWIR}_2,\mathrm{NIR},\mathrm{Red})/\max(\mathrm{SWIR}_2,\mathrm{NIR},\mathrm{Red})$	Pekel <i>et al.</i> (2014)
Hue (HSV)	$\begin{array}{l} 0, \text{if } V = \min(\text{SWIR}_2, \text{NIR}, \text{Red}) \\ \hline & \left(\frac{60^\circ \times (\text{NIR} - \text{Red})}{(V - \min(\text{SWIR}_2, \text{NIR}, \text{Red})} + 360^\circ \right) \mod 360^\circ, \text{if } V = \text{SWIR}_2 \\ \hline & \left(\frac{60^\circ \times (\text{Red} - \text{SWIR}_2)}{V - \min(\text{SWIR}_2, \text{NIR}, \text{Red})} + 120^\circ, \text{if } V = \text{NIR} \\ \hline & \frac{00^\circ \times (\text{SWIR}_2 - \text{NIR})}{V - \min(\text{SWIR}_2, \text{NIR}, \text{Red})} + 240^\circ, \text{if } V = \text{Red} \end{array}$	Pekel <i>et al.</i> (2014)
TWI – Topographic Wetness index	$\ln(\alpha/\tan\beta)$; α is upslope area per unit contour length (m) which calculated as (flow accumulation + 1) × (cell size); β is the slope	Beven and Kirkby (1979)
	expressed in radians	

3.4.4 Rule-based regression tree models

The regression-tree (RT) algorithm produces a rule-based model for predicting a single continuous response variable from one or more explanatory variables. A regression tree is built through a process known as recursive partitioning, which is an iterative process of splitting the data into homogeneous subsets. At each split or node, a splitting criterion is used to determine which explanatory variable is most suited to divide the training data. In this study, we applied two popular machine learning approaches, *i.e.*, Random Forest (RF) and Cubist, to examine the relationship between surface water fraction and predictor variables, and to identify the relative importance of the variables to predict the surface water fraction. Previous studies have demonstrated that these two approaches can obtain a high accuracy for deriving fractional cover of a single specific land surface component (*e.g.*, Huang *et al.* 2016b; Walton 2008; Xia *et al.* 2017).

RF is a classification and regression technique introduced by Breiman (2001). It is based on the Classification and Regression Trees (CART; Breiman *et al.* 1984), which are a series of binary rule-based decisions that dictate how an input is related to its predictor variables. In contrast to CART many independent regression trees are generated in RF instead of one. Each tree is grown with a random subset of predictor variables at each node and with a random selection of a bootstrapped sample from the original training dataset with the aim of avoiding overfitting (Breiman 2001). About one-third of the observations in the training dataset are not used to grow the tree. These observations are considered the out-of-bag (OOB) observations and can be used to estimate variable importance. The final prediction for a single estimate is the average of the predictions of all trees, which generally produce more accurate results for the RT implementation when compared to the CART approach (Breiman 2001).

Cubist is a tool for generating rule-based regression trees developed by Quinlan (1993). A rule-based model developed by Cubist contains a set of conditional rules that partition the data space into smaller regions, each of which is linked to a multivariate linear regression model predicting the dependent from the predictor variables. This differs from RF, which contains only a single value at each terminal node. To optimize accuracy and reduce instability of the model prediction, the Cubist model uses a boosting-like procedure called "committees" where a series of model trees are developed in sequence. Each member of the committee predicts the target value for a case and the members' predictions are averaged to give a final prediction (Quinlan 1993). Another innovation in Cubist is the use of nearest-neighbors to adjust the predictions from the rule-based model.

	1			
Model		Training	Number of	Inclusion of
		data	samples	temporal variables
	$RF_localT_1_single$	T_1	3463	No
	$Cubist_localT_1_single$	T_1	3463	No
	$RF_localT_1_temporal$	T_1	3463	Yes
Local	$Cubist_localT_1_temporal$	T_1	3463	Yes
models	$RF_localT_2_single$	T ₂	4352	No
	$Cubist_localT_2_single$	T ₂	4352	No
	$RF_localT_2_temporal$	T ₂	4352	Yes
	Cubist_localT2_temporal	T ₂	4352	Yes
	RF_global_single	T_1 and T_2	7815	No
Global	Cubist_global_single	T_1 and T_2	7815	No
models	RF_global_temporal	T_1 and T_2	7815	Yes
	Cubist_global_temporal	T_1 and T_2	7815	Yes

Table 3.3. Descriptions of models, including the training dataset and temporal variables used to build each model type.

After an initial model tree is created and a sample is predicted by this model tree, Cubist can find its nearest neighbors, *i.e.*, the most similar samples in terms of predictor variable values, and determine the average of these training set points (Kuhn *et al.* 2016; Quinlan 1993). We tuned the models over different values of "committees" and "neighbors" ("committees" was set to be 0, 10, 20, 50, 100, and "neighbors" was set to be 0, 1, 5, 9) through a 10-fold cross-validation on our training data and selected the optimal values which produced the best performance in terms of RMSE.

Twelve models were built and tested to assess how accurate the RF and Cubist regression tree algorithms could estimate surface water fraction (see Table 3.3). All models were built using the randomForest (Liaw and Wiener 2002) and Cubist (Kuhn *et al.* 2016) packages that are implemented in the R software (R Core Team 2013). The twelve models tested the effects of two factors, including:

 the need for fine-tuning the model to local conditions: global models are based on the combined training dataset from both Landsat tiles whereas local models are based on training data from a single Landsat tile; (2) the inclusion of annual variables as predictors that summarize temporal characteristics of the single-date predictors within the year: single date models only use variables obtained from the MODIS composite corresponding to the same acquisition date as the Landsat image, whereas temporal models incorporate annual summaries of the temporal characteristics of the predictor variables.

3.4.5 Model accuracy assessment, sensitivity analysis, and model prediction

The twelve models were evaluated on both the training data that were used to generate the models, and the validation data from each Landsat tile. Three statistical measures were used to assess model performance: the coefficient of determination (R^2), mean absolute error (MAE) and root mean square error (RMSE). The robustness and transferability of the local models were explored by applying the model developed for T_1 to T_2 and vice versa, and evaluating the results against the same validation dataset.

To assess the relative importance of the input variables in the two best-performing models, a different approach was followed for RF and Cubist models. RF calculates the variable importance as increased percentage of mean squared error (%IncMSE) that arises from randomly permuting the values of a single variable in the OOB data. For each tree, the prediction error (*i.e.*, MSE) on the OOB portion of the data is recorded. Then the same is done after permuting the specific predictor variable. The differences between the two are subsequently averaged for all trees and normalized by the standard deviation of the differences (Liaw and Wiener 2002). Cubist summarizes the individual variable contribution to the model by assessing: (1) the percentage of rule conditions for which the variable is used as input, and (2) the percentage of times for which the variable is used in the multivariate linear model related to the conditional rule (Kuhn *et al.* 2016).

To understand if particular surface water fractions were modelled more accurately than others, we assessed model accuracy according to different water fraction categories in 20% increments by calculating the distribution of signed deviation (SD) for the two bestperforming models. Furthermore, we tested how the variability in surface water presence impacted the model's prediction performance. For this, we stratified the validation samples into two classes (*i.e.*, permanent and fluctuating water surfaces). Based on time series of Landsat CFmask, we selected the five largest and the five smallest values of surface water fraction for each sample location. A measure of relative variability (RV) in surface water fraction was calculated for each location as: $RV = (\overline{5 \ largest \ values} - \overline{5 \ smallest \ values})/\overline{5 \ largest \ values}$ (3.1)

Samples for which RV values were less than 0.2 were considered to be permanent water, RV values larger than 0.5 were regarded as fluctuating water. Samples with intermediate values (0.2–0.5) were not considered. Subsequently we extracted the R^2 , RMSE and MAE (considering only these 10 most extreme values per sample location) to assess if the extreme water fractions for each location are accurately represented by the models.

The two best-performing models were selected to produce time series surface water fraction maps for the study areas.

3.5 Results

Each of our RF models was constructed using 500 trees, and for each tree, one third of all input variables were selected randomly at each split during the model construction process. For Cubist models, we used a 20-member committee and 9-neighbor as these values produced the best performance following an iterative testing of different settings during which we found that adding more committees or neighbours had little effect on the accuracy. The performance of the twelve models is summarized in Table 3.4 and Figure 3.4. The local models show a close match with the training data (R² above 0.95 for all local models). When testing using the validation data from the same Landsat tile where models were trained, the results were comparable (R² above 0.81 for all local models). However, when applying the local temporal models developed for T₁ to T₂ and vice versa, the R² was greatly reduced (*e.g.*, the R² for the Cubist local temporal models were and the background were very different between the two areas with contrasting environments. In other words, our results indicated that a model developed for one region cannot easily be transferred for accurate mapping of surface water elsewhere.

Monitoring the Dynamics of Surface Water Fraction from MODIS Time Series



Figure 3.4. Density scatterplots of actual versus (a) RF_global_temporal model and (b) Cubist_global_temporal model predicted surface water fraction for the training (left) and validation datasets (right) from both Landsat tiles. The plots show the local density around each sample point. High density is shown in red and low density is shown in blue.

Interestingly, when combining training data for the two regions in a single global model, the accuracies were comparable to the local models (Table 3.4). For example, the Cubist_global_single model had an R² of 0.90 between T₁ reference and model-estimated surface water, with an RMSE of 11.59% and MAE of 8.18% (Cubist local model resulted in $R^2 = 0.90$, RMSE = 11.96%, MAE = 8.60%). For the T₂ reference data the accuracies were similar though more accurate than any of the single-date local models. This finding is positive in the sense that it suggests that accurate global models can be constructed as long as the training data are collected that properly represent the various environmental conditions found in the application area.

By incorporating temporal variables, both RF and Cubist global temporal models showed improvements in the accuracy measures as compared to the single date models (Table 3.4). The R² values for the RF and Cubist global temporal models were consistently larger than 0.90 for both T_1 and T_2 . This suggests that temporal variables provide important information for predicting water fraction (as also shown later in Figure 3.6).

predicted surface water	fraction	i and actual c	lata for all mode.	ls.					
Model	Traini	ng data		Valida (3245	tion data_T samples)		Valida (4165	tion data_T2 samples)	
	R^2	RMSE (%)	MAE (%)	R^2	RMSE (%) MAE (%)	R^2	RMSE (%)	MAE (%)
$RF_localT_{1_single}$	66.0	4.13	2.31	0.89	12.13	8.81	0.79	17.45	13.70
Cubist_localT1_single	0.96	7.43	4.52	0.90	11.96	8.60	0.74	19.00	14.11
$RF_localT_1_temporal$	0.99	3.46	1.65	0.91	10.39	7.78	0.65	25.87	23.49
Cubist_localT1_temporal	0.97	6.57	3.57	0.92	10.43	7.55	0.35	31.87	28.85
RF_localT_2 _single	0.98	4.12	2.46	0.85	16.77	12.13	0.83	15.12	9.97
Cubist_localT2_single	0.95	7.50	4.79	0.84	15.88	11.45	0.81	16.56	10.93
$RF_localT_2_temporal$	0.99	2.65	1.35	0.75	24.39	21.58	0.90	11.98	8.15
Cubist_localT2_temporal	0.98	5.10	3.06	0.37	29.85	25.62	0.90	11.59	7.71
RF_global_single	0.99	4.27	2.30	0.90	12.06	8.54	0.85	14.02	9.58
Cubist_global_single	0.93	9.56	6.26	0.90	11.59	8.18	0.86	13.70	8.98
RF_global_temporal	0.99	3.05	1.56	0.91	10.99	7.79	0.91	11.83	8.06
Cubist_global_temporal	0.96	7.00	4.25	0.92	10.39	7.52	0.90	11.54	7.79

67

Chapter 3



Figure 3.5. Distribution of signed deviation (SD) of predicted surface water fraction using RF (a) and Cubist (b) global temporal model by surface water categories for validation data. Positive SD values indicate model overestimations and negative values indicate model underestimations.

Figure 3.5 provides the distribution of signed deviation (SD) per surface water fraction category. Both the RF and Cubist global temporal models obtained the highest accuracy for the 0% water fraction, followed by 0%–20% water fraction, and the lowest accuracy for the 60–80% water fraction. For example, for the Cubist global temporal model (Figure 3.5b), 63% of the validation data fall within +–0.05 deviation and 86% fall within +–0.10 deviation for 0%–20% water fraction, while 52% of the predictions had an absolute difference larger than 0.10 for the 60%–80% water fraction. For the RF global temporal model the deviation distributions were similar; 65% fall within +–0.05 deviation and 84% fall within +–0.10 deviation for the 0%–20% water fraction. Figure 3.5 reveals a general tendency of the models to underestimate the larger values (*i.e.*, 60%–80% to 100%). This implies that large surface water fractions are often not accurately captured.

The relative importance of the predictor variables for estimating surface water fraction is summarized in Figure 3.6. The plots listed the 20 most important variables according to RF and Cubist global temporal models. TWI and the NIR band are the predictors of greatest importance for both models. This corroborates reports from previous studies that

have identified topographic indices as important variables for mapping surface water (Drake *et al.* 2015; Guerschmann *et al.* 2011; Margono *et al.* 2014). This is as expected because surface water is usually found at depressions in the landscape. The importance of NIR for mapping surface water fraction in flood applications was also highlighted by Sun *et al.* (2012). Surprisingly, NIR is the single non-static variable of considerable importance for mapping and monitoring surface water fraction in RF global temporal model (Figure 3.6a), which makes it the only variable for explaining the fluctuating variation of surface water fraction (*i.e.*, seasonal and intra-annual fluctuation). In the Cubist global temporal model, NIR is the single non-static variable that is important when splitting the data (Figure 3.6b), but dynamic variables are the most important in the linear models (Figure 3.6c).



Figure 3.6. Twenty predictor variables with the highest relative importance for estimation of surface water fraction. Importance is measured as (%IncMSE) with the RF_global_temporal model (a), variable usage (%) in the rule conditions (b) and in the models (c) with the Cubist global temporal model.

Table 3.5. Coefficient of determination (R^2) , root mean square error (RMSE) and mean absolute error (MAE) between predicted surface water fraction and actual data using RF and Cubist temporal model for different types of water permanence.

Model	Permanent water		Fluctuating water		water	
	R^2	RMSE	MAE	R^2	RMSE	MAE
		(%)	(%)		(%)	(%)
RF_global_temporal	0.67	13.35	10.07	 0.75	16.00	10.87
Cubist_global_temporal	0.70	12.75	9.50	0.80	14.66	9.98

Table 3.5 shows the model accuracies for permanent and fluctuating water. Because we included only a subset of the validation data (as explained in Section 3.4.5), the accuracies were lower compared to the overall accuracies. The results show that both models had slightly lower RMSEs and MAEs for permanent water as compared to fluctuating water. This indicates that the models developed are not only accurate for the static mapping of surface water fraction, but can instead be applied effectively for monitoring the dynamics of surface water fraction. This is illustrated in Figure 3.7 where we show the surface water dynamics resulting from RF and Cubist global temporal model for two small lake systems with a known strong fluctuating variability in water surface, *i.e.*, Laguna Salada de Chiprana (7a) in Ebro basin, northern Spain and Reserva Natural Complejo Endorreico Lebrija-Las Cabezas (7b) in Andalusia, southern Spain. The size of the first lake system is approximately one MODIS pixel and the second about 0.8 MODIS pixel. These two lakes are identified as Ramsar wetland sites (http://www.ramsar.org/wetland/spain). Laguna Salada de Chiprana is a permanent and deep saline lake with a maximum extent in May and minimum extent in September (De Wit et al. 2013). For Lebrija-Las Cabezas we assessed the total extent of four comprised small and shallow water bodies (i.e., Lagunas del Pilón, Cigarrera, Galiana, and Peña) that are usually flooded during December-May, but after dry out due to lack of rainfall and high evaporation rates. The surface water extent estimated from Cubist global temporal model corresponds well with the Landsat reference data from CFmask bands and from Global Surface Water data which was produced by the Joint Research Centre of the European Commission (JRC: Pekel et al. 2016) for these two lake systems. Although the seasonal pattern was appropriately represented also by the RF model, it resulted in large overestimations of very small surface water extent, as shown in Figure 3.7b and a consequently poor correspondence with surface water extent estimates from Landsat. Even if a more detailed

validation of these results with *in situ* data would be desirable (Li *et al.* 2015), Figure 3.7 illustrates that the detailed information on seasonal, inter-annual and long-term changes in surface water extent of these small lakes are not fully captured by Landsat due to significant temporal data gaps. This detail was captured with our approach using 500 m resolution MODIS imagery, with a better performance when using the Cubist method.



Figure 3.7. Time series of surface water extent for two Ramsar sites used as examples of the results from the Cubist (blue dots) and RF (red dots) global temporal model for the period from 2000 to 2015. (a) Laguna Salada de Chiprana (b) Reserva Natural Complejo Endorreico Lebrija-Las Cabezas. The size of the first site is approximately one MODIS pixel and the second about 0.8 MODIS pixel. Yellow and green asterisks represent Landsat reference data from CFmask bands and from the JRC Global Surface Water product, respectively.

3.6 Discussion

The rule-based regression tree models using RF and Cubist provided accurate estimation of the surface water fraction ($R^2 > 0.9$, RMSE < 11.5%). This can be attributed to the effective partitioning of the data into homogeneous subsets, which is a principal characteristic of the applied models (Huang and Townshend 2003). Our reported accuracies compare favorably to previous studies that aimed at estimating surface water fraction from medium-resolution imagery (MODIS or similar). For instance, in a study

of Yukon Flats National Wildlife Refuge, Weiss and Crabtree (2011) reported an R² of 0.625 between reference and modeled surface water fraction using simple linear regression model. Guerschmann et al. (2011) developed a logistic regression model for water fraction mapping on the Australian continent. An R² of 0.90 was reported for the training data but a lower R² of 0.70 was reported for other test regions. Similar to our findings, other studies also found that RF and Cubist often overestimate small values and underestimate large values when estimating fractional cover of land surface components such as surface water fraction (Huang et al. 2014b) and in other applications (Wang et al. 2017). One explanation for this behavior could be that RF averages the target values with each terminal nodes to give a final prediction, and the averaging reduces the range of surface water fraction in the final outcome. As a consequence, RF cannot predict beyond the range of response values in the training data, nor predict the entire range of response values. Cubist has a better performance as it predicts the terminal nodes using a linear model as opposed to the simple average in RF. Prediction biases commonly result from statistical methods including RF (Huang et al. 2014b), due to highly unbalanced distributions of the percentage cover of those land cover components in real world like impervious surfaces and water bodies, which typically account for small portions of the total area. Further improvement in Cubist model could be achieved by setting a "extrapolation" parameter which controls to what extent predictions are adjusted to be consistent with the range observed in the training set (Kuhn et al. 2016).

Training data collection is crucial for effective application of machine learning algorithms like RF and Cubist. The applicability of empirical approaches remains often confined to relatively small study areas, while our study shows that RF and Cubist approaches developed for one site could not be accurately transferred to the other site. Instead, by collating more training data from the two sites we were able to construct comparable models to local models. This is probably because machine learning approaches like Cubist can effectively take care of tuning a global model locally, as long as it is fed with sufficient good-quality training data that comprise the various environmental conditions encountered across the larger area. In this sense, we are confident that our approach for constructing global models can be expanded effectively to wider areas, potentially resulting in an accurate method for global monitoring of water fraction from MODIS or similar imagery.

For this to happen, training data are required across the globe. In this study we achieved this with CFmask, which is globally available for all collected Landsat images from the USGS. For this study we found a high accuracy (>92%) when comparing against very

high resolution imagery from Google Earth. In addition, an alternative global temporal dataset at 30 m resolution on water presence is now available that also uses the entire Landsat archive (Pekel *et al.* 2016). The authors reported a high accuracy (commission accuracy of 99.45% and omission accuracy of 97.01%) when a validation was performed by using high-resolution satellite imagery and aerial photography spanning across geographic regions (global) and across different time periods. This dataset could provide an improvement possibly resulting in a better training of the regression tree models. While it may be a large step to move from a two Landsat-tile model to a truly global model, intermediate upscaling test could be performed to large eco-regions, such as the Mediterranean.

Other adaptations that may further improve the accuracy of surface water fraction monitoring include the selection of imagery, and the further exploration of alternative machine learning approaches. In this study we used MODIS Collection 5 data. This was the latest version of the MCD43A4 when starting the study, but now Collection 6 is available with quality improvements (Schaaf and Wang. 2015). In addition, our approach may also work well with non-MODIS datasets like Proba-V and Sentinel-3, even if their archive starts later. Research comparing algorithms have shown that regression trees are often among the top performers across a range of applications on fractional surface cover mapping (Drzewiecki 2016; Xia *et al.* 2017), whereas alternative machine learning algorithms such as multivariate adaptive regression splines and artificial neural networks may hold great potential and would require further testing (Weng 2012).

This paper provides an important step towards large-scale mapping of surface water/wetland extent at high temporal resolution and at a relatively high spatial resolution by depicting each pixel as a percent coverage of surface water. Given the good accuracies obtained here when incorporating dissimilar study sites in a single model, our approach has potential for application at global scale. Apart from successfully mapping the locations and extents of permanent water, the proposed methodology allows for monitoring the temporal changes of fluctuating water surfaces. This is meaningful for a variety of potential applications of the proposed method, such as mapping temporary and ephemeral surface water bodies, monitoring extreme events such as floods and drought , and estimating global loss of wetland extent by including the loss of small-size and temporary wetlands, which are likely to contribute much more to the decline of wetland biodiversity than the loss of an equivalent area in larger wetlands (Deane *et al.* 2017; Semlitsch and Bodie 1998). We note that the application in wetlands with inundated vegetation or with floating vegetation may remain challenging, particularly because these

(parts of) wetlands will likely be omitted in Landsat-derived training datasets. Further research is needed into how such conditions may affect model accuracy or require model adaptations, even though we expect that the moisture presence will effectively be captured in MODIS-derived indices. Relevant spatial information that could be extracted from temporal surface water estimates include 1) dates of water appearance and disappearance within a given year, 2) the period during which water is detected, 3) seasonal and inter-annual variability of surface water extent for ephemeral wetlands, and 4) the timing, frequency, and extent of events such as floods and drought, lake expansion and retreat or river-channel migration. Such information can provide insights for understanding how climate change and human actions affect the dynamics of surface water, and has potential to greatly contribute to the management and conservation of biodiversity and other ecosystem services associated with terrestrial surface water and wetlands.

3.7 Conclusions

In this study, we evaluated two ruled-based regression tree methods, RF and Cubist, for assessing surface water fraction from MODIS imagery and a DEM for two contrasting environments in Spain. Both models achieved high accuracies ($R^2 > 0.91$, RMSE < 11.47%) when a single model was trained using data from both study areas. A prerequisite for the method to be effective is to build a reliable training dataset, which is usually the most time consuming element for applying machine learning approaches. The time series of Landsat-derived CFMask water fraction used in this paper proved to be an efficient basis for building such a training set, and showed a high accuracy versus high-resolution imagery from Google Earth. We also showed that tuning model parameters for each study area is not required, but that a single global model can provide accurate results, as long as the training data comprise the different environmental conditions across the spatial domain for which the model will be applied.

Our study complements existing studies on large-scale and global surface water monitoring from MODIS data that focused on binary land/water classifications. Contrary to most of those studies, we incorporate small water bodies (*i.e.*, less than one MODIS pixel size) in the analysis by estimating surface water fraction. In this way, we could derive 16-day interval surface water fraction maps for 16 years of MODIS data at 500 m pixel resolution. This long-term mapping with a high temporal density also complements existing high-resolution Landsat-based studies especially for areas that suffer from longer data gaps and thus cannot provide detailed information on seasonal, inter-annual and

long-term changes. We expect that our approach and resulting maps could serve as a tool for (a) operational surface water monitoring especially in arid and semi-arid areas; (b) updating and refining of existing wetland and habitat databases; and (c) evaluating ecosystem services and biodiversity assessments.

Monitoring the Dynamics of Surface Water Fraction from MODIS Time Series

Chapter 4

A New Dense 18-Year Time Series of Surface Water Fraction Estimates from MODIS for the Mediterranean Region^{*}

^{*} This chapter is based on: Li, L., Skidmore, A., Vrieling, A., & Wang, T. (2019). A new dense 18year time series of surface water fraction estimates from MODIS for the Mediterranean region. *Hydrology and Earth System Sciences*, 23, 3037–3056.

Abstract

Detailed knowledge on surface water distribution and its changes is of high importance for water management and biodiversity conservation. Landsat-based assessments of surface water, such as the Global Surface Water (GSW) dataset developed by the European Commission Joint Research Centre (JRC), may not capture important changes in surface water during months with considerable cloud cover. This results in large temporal gaps in the Landsat record that prevent the accurate assessment of surface water dynamics. Here we show that the frequent global acquisitions by the Moderate Resolution Imaging Spectrometer (MODIS) sensors can compensate for this shortcoming, and in addition allow for the examination of surface water changes at fine temporal resolution. To account for water bodies smaller than a MODIS cell, we developed a global rule-based regression model for estimating the surface water fraction from a 500 m nadir reflectance product from MODIS (MCD43A4). The model was trained and evaluated with the GSW monthly water history dataset. A high estimation accuracy ($R^2 = 0.91$, RMSE = 11.41 %, and MAE = 6.39 %) was achieved. We then applied the algorithm to 18 years of MODIS data (2000-2017) to generate a time series of surface water fraction maps at an 8 d interval for the Mediterranean. From these maps we derived metrics including the mean annual maximum, the standard deviation, and the seasonality of surface water. The dynamic surface water extent estimates from MODIS were compared with the results from GSW and water level data measured in situ or by satellite altimetry, yielding similar temporal patterns. Our dataset complements surface water products at a fine spatial resolution by adding more temporal detail, which permits the effective monitoring and assessment of the seasonal, inter-annual, and long-term variability of water resources, inclusive of small water bodies.

4.1 Introduction

Terrestrial surface water bodies such as lakes, reservoirs, and rivers cover approximately 3 % of the global land mass. They play a crucial role in the global hydrological cycle, biodiversity conservation, and climate process (Chahine 1992; Tranvik *et al.* 2009). Detailed knowledge on surface water distribution, and its seasonal, inter-annual, and long-term variability can serve as an important source for water management (Cole *et al.* 2007), ecosystem assessment, and biodiversity conservation (Turak *et al.* 2017). Remotesensing data have increasingly been used to monitor surface water changes, and powerful methods and tools have been developed for analyzing Earth observation data. However, existing approaches for monitoring the surface water extent are limited either in geographic scope, temporal extent of the record, or with respect to the temporal frequency of observations.

At the global scale, several static datasets exist that provide information on the spatial extent of water bodies and wetlands. For example, the Global Lakes and Wetlands Database (GLWD: Lehner and Doll 2004) was based on historical maps and has a spatial resolution of 30 arcsec (approx. 1 km). Carroll *et al.* (2009) combined the Shuttle Radar Topography Missions (SRTM) Water Body Data (SWBD) with 250 m Moderate Resolution Imaging Spectrometer (MODIS) reflectance data to produce a global static map of surface water for circa 2000–2002. Global Landsat-based static surface water datasets include the 3 arcsec (~ 90 m) Water Body Map (G3WBM: Yamazaki *et al.* 2015) and the Global Land Cover Facility (GLCF) inland surface water dataset at a 30 m resolution for 2000 (Feng *et al.* 2015).

Even though static water maps are adequate for some applications there is an increasing demand for information on the spatiotemporal variability of inland water bodies and their long-term evolution (Belward 2016). Dynamic mapping and monitoring of the surface water extent have been explored using optical sensors featuring fine (10–30 m) to medium (250–500 m) spatial resolutions. At fine spatial resolution, several studies have recently presented interesting results on long-term variability of surface water with the entire Landsat archive at regional (Halabisky *et al.* 2016; Heimhuber *et al.* 2016), continental (Mueller *et al.* 2016), and global scales (Donchyts *et al.* 2016; Pekel *et al.* 2016). The European Commission Joint Research Centre's (JRC) Global Surface Water (GSW) dataset (Pekel *et al.* 2016) quantifies changes in global surface water over the past 32 years with a monthly time interval. This product allows for the analysis of surface water dynamics over long time periods at fine spatial resolution, but only provides information

on monthly changes in surface water. Moreover, the Landsat archive also contains data gaps and temporal discontinuities depending on the geographical location (Pekel *et al.* 2016). This is due to both the limited number of acquisitions during specific time intervals, and the location- and time-dependent persistency of cloud cover. These data gaps affect the accuracy of the seasonality information (Yamazaki and Trigg 2016). To better represent water bodies with short hydroperiods and short-duration flooding, it is critical to account for such gaps when monitoring surface water. In recent years, the revisit time of fine-resolution sensors has increased (*e.g.*, Sentinel-2 has offered a 5 d repeat since March 2017: Du *et al.* 2016). However, these data cannot yet be used to create long-term (> 10 year) consistent time series at short time intervals.

Moderate resolution imagery derived from satellite sensors such as MODIS provides daily observations over long time-spans and as such has the potential to construct longterm and dense time series of surface water over large regions. Many studies have explored the use of MODIS in mapping water body dynamics at regional to continental scales (Kaptue et al. 2013; Pekel et al. 2014; Sharma et al. 2015) using binary classification methods. At the global scale, Khandelwal et al. (2017) used MODIS multispectral data to map the global extent and temporal variations of 94 large reservoirs at a 500 m resolution and at an 8 d interval from 2000 to 2015. The recent Global Climate Observing System (GCOS) report states that essential climate variables (ECVs) need to be established for water extent and lake ice cover products, ideally with daily temporal resolution (Belward 2016). To address this requirement, the first daily global dataset of inland water bodies at a 250 m spatial resolution from 2013 to 2015 was developed by (Klein et al. 2017). This work advanced surface water mapping using remote sensing, due to its dense temporal resolution, and enhanced our understanding of rapid water changes caused by extreme climate change and human activities. However, like other surface water mapping efforts based on binary classification methods (e.g., Khandelwal et al. 2017; Mohammadi et al. 2017), this product omits lakes and narrow rivers that only cover a portion of a MODIS resolution cell.

To overcome this limitation and incorporate small water bodies, several researchers have attempted to predict sub-pixel surface water estimates of MODIS by providing the water fraction in each pixel using techniques such as linear spectral mixture modeling (*e.g.*, Hope *et al.* 1999; Li *et al.* 2013b; Olthof *et al.* 2015) and machine learning (*e.g.*, Li *et al.* 2018a; Rover *et al.* 2010; Sun *et al.* 2012) for small areas. However, the utility and efficiency of these methods have rarely been explored for the estimation of the surface water fraction for larger areas. In our previous work (Li *et al.* 2018a), we explored the

use of rule-based regression models over two small areas on the Iberian Peninsula and concluded that a single global regression model can provide accurate surface water estimates across areas with different environmental conditions as long as it is fed with training data that comprise these various conditions. Consequently, we concluded that this approach has the potential to be applied over much larger areas. Therefore, the aim of this paper is to explore the utility and efficiency of a rule-based regression model for the estimation of the surface water fraction for the Mediterranean region, and to develop a new surface water fraction dataset for the Mediterranean region using fine temporal resolution MODIS data as input for the effective assessment of seasonal, intra-annual, and long-term surface water dynamics inclusive of small water bodies. Our specific objectives are as follows:

- (1) to develop an approach for the estimation of the surface water fraction for the Mediterranean region at a fine temporal resolution from MODIS data;
- (2) to generate an 8 d interval time series of surface water fraction maps for the Mediterranean from 2000 to 2017, and to use that to derive a series of ecologically relevant metrics;
- (3) to compare our dataset with an existing dataset (*i.e.*, JRC's GSW) and water level data to assess how they compare in space and time.

4.2 Study area

We loosely defined the Mediterranean in this study as the region that is contained within 10 MODIS grid tiles, which together cover all coastal areas of the Mediterranean and the Black Sea, including a significant portion of their inland areas (Figure 4.1). This boundary is defined based on the combination of (1) the definition of the Mediterranean region by the Mediterranean Wetland Observatory (MWO) project, *i.e.*, 27 Mediterranean countries are included by MWO; (2) the inclusion of areas with a large amount of Ramsar wetlands; and (3) the exclusion of southern parts of north African countries (*i.e.*, Morocco, Algeria, Libya, and Egypt) that comprise few water bodies according to the maximum water extent over 32 years from JRC's GSW product.





Figure 4.1. Study area and sample locations.

The study region covers 13 climate zones as defined by Peel *et al.* (2007) (Figure 4.1). Numerous water bodies of different types are found in this region, including large coastal lagoons, fresh, brackish or salt marshes, riverine forests and reed beds, flood plains and wet meadows, mountainous lakes and surrounding wetlands, salted lakes, temporary marshes, and streams (Costa *et al.* 1996). Our study area accounts for 25 % of the world's Ramsar sites that contain a great ecological, social, and economic value, especially as they provide habitat, reproduction, and migration stopover sites for numerous bird species (Galewski 2012). A good number of the water bodies and wetlands in the region are small, shallow, and highly variable between seasons and years due to weather effects and human activities (Costa *et al.* 1996).

Many Mediterranean water resources are degraded mainly due to urbanization, agricultural reclamation, increasing water use for irrigation, and hydraulic works such as dams, dikes, river channeling, and drainage and irrigation networks (Batalla *et al.* 2004). A number of projects and programs have performed monitoring of surface water and wetlands in the Mediterranean region, such as MWO (http://medwet.org/, last access: 10 July 2019) and the GlobWetland initiative (http://webgis.jena-optronik.de/, last access: 10 July 2019), which highlighted the importance of protecting Mediterranean water resources. However, these projects either performed wetland mapping for a few moments in time (*e.g.*, GlobWetland only covered 1975, 1990, and 2005), or were limited to specific water bodies and wetlands instead of the whole landscape. Although surface

water dynamics in the Mediterranean can be analyzed at fine spatial resolution with JRC's GSW, it has large spatial and temporal gaps. Figure 4.2 shows the percentage of pixels with valid observations in JRC's GSW monthly water history dataset for each month between January 2000 and October 2015 calculated over the entire Mediterranean area. The figure illustrates that no valid observation exists for December in the years 2000–2015 in Mediterranean areas according to the GSW monthly water history map. In addition, less than 10 % of the Mediterranean area has observations for January.



Figure 4.2. Percentage of pixels with valid observations in JRC's Global Surface Water (GSW) monthly water history dataset for each month between January 2000 and October 2015, taken as a spatial average for the entire Mediterranean region as displayed in Figure 4.1.

4.3 Data

Table 4.1 summarizes all datasets used in this study. They include a number of sources used to derive model input variables, training data for building the model, validation data for model accuracy assessment, and other existing surface water products against which we compared our products. Details are provided in the following sections.

4.3.1 MODIS data

The main input dataset in this study is the MODIS Terra and Aqua nadir BRDF-adjusted reflectance (NBAR) product (MCD43A4, V006). This product provides 500 m resolution surface reflectance data for each of the MODIS bands (1-7) corrected to a common nadir view geometry at the local solar noon zenith angle using a bidirectional reflectance distribution function (BRDF) model (Schaaf and Wang. 2015). Compared with the previous collection (V005) that had an 8 d frequency, the V006 collection was retrieved on a daily basis. Each daily value is a result of compositing information obtained during 16 d of observations, which are weighted as a function of the quality, the observation coverage, and the temporal distance from the day of interest. Each daily V006 retrieval is the center (*i.e.*, the ninth day) of the moving 16 d input window (Schaaf and Wang. 2015). We also used the MCD43A2 (V006) Bidirectional Reflectance Distribution Function and Albedo (BRDF/Albedo) Quality dataset to filter out pixels with snow and ice in the MCD43A4 product. This dataset has the same temporal and spatial resolution as MCD43A4 (i.e., daily 500 m resolution), and contains quality information for the corresponding MCD43A4 NBAR product including snow and ice presence (Schaaf and Wang 2015).

Data/product name	Temporal	Spatial	Purpose in the study
	resolution	resolution	
MCD43A4, V006	daily	500 m	Generation of predictor variables and
(MODIS/Terra and Aqua			production of time series of water
Nadir BRDF-Adjusted			fraction maps
Reflectance)			
MCD43A2, V006	daily	500 m	Snow and ice mask
(MODIS/Terra and Aqua			
BRDF/Albedo Quality)			
GSW monthly water	monthly	30 m	Generation of training and validation
history dataset			datasets, and thematic products
GSW maximum water	static	30 m	Define sampling strata; Exclusion of
extent map			non-water samples from training
			locations
GSW water transitions	static	30 m	Define sampling strata
map			
Digital elevation model	static	90 m	Generation of predictor variables;
from Shuttle Radar			Identification of sloping terrain and
Topography Mission			terrain shadows
(SRTM) 3 v4.1			
USGS Landsat archive	16-day	30 m	Link the GSW monthly history datasets
			to a single date of cloud-free Landsat
			acquisition because the exact date of
			observation is not included in the GSW
			dataset
MCD12Q1 (MODIS Land	static	500 m	Identification of building shadows
Cover Type product)			
Land Water Mask derived	static	250 m	Comparison of products
from MODIS and SRTM			
(MOD44W)			
Water level from satellite	10-day	-	Validation of results
altimetry			
Water level from in situ	daily	-	Validation of results

 Table 4.1. Input and reference datasets used in the study.

For this study, we downloaded the daily files of MCD43A4 and MCD43A2 for 2000, 2003, 2006, 2009, 2012, and 2015. As explained in Section 4.4.1, all the available dates (or months) of the GSW monthly water history dataset in these 6 years over the sample locations were used for building training and validation data. Instead, when producing the surface water fraction time series, we collected the MCD43A4 and MCD43A2 files using an 8 d time step from February 2000 to December 2017, as processing daily files for the large study area and the 18-year time period would become too time- and memoryconsuming. The 8 d repeat coverage is considered to be a minimum for effectively capturing water bodies with short hydroperiods while simultaneously accounting for frequent cloud cover (Guerschmann et al. 2011; Wulder et al. 2016). All images were downloaded from NASA Earthdata Search website the (https://search.earthdata.nasa.gov/search, last access: 10 July 2019).

4.3.2 Global Surface Water (GSW) dataset

To generate training and validation data for modeling the surface water fraction, we used the GSW dataset (Pekel *et al.* 2016). This dataset provides the global distribution of the surface water extent at a monthly time interval from March 1984 to October 2015 (380 months) at a 30 m spatial resolution, and also includes a series of thematic maps summarizing different facets of the spatial and temporal dynamics of surface water over 32 years. This dataset is derived from the entire archive of Landsat 5 Thematic Mapper (TM), the Landsat 7 Enhanced Thematic Mapper-plus (ETM+), and the Landsat 8 Operational Land Imager (OLI). Water detection was performed using a dedicated expert system, which was a procedural sequential decision tree that used both the multispectral and multitemporal attributes of the Landsat archive as well as ancillary data layers (Pekel *et al.* 2016). Based on a validation with very high resolution satellite and aerial imagery, the authors reported a high mapping accuracy with a commission accuracy of 99.45 % and an omission accuracy of 97.01 % (Pekel *et al.* 2016).

The GSW monthly water history dataset is available in Google Earth Engine (Gorelick *et al.* 2017) as an image collection containing 380 images, one for each month between March 1984 and October 2015. Each image provides a binary classification of water presence, or indicates if no valid (cloud-free) Landsat observations were available for a specific pixel and month. For comparison with MODIS data, we used the monthly water history datasets between February 2000 and October 2015, which resulted in 189 images.

Several GSW thematic maps were derived from the GSW monthly water history dataset (Pekel *et al.* 2016). In this study, we used two thematic maps, the maximum water extent

map and the water transitions map, for the Mediterranean region from the data access website (https://global-surface-water.appspot.com/download, last access: 10 July 2019). The maximum water extent map indicates whether each 30 m grid cell was ever detected as water over the 32-year period. The transition map contains 10 water classes: permanent, new permanent, lost permanent, seasonal, new seasonal, lost seasonal, seasonal to permanent, permanent to seasonal, ephemeral permanent, and ephemeral seasonal. It provides information on both intra- and inter-annual variability of surface water (Pekel *et al.* 2016).

4.3.3 Terrain data

Terrain data are useful for predicting the locations for water bodies (Drake *et al.* 2015; Grabs *et al.* 2009). In this study, we used the near-global Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) distributed by the Consortium for Spatial Information of the Consultative Group of International Agricultural Research (CGIAR-CSI). This product has a \sim 90 m resolution and is a post-processed derivative to address areas of missing data in the original SRTM DEM made by the National Aeronautics and Space Administration (Jarvis *et al.* 2008). The most recent version of this product is SRTM3 v4.1 and is freely available from http://srtm.csi.cgiar.org/ (last access: 10 July 2019).

4.3.4 Satellite altimetry and *in situ* water level

We obtained water levels from the U.S. Department of Agriculture Global Reservoir and Lake Monitoring (GRLM) website (http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir, last access: 10 July 2019). This site provides time series of water level variations for some of the world's largest lakes and reservoirs, mainly greater than 100 km². The GRLM utilizes near-real time data from the Jason-3 mission, and archive data from the Jason-2/OSTM, Jason-1, Topex/Poseidon, and ENVISAT satellites. We also obtained daily *in situ* gauge observations for Fuente de Piedra Natural Reserve in southern Spain, which were also used in Li *et al.* (2015).

4.3.5 Additional data

We utilized the MODIS land cover type product (*i.e.*, MCD12Q1: Friedl *et al.* 2010) to identify and mask areas that potentially have commission errors related to building shadows. To assess the spatial accuracy of MODIS derived maps, we also used the land

water mask derived from MODIS 250 m and SRTM data (MOD44W: Salomon *et al.* 2004) for comparison.

4.4 Methods

4.4.1 Approach for deriving the surface water fraction

The approach used to derived the surface water fraction builds on our previous work (Li *et al.* 2018a) with considerable improvements regarding input data, training data, and commission error processing. We explored the use of MODIS spectral information and a topographic metric for estimating the surface water fraction over two study areas in Spain via the use of rule-based regression models and concluded that a single global regression model can be effectively tuned locally as long as it is fed with training data that comprise the various environmental conditions encountered across the larger area (Li *et al.* 2018a). In this sense, the approach for constructing a global model can be expanded effectively to wider areas such as the Mediterranean region. The following subsections and Figure 4.3 describe the individual steps of the approach in detail.

4.4.1.1 Selection of sample locations for training and validation

Sample locations were selected using a two-stage stratified random sampling method. First, a total of 13 strata were defined based on climate zones (see Figure 4.1). We created sampling blocks by partitioning the study area into $\sim 5 \text{ km} \times 5 \text{ km}$ grids (*i.e.*, 10×10 MODIS pixels as one block) and assigned each block to the climate zone within its spatial footprint. Blocks that covered more than one climate zone and those that contained no surface water based on the JRC maximum water extent product were excluded from our sample. We then selected 1400 blocks ($\sim 2\%$ of all resulting blocks) using stratified random sampling.



Figure 4.3. Diagram of our approach for deriving the surface water fraction.

Second, we created 500 m ×500 m grids corresponding to the MODIS geometry in each of the 1400 blocks. A total of 14 strata were defined based on the combination of water fraction categories and water permanence types. Specifically, we first divided all grid cells into seven water fraction categories (0 %, 0 %–20 %, 20 %–40 %, 40 %–60 %, 60 %–80 %, 80 %–100 %, and 100 %) according to the aggregated GSW maximum water extent. Then for each category, we further classified water permanence types based on the aggregated GSW water transitions map. For the 20 %–40 %, 40 %–60 %, 60 %–80 %, 80 %–100 %, or 100 % categories, grid cells were further classified as fluctuating water if they contained more than 20 % fluctuation water otherwise they were assigned as permanent water. For the 0 %–20 % category, grids with 0 % permanent water were classified as fluctuating water, whereas grids with 0 % fluctuation water were not assigned due

to a very low water fraction. In the end, the 14 strata were as follows: 100 % permanent, 100 % fluctuating, 80 %-100 % permanent, 80 %-100 % fluctuating, 60 %-80 % permanent, 60 %-80 % fluctuating, 40 %-60 % permanent, 40 %-60 % fluctuating, 20 %-40 % fluctuating, 0 %-20 % permanent, 0 %-20 % fluctuating, 0 %-20 % no water class, and 0 % no water class. From each of the 14 strata, we randomly selected 500 grid cells. This resulted in a set of 7000 MODIS-scale reference grid cells (shown in Figure 4.1), which were further split into 3500 training and 3500 validation locations using random sampling from each strata.

Sampling times were selected at a constant interval of 3 years (*i.e.*, 2000, 2003, 2006, 2009, 2012, and 2015). All available dates/months in the GSW monthly history datasets from those selected years were used.

4.4.1.2 Building training and validation datasets

The GSW monthly water history maps were used for generating training and validation data. Specifically, the 30 m monthly water history maps from all sampling years/months were aggregated to the 500 m resolution for all sample locations in GEE by dividing the 30 m water pixels by the total number of 30 m pixels within each 500 m resolution cell, resulting in a surface water fraction that we used as a reference. Given that the exact dates of these monthly water history maps are not provided with the GSW product, we linked these reference estimates to the USGS Landsat archive that GSW used as its input. For each combination of location/month, we retained only those reference estimates for which the location was covered by a single Landsat tile acquired during that month. If multiple Landsat tiles existed in that month for that location, we only retained the reference estimates if all but one Landsat tile had 100 % cloud cover. In this way, we could accurately assign a precise date to the retained reference estimates.

Surface water fraction estimates derived from all months of the sample years for training locations were used as the training dataset, and the estimates from all sample months for validation locations were used as the validation dataset.

4.4.1.3 Modeling surface water fraction and accuracy assessment

The surface water fraction was estimated using MODIS spectral information and derived water indices, and a topographic metric via a rule-based regression model. All predictor variables (Table 4.2) evaluated by Li *et al.* (2018a) were used as input for the estimation of surface water fraction. In addition, the annual mean, minimum, maximum, standard

deviation, and the coefficient of variation (CV) of each MODIS-derived predictor variable were also included as input in the model. These temporal summaries were demonstrated to be an important input for predicting surface water fraction in our previous study Li *et al.* (2018a).

Cubist regression models (Quinlan 1993) contain a set of conditional rules that partition the data space into smaller regions, each of which is linked to a multivariate linear regression model that can predict the explanatory variable (here surface water fraction). Following the findings of our earlier work (Li et al. 2018a), we used a single global Cubist regression model, but trained it with data collected from across the study area to tune the model to local conditions. In the global Cubist regression model, two parameters can be defined to optimize accuracy and reduce instability of the model prediction. The first is called "committees" indicating that multiple model trees are developed in sequence. Each member of the committee predicts the target value and the members' predictions are averaged to give a final prediction (Quinlan 1993). The second parameter is "neighbors" and allows the Cubist model to group similar samples in terms of predictor variable values, and determine the average prediction of these training samples (Kuhn et al. 2012; Quinlan 1993). We tuned the models in the R software (R Core Team 2013) over different values of "committees" and "neighbors" ("committees" was set to be 0, 10, 20, 50, 100, and "neighbors" was set to be 0, 1, 5, 9) through a 10-fold cross-validation on our training data and selected the values that produced the smallest root mean square error (RMSE).

The resulting model was evaluated on both the training data that were used to generate the model and the independent validation data (Section 4.4.1.2). Three statistical measures were used to assess model performance: the coefficient of determination (R^2), mean absolute error (MAE) and the RMSE.

Table 4.2 . Overview of the predictor variables used in this study. Mo 1250 nm), SWIR ₂ (1628–1652 nm), and SWIR ₃ (2105–2155 nm).	DDIS shortwave infrared bands are referre	d to as SWIR ₁ (1230–
Predictor variable	Formula	Reference
MODIS individual bands (Red, NIR, Blue, Green, SWIR1, SWIR2, SWIR3)	1	
NDVI – Normalized Difference Vegetation Index	(NIR - Red)/(NIR + Red)	Tucker (1979)
NDWI – Normalized Difference Water Index	(Green - NIR)/(Green + NIR)	McFeeters (1996)
MNDWI – Modified Normalized Difference Water Index	$(Green - SWIR_2)/(Green + SWIR_2)$	Xu (2006)
NDWI – Normalized Difference Water Index (referred as LSWI _{B5})	$(NIR - SWIR_1) / (NIR + SWIR_1)$	Gao (1996)
LSWI – Land Surface Water Index (referred as LSWI _{B6})	$(NIR-SWIR_2)/(NIR+SWIR_2)$	Xiao <i>et al.</i> (2002a)
TCWI – Tasseled Cap Wetness Index	$0.10839 \times \text{Red} + 0.0912 \times \text{NIR} + 0.5065 \times$ Blue + 0.404 × Green + 0.241 × SWIR ₁ - 0.4658 × SWIR ₂ - 0.5306 × SWIR ₃	Zhang et al. (2002)
TCBI – Tasseled Cap Brightness Index	0.3956 × Red + 0.4718 × NIR + 0.3354 × Bh + 0.3834 × Green + 0.3946 × SWIR ₁ + 0.343 × SWIR ₂ + 0.2964 × SWIR ₃	te Zhang <i>et al.</i> (2002) 4

92

Predictor variable	Formula	Reference
Value (HSV)	max(SWIR2, NIR, Red)	Pekel et al. (2014)
Saturation (HSV)	$1-\min(SWIR_2,NIR,Red)/\max(SWIR_2,NIR,Red)$	Pekel <i>et al.</i> (2014)
Hue (HSV)	$\begin{array}{l} 0, \text{if } V = \min(\text{SWIR}_2, \text{NIR}, \text{Red}) \\ \hline \left(\begin{matrix} 60^\circ \times (\text{NIR} - \text{Red}) \\ \overline{(V - \min(\text{SWIR}_2, \text{NIR}, \text{Red})} + 360^\circ \right) \mod 360^\circ, \text{if } V = \text{SWIR}_2 \\ \hline 60^\circ \times (\text{Red} - \text{SWIR}_2) \\ \hline \overline{V - \min(\text{SWIR}_2, \text{NIR}, \text{Red})} + 120^\circ, \text{if } V = \text{NIR} \\ \hline \hline 60^\circ \times (\text{SWIR}_2, \text{NIR}, \text{Red}) \\ \hline \overline{V - \min(\text{SWIR}_2, \text{NIR}, \text{Red})} + 240^\circ, \text{if } V = \text{Red} \\ \hline \hline V - \min(\text{SWIR}_2, \text{NIR}, \text{Red}) \\ \hline \end{array}$	Pekel <i>et al.</i> (2014)
TWI – Topographic Wetness index	$\begin{split} &\ln(\alpha/\tan\beta); \ \alpha \ \text{is upslope area per unit contour length (m) which calculated as} \\ & (flow accumulation + 1) \times (cell size); \ \beta \ \text{is the slope} \\ & expressed in radians \end{split}$	Beven and Kirkby (1979)

4.4.1.4 Mapping and post processing

We applied the resulting model to the MCD43A4 V006 data from February 2000 to December 2017 to produce gridded time series of surface water fraction for the Mediterranean region with an 8 d time step resulting in 46 maps per year (except for 2000, which only contained 39 images).

Recent studies on surface water detection using optical sensors showed that multiple sources of commission errors exist, such as terrain and building shadows (Klein *et al.* 2017; Pekel *et al.* 2016). These errors can be corrected using masks derived from auxiliary data (Klein *et al.* 2017; Pekel *et al.* 2016). In this study we addressed two sources of commission errors: shadows from buildings and identified surface water presence that is unlikely on sloping terrain. Specifically, a slope map was derived from the SRTM DEM by calculating the maximum rate of change in elevation from each raster cell to their eight neighbors. We then used a threshold of 5° to identify steep locations where it is unlikely to find surface water but could have been detected as water by our model, for example, due to spectral confusion between water and terrain shadows (Yamazaki *et al.* 2015). In the case of building shadows, we used the urban class of the MODIS classification product MCD12Q1 (Friedl *et al.* 2010) to assign areas potentially affected by building-induced shadows. Pixels were reassigned to the 0% water fraction for slopes steeper than 5° and for areas classified as urban in MCD12Q1, except for places where water was present according to the GSW maximum water extent.

4.4.2 Generation of the surface water fraction metrics and comparison of products

Based on gridded time series of surface water fraction, we derived a series of ecologically relevant metrics that capture both the intra- and inter-annual variability and changes, and further compared these metrics with GSW-derived thematic products. The readily available GSW thematic products were derived from 32 years of data (between March 1984 and October 2015); thus, they cannot be directly compared against our MODIS-based results for 2000–2017. Therefore, we reproduced the GSW thematic maps using the GSW monthly water history dataset from the overlapping period of these two datasets (*i.e.*, February 2000–October 2015). We then also computed the MODIS-derived temporal metrics for this period. In total, 189 GSW water history images and 720 surface water fraction maps were incorporated for the generation of these thematic maps. The following metrics were generated:

- (1) The annual maximum and mean annual maximum surface water fraction between 2000 and 2015: the GSW monthly water history maps were summarized for each year in GEE to calculate the annual maximum surface water extent. We then aggregated the results of each year to the MODIS resolution and averaged all years to derive the mean annual maximum surface water fraction. To assess the spatial agreement of the mean annual maximum surface water fraction derived from MODIS and GSW, we calculated the surface water area (in km²) from the mean annual maximum surface water fraction using a threshold continuum. Specifically, the surface water fraction was partitioned using nine threshold values set in 10 % increments from 0 % to 100 %. All pixels with a surface water fraction greater than or equal to the threshold were summed and then multiplied by the MODIS pixel size. We then compared the water area (in km²) derived from the mean annual maximum surface water fraction based on GSW and MODIS across a different continuum of threshold values. We also compared the water area with the 250 m static water mask from MOD44W.
- (2) The standard deviation of the annual maximum: a measure of the inter-annual variability of water presence.
- (3) The seasonality: a measure for the seasonal and intra-annual variability of water presence. We calculated the number of times (*i.e.*, the water occurrence) a given pixel displayed standing water above a certain water fractional threshold for a single hydrological year (October 2014 to September 2015), as GSW has a relatively large number of valid observations for this year (Figure 4.2), and then classified different types of water permanence. The water occurrence was calculated for each grid cell as a fraction of the number of times water was present relative to the total valid observations (*i.e.*, not affected by clouds). We adopted the classification criteria of Guerschmann *et al.* (2011) and further modified it to be applicable to the Mediterranean region (Table 4.3). As the classes are not mutually exclusive, they are prioritized in the order shown in Table 4.3.

Table 4.3. Seasonality classification criteria based on surface water fraction and occurrence. The occurrence indicates how often the specified surface water fraction is reached for a single hydrological year (October 2014–September 2015).

Class name	Surface water fraction	Water occurrence
Permanent water	$\geq 70\%$	\geq 90%
Semipermanent water	$\geq 70\%$	70%-90%
Intermittent water	$\geq 70\%$	20%-70%
Infrequent inundation	$\geq 70\%$	1%-20%
Mixed permanent and semipermanent water	30%-70%	$\geq 70\%$
Mixed intermittent water	30%-70%	20%-70%
Mixed infrequent inundation	30%-70%	1%-20%
Never inundated	< 30%	$\geq 0\%$

4.4.3 Demonstrating the representation of surface water dynamics by the new MODIS dataset

To assess the performance of our MODIS-derived product for monitoring temporal variations in the surface water extent, we selected three lakes with fluctuating water presence. These three sites have varying sizes, and different geographic locations and temporal dynamics, and have also been listed in the International Conventions on Wetlands (known as Ramsar) given their importance for staging and wintering waterfowl. The three sites are as follows:

- The Fuente de Piedra lake, Spain, which is a shallow and saline lake, has a maximum area of 13.6 km². It experiences strong seasonal, inter-annual, and intra-annual variations of water level and inundation extent (Li *et al.* 2015);
- (2) Lake Sabkhat al-Jabbul, Syria, which is a large, permanent saline lake that is surrounded by semiarid steppe. At high water levels, it contains two islands; it traditionally floods in the spring and shrinks back during the summer and autumn but seldom dries out completely (JAES-CC 2010);
- (3) The coastal marshland complex of Doñana, Spain, which is separated from the ocean by an extensive dune system and is subject to seasonal and inter-annual variations in water level (De Castro and Reinoso 1997);

For the first two lakes, we compared the time series of the MODIS-derived surface water area with that from GSW, and further compared these against water level data from
satellite altimetry or *in situ* measurements. We calculated the Spearman rank correlation (ρ) between water level and water area derived from MODIS SWF and JRC's GSW data to assess the correspondence between these datasets. For Doñana, we compared the monthly spatial distribution of the surface water extent derived from MODIS and JRC's GSW. To ensure the accuracy of the area calculations, we only calculated an area for times when it contained at least 95 % of valid data from MODIS SWF and JRC's GSW data.

4.5 Results

4.5.1 Model performance

Following the model tuning of the Cubist regression model (see Section 4.4.1.3), we found that a 20-member committee and 9-neighbor model resulted in the smallest RMSE between the actual (GSW-derived) water fraction and our MODIS-based water fraction estimates. The addition of more committees or neighbors had little effect on the accuracy.

Table 4.4. Statistical measures between the predicted surface water fraction and the actual data for training and validation data, and for different types of water permanence using validation data.

	\mathbb{R}^2	RMSE (%)	MAE (%)
Training data	0.93	9.79	5.61
Validation data	0.91	11.41	6.39
Permanent water	0.90	12.07	6.38
Fluctuating water	0.85	12.60	8.11

Table 4.4 shows the statistical measures between the predicted and actual surface water fraction. The model predicted surface water fraction shows good agreement with the actual value with an R^2 of 0.93 for the training data. When testing using the independent validation dataset, the R^2 value is only slightly smaller, and the RMSE and MAE are slightly larger, suggesting that the Cubist regression model does not suffer from overfitting. The RMSE and MAE for fluctuating water were slightly larger than for permanent water (Table 4.4). This suggests that the developed model not only provides accurate results for the static mapping of the surface water fraction, but it can also be applied effectively for monitoring the dynamics of fluctuating surface water.

4.5.2 Surface water fraction metrics and comparison of products

The mean annual maximum surface water fraction values generated from GSW and the MODIS-derived product over the 2000–2015 period are displayed in Figure 4.4. Overall, the two maps are in good agreement. Visual comparison indicates that our MODISderived product is able to detect narrow rivers with widths covering a couple of MODIS pixels, such as the Danube, Euphrates, Po, Rhine, and Tagus rivers. Large differences are evident for some low surface water fraction regions such as parts of Hungary and the Ukraine (Figure 4.4). These differences likely correspond to the presence of wet meadows, salt marshes, and floodplains along large rivers, which are usually saturated and inundated with water during most of the vegetative season (Šefferová Stanová et al. 2008; Stefan et al. 2016). Around the Po River in Italy, where rice paddies are seasonally present, our MODIS product also shows larger surface water fractions (Figure 4.4c). This result suggests that the MODIS-derived surface water fraction has enhanced sensitivity to surface water in wetland areas with emerged vegetation, and this could be attributed to the fact that several predictor variables such as LSWI (Xiao et al. 2002a) are also sensitive to vegetation water content (Li et al. 2015). Table 4.5 confirms that the total surface water areas for the Mediterranean calculated from MODIS are more comparable to the GSW results when only considering areas with a higher surface water fraction. For example, when only accounting for pixels with surface water fractions equal to or greater than 50%, the total surface water areas for the Mediterranean based on both datasets are similar (75 107 km² for GSW versus 73 444 km² for MODIS). In comparison, only 70 543 km² of water was detected in this region based on the 250 m static water mask from MOD44W. This implies that our MODIS product detects more surface water than other coarse-resolution binary maps. Nonetheless, our MODIS product detects less surface water than GSW for larger thresholds (\geq 50 %), whereas it detects much more surface water than GSW for small thresholds (≤ 20 %). This confirms an earlier finding that machine learning approaches such as Cubist and random forest often underestimate large values and overestimate small values when estimating the fractional cover of land surface (e.g., Huang et al. 2014b; Li et al. 2018a; Wang et al. 2017). In addition to the effects of mixed pixels (Klein et al. 2017), the most obvious reason for this is because regression techniques used in such approaches fit linear equations to relationships that may not be linear over the entire range of values.

Threshold for surface Total surface water areas Total surface water areas (km²) based water fraction (km²) based on GSW on MODIS surface water fraction 90% 48 718 47 145 55 887 80% 51 778 70% 62 371 58 996 60% 68 855 65 993 50% 75 107 73 444 40% 81 220 82 417 30% 87 217 96 239 20% 93 207 142 531 10% 99 260 284 916

Table 4.5. Comparison of total surface water area (in km^2) as determined from JRC's GSW and MODIS mean annual maximum surface water fraction maps for different thresholds.





Figure 4.4. Mean annual maximum surface water fraction maps as obtained from (a) JRC's GSW and (b) MODIS time series surface water fraction for the period from 2000 to 2015, and (c) the difference between the two maps. Positive difference values indicate that our MODIS dataset detected a larger water fraction than JRC's GSW.





Figure 4.5. Standard deviation of the surface water fraction as calculated from (a) JRC's GSW and (b) MODIS annual maximum surface water fraction maps for the period from 2000 to 2015.

Figure 4.5 shows the standard deviation of the annual maximum surface water fraction. It indicates that areas of large inter-annual variability agree between MODIS-based and GSW-based results. Both indicate a larger variability in the surface water fraction in semiarid and desert climate zones, particularly in the north of Algeria, for the Volga Delta in the Caspian depression, and along the Tigris and Euphrates rivers of Iraq.





Figure 4.6. Seasonality information derived from time series of the (a) JRC's GSW and (b) MODIS surface water fraction for a single year (October 2014 to September 2015).

Figure 4.6 displays the seasonality metric for the entire study area, with details for two selected sites shown in panels (a) and (b) of Figures 4.7 and 4.8. Fuente de Piedra is a seasonally flooded lake which usually dries out completely in summer (May–September) (Batanero *et al.* 2017; Li *et al.* 2015) with the exception of extremely wet years (*e.g.,* 2010, 2011, 2013: Rodriguez-Rodriguez *et al.* 2016) when water was present throughout the whole year. The differences in water seasonality for Fuente de Piedra between the two products (Figure 4.7a, b) can be attributed to the fact that GSW lacks observations in wet seasons, resulting in a reduced water occurrence compared with our MODIS product. The discontinuities in the Landsat record between seasons can affect the accuracy of seasonality information, which has also been demonstrated by Klein *et al.* (2017) and Pekel *et al.* (2016). Permanent water is present in parts of Lake Sabkhat al-Jabbul, but the

larger central portion of the lake has highly dynamic intermittent water (Figure 4.8). Mixed permanent and semipermanent waters are mostly found on the edge of permanent water and in narrow rivers.

4.5.3 MODIS-derived surface water dynamics for selected lakes

Panels (d) of Figures 4.7–4.8 show the time series of the surface water extent detected by MODIS and GSW for two selected lakes. Fuente de Piedra (Figure 4.7d) experiences large temporal variability in the surface water extent throughout the year, which is well represented by our MODIS product with 461 time steps (Table 4.6). This variability corresponds closely to the *in situ* water level data ($\rho = 0.95$). Note that in extreme wet years (i.e., 2010, 2011, and 2013), the lake remained flooded throughout the year without increasing in size with regard to water level changes (Rodriguez-Rodriguez et al. 2016). The water extent derived from MODIS SWF also matched closely to that from GSW (Figure 4.7c, d; r = 0.95). GSW only had 73 valid time steps with most observations in dry seasons (*i.e.*, June to October); thus, it did not allow for the appropriate capture of the seasonal dynamics, particularly for the November-March period when the lake usually reaches its full surface water extent. Time series of the water extent of Lake Sabkhat al-Jabbul as determined by our MODIS product showed a relative high correlation with water level data ($\rho = 0.69$). It also showed good agreement with the water extent derived from GSW, including the seasonal peak extent and the minimum surface water extent during the dry season (Figure 4.8c, d; r = 0.88). This implies that the coarse 500 m MODIS data not only provide more detailed temporal information (563 MODIS surface water fraction time steps versus 69 GSW time steps) (Table 4.6), but they also give accurate estimations of the surface water area compared with the results from Landsat 30 m resolution data.

Lakes	MODIS surface water fraction	JRC's GSW
Fuente de Piedra	461	73
Lake Sabkhat al-Jabbul	563	69
Doñana	714	70

Table 4.6. Number of valid temporal observations for the three lakes based on the MODIS surface water fraction and JRC's GSW between February 2000 and October 2015.

A New Dense 18-Year Time Series of Surface Water Fraction Estimates from MODIS



Figure 4.7. Seasonality information derived from (a) JRC's GSW and (b) the MODIS surface water fraction from a single year (October 2014 to September 2015). The colors in (a) and (b) are the same as in Figure 4.6. (c) A scatterplot of the water area obtained from JRC's GSW versus that from MODIS SWF. r represents the Pearson correlation between two datasets. (d) A comparison of time series of the surface water area (in km^2) derived from JRC's GSW (shown using green asterisks) with MODIS surface water fraction (shown using blue dots) from 2000 to 2015, along with in situ water level data (shown using orange dots) for Fuente de Piedra, Spain. ρ represents the Spearman rank correlation between the water level and water area.



Figure 4.8. As in Figure 4.7, but for Lake Sabkhat al-Jabbul, Syria. The water level for this lake is computed from Jason-2/OSTM altimetry and repeats every 10 d.

Figure 4.9 compares the MODIS monthly surface water fraction with the GSW monthly water history for Doñana, Spain. The visual comparison shows that the distribution of the MODIS surface water fraction agrees well with the GSW monthly water maps. The 500 m MODIS surface water fraction is able to capture the spatial patterns as detected from the high-resolution Landsat-based GSW dataset. Seasonal drying out and flooding of the wetland is well detected with the MODIS-derived surface water fraction, whereas GSW lacks temporal details, for example, during large water extents in November and December. MODIS also captures the timing of the maximum water extent (*i.e.*, in January) and water retreat (*i.e.*, in July). This example highlights that the information provided by our MODIS product can contribute to a better understanding of surface water dynamics.



A New Dense 18-Year Time Series of Surface Water Fraction Estimates from MODIS

Figure 4.9. Comparison of monthly water distribution based on (a) the Landsat-based GSW dataset and (b) the MODIS surface water fraction for Doñana, Spain, for 2011.

4.6 Discussion

We estimated the surface water fraction for the Mediterranean region from MODIS data, and improved on previous efforts to estimate surface water fraction from mediumresolution imagery (MODIS or similar). The prediction accuracy of our model (R^2 = 0.91, RMSE = 11.41 %, and MAE = 6.39 %) is higher than the R² of 0.625 reported by Weiss and Crabtree (2011), who used a linear regression model, and the R² of 0.7 reported by Guerschmann *et al.* (2011), using a logistic regression model. This research successfully expanded our previous work (Li *et al.* 2018a) by upscaling it from a relatively small region to the whole Mediterranean while retaining a similar high accuracy (both achieved an R² of 0.91). This attributes to the feasibility and robustness of Cubist regression modeling with respect to dealing with different environmental conditions when training data are collected across a wide geography resulting in varying spectral characteristics.

Secondly, we generated surface water fraction maps at high temporal frequency, which is an advantage over existing fine-resolution datasets. Our MODIS-derived surface water

Chapter 4

fraction product accurately displays the spatiotemporal variability of surface water. The comparison with the GSW 30 m product and water level data (Figures 4.7-4.9) reveal that our product can efficiently monitor the seasonal, intra-annual, and long-term surface water dynamics with good spatial and temporal accuracy. For example, it complements the GSW by allowing for the detection of inundation and recession processes over short time periods and for the better representation of seasonality changes and temporal trends over longer periods. The MODIS surface water fraction can also accurately detect the spatial distribution of surface water inclusive of small water bodies (less than one MODIS pixel) and narrow rivers, which are missing in other coarse-resolution products using binary classification method such as MOD44W (Salomon et al. 2004). Metrics derived from the MODIS surface water fraction and GSW time series can reflect different facets of surface water dynamics. The accuracy of some metrics, such as seasonality, relies on the number of valid observations and the temporal interval. Landsat-derived metrics might be problematic in areas with large temporal gaps caused by persistent cloud cover, as shown in this study (Figure 4.8) and previously by Pekel et al. (2016) and Klein et al. (2017). Although cloud coverage also limits MODIS observations, the probability of obtaining cloud-free observations is higher (Table 4.6) due to the daily acquisitions and consequent temporal compositing possibilities. With these advantages, we expect that our MODIS surface water fraction product could fill in important information on surface water for areas and time periods for which cloud-free Landsat acquisitions are few or nonexistent.

Our MODIS-derived surface water fraction product also has limitations. Firstly, it is designed to detect only open surface water; therefore, it may not effectively capture water bodies covered by dense vegetation, such as swamps, lakes with considerable coverage of aquatic vegetation, and inundated dense forests. Secondly, the MODIS surface water fraction product overestimates small surface water fractions of less than 20 % (Table 4.5), which has also been found in previous studies on surface water fraction mapping (Li *et al.* 2018a; Parrens *et al.* 2017). This overestimation might be attributed to the mixed spectral response of pixels with different land cover types, as already demonstrated by many studies (*e.g.,* Guerschmann *et al.* 2011; Klein *et al.* 2017). Further work could consider reassigning pixels with less than 20 % surface water to 0 % water fraction for locations where water is never present according to the GSW maximum water extent. Thirdly, the auxiliary layers utilized for the identification of potential areas of commission errors also appear to have some limitations. For example, some urban areas might be not mapped in MCD12Q1, and areas of cloud cover and cloud shadow might

not be completely removed from the MCD43A4 product. The importance of these limitations may diminish as the quality of these auxiliary layers improves or dynamic datasets rather than static layers are incorporated (*e.g.*, Global Human Settlement Layer: Pesaresi *et al.* 2016). Fourthly, the MCD43A4 product also suffers from many missing values, especially in regions with large amounts of precipitation, aerosol concentrations, or snow and ice coverage (Klein *et al.* 2017). Future research should focus on combining other moderate-resolution data (*e.g.*, MOD09) for areas with missing data to ensure a gap-free reconstruction of inland water development for the past and future.

This work can be further scaled up over much larger regions and for shorter (*e.g.*, daily) time intervals. A high-quality training dataset is crucial for the effective application of the rule-base regression model and the collection of such training data is time consuming (Sun *et al.* 2012). In this paper, time series of the GSW monthly water history dataset proved to be an efficient basis for building a reliable training dataset. Considering that GSW is globally available, we are confident that our approach can be scaled to monitor the surface water fraction globally with MODIS data. Although the surface water fraction maps were produced with an 8 d time step, the model developed in this paper was actually trained using daily MODIS data and could be directly applied to that temporal resolution. The resulting daily surface water fraction maps could be of high interest for ecological and hydrological research. The recent GCOS report requires water extent and lake ice cover with a daily temporal resolution and a 20 m and 300 m spatial resolution, respectively (Belward 2016). Our approach makes this requirement for the coarser of the two resolutions within reach.

Our MODIS surface water fraction dataset may benefit a large number of applications. For example, it could be used as a monitoring tool for analyzing hydrologic extremes such as floods and droughts, detecting abnormal changes of wetland hydrology, capturing short-duration events, identifying newly formed and disappearing water bodies, and estimating global water loss. The MODIS surface water fraction dataset may help to improve the calibration and validation of hydrological models. For example, the water area can help to estimate a series of hydrological parameters such as water discharge (Huang *et al.* 2018b) and water volume (Busker *et al.* 2019; Cael *et al.* 2017; Duan and Bastiaanssen 2013; Tong *et al.* 2016). This would be particularly useful for areas where *in situ* measurements are sparse or inaccessible. Closely monitoring hydrological variability, and human activities affect the dynamics of surface water and human livelihoods (Tulbure and Broich 2019; Zhang *et al.* 2019). It may also provide new insights for understanding

how surface water dynamics further influence climate. For example, lake expansion and the creation of new dams can alter local and regional precipitation patterns (Degu *et al.* 2011; Ekhtiari *et al.* 2017; Hossain *et al.* 2009). Similarly, our long-term records and derived metrics have the potential to contribute to the management and conservation of biodiversity and other ecosystem services associated with terrestrial surface water and wetlands.

4.7 Conclusions

We derived an 8 d, 500 m resolution surface water fraction product over the Mediterranean for the period from 2000 to 2017 by applying a global Cubist regression tree model to MODIS and SRTM data. We validated the results with JRC's Landsatderived GSW dataset, which resulted in a high overall accuracy ($R^2 = 0.91$, RMSE = 11.41 %, and MAE = 6.39 %). The MODIS-derived surface water fraction showed a good spatial and temporal correspondence with JRC's GSW. Comparison with satellite altimetry and *in situ* water level data for selected lakes demonstrated the ability of MODIS surface water fraction to effectively monitor seasonal and inter-annual changes in surface water extent. Our dataset provides a consistent, long-term record (18 years) of 8 d water fraction dynamics for the Mediterranean region, and complements fine spatial resolution surface water products, especially in regions where such products have long temporal and spatial data gaps due to both the limited number of acquisitions and persistent cloud cover. Our approach is also promising for monitoring surface water fraction at the global scale and at a daily interval. A New Dense 18-Year Time Series of Surface Water Fraction Estimates from MODIS

Chapter 5

Evaluation of a New 18-year MODIS-Derived Surface Water Fraction Dataset for Constructing Mediterranean Wetland Surface Water Dynamics^{*}

^{*} This chapter is based on: Li, L., Vrieling, A., Skidmore, A., & Wang, T. (Submitted). Evaluation of a new 18-year MODIS-derived surface water fraction dataset for constructing Mediterranean wetland surface water dynamics. *Journal of Hydrology*.

Abstract

Wetlands are among the most biodiverse ecosystems in the world, due largely to their dynamic hydrology. Frequent observations by satellite sensors such as the Moderate Resolution Imaging Spectrometer (MODIS) allow for monitoring the seasonal, interannual and long-term dynamics of surface water extent. However, existing MODIS-based studies have only demonstrated this for large water bodies (>100 km²) despite the ecological importance of smaller-sized wetland systems. In this paper, we assessed if the temporal dynamics of surface water extent can be effectively captured with MODIS for a wide range of water body sizes, specifically 0.01 km² and larger. We constructed the temporal dynamics of surface water extent for 340 individual water bodies in the Mediterranean region between 2000 and 2017, using a previously developed 8-day 500 m MODIS surface water fraction (SWF) dataset. We then evaluated how MODIS SWF time series corresponded to water extent derived from a Landsat-based dataset and to satellite-altimetry derived water level data. Results showed that although correlations between MODIS- and Landsat-derived SWF were poor for relative static water bodies (r = 0.17), they increased for more dynamic water bodies (r = 0.81). For dynamic water bodies larger than 100 km², time series of water extent derived from MODIS SWF showed good correlation with both Landsat ($r \ge 0.76$) and water level data ($\rho \ge 0.63$). Our MODIS SWF dataset can also effectively monitor the variability of very small water bodies ($< 1 \text{ km}^2$) when comparing with Landsat data as long as variability in their area was high. We conclude that MODIS SWF is a useful product to help understand hydrological dynamics for both small and larger-sized water bodies, and to monitor their seasonal, intermittent, inter-annual and long-term changes.

5.1 Introduction

Wetlands are among the most biodiverse ecosystems in the world, due largely to their dynamic hydrology (Hils 1997). The hydrological cycle of flooding and drying within a wetland is important for the formation and maintenance of specific wetland habitat types and the species that they support (Hails 1997). The hydrological dynamics also reflect spatial and temporal changes in a number of environmental factors including anomalous rainfall-driven flood events (Cian et al. 2018), seasonal thawing and snowmelt in spring (Watts et al. 2012) and longer-term environmental changes (Street and Grove 1976). Characterizing hydrological variations is a prerequisite for understanding hydrological and ecological processes across different wetland types, and to identify abnormal changes in wetland hydrology in relationship to normal variability of wetland hydrology general patterns. However, accurate information on wetland dynamics is scarce, particularly for small-sized wetlands (Ogilvie et al. 2018a). Despite their size, small wetlands in arid, semi-arid and Mediterranean regions often act as critical refuge and breeding areas, offer food sources for wildlife, and harbor many plant and animal species that would otherwise not survive in the surrounding landscape (Calhoun et al. 2017; McCulloch et al. 2003). Efficient, accurate and robust tools for monitoring the hydrology of wetlands over large areas are urgently needed for management and conservation strategies (Finlayson et al. 1999; Turak et al. 2017).

Remote sensing has been used as a cost effective tool to monitor the changes of surface water in space and time (Alsdorf *et al.* 2007; Ozesmi and Bauer 2002). The challenge is the trade-off between temporal and spatial resolution of remotely sensed imagery. High spatial resolution has proved to provide accurate estimation of the location and extent of surface water while high temporal resolution imagery is effective in monitoring and analyzing the dynamics of surface water.

Many attempts to monitor surface water changes with 10-30 m resolution optical data have been reported including spectral data from Landsat (Borro *et al.* 2014; Doña *et al.* 2016; Jin *et al.* 2017; Tan *et al.* 2019) and Sentinel-2 (Kaplan and Avdan 2017a; Yang *et al.* 2018b) for specific locations. For larger spatial and longer time scales, several efforts exist to map and monitor surface water changes with Landsat data at regional (Halabisky *et al.* 2016; Heimhuber *et al.* 2016), continental (Mueller *et al.* 2016), and global scale (Donchyts *et al.* 2016; Pekel *et al.* 2016). These datasets allow for analyzing surface water dynamics at fine spatial resolution and over a long time frame. However, the combination of clouds and revisit time of the used satellites results in data gaps and temporal

discontinuities in the data (Pekel *et al.* 2016). The limited number of observations does not provide enough detail for understanding seasonal patterns and dynamics of interannual and long-trend response, much less to determine if measured changes in the surface water extent represent natural seasonal variability, or abnormal changes in wetland hydrology.

Moderate to coarse resolution imagery are acquired with shorter revisit times and provide long-term and dense time series data. Therefore, they have been widely used for monitoring changes in surface water. For example, Chen et al. (2014) investigated the spatial-temporal patterns of abrupt changes of Poyang Lake using time-series of the Moderate Resolution Imaging Spectroradiometer (MODIS) 16-day Normalized Difference Vegetation Index (NDVI) product. Using the same product and unsupervised classification, Gao et al. (2012) estimated historical area changes in 34 large reservoirs around the globe. Khandelwal et al. (2017) applied supervised classification on 16-day 500 m MODIS surface reflectance time series data to monitor changes in surface water extent for a global set of 94 large reservoirs. All of these studies are based on binary classification methods and are confined to large water bodies with an area greater than 100 km². To overcome the limitation of coarse spatial resolution and incorporate small water bodies, we have developed a dense 18-year time series dataset of surface water for the Mediterranean region in a previous paper (Li et al. 2019). This dataset estimated the surface water fraction (SWF) for each 500 m MODIS pixel (using nadir reflectances from the MCD43A4 product) with a rule-based regression model. By assessing water fractions, the dataset allows for mapping and monitoring surface water smaller than a MODIS pixel. Although in that paper we briefly illustrated the dataset's options for monitoring temporal changes, a formal assessment of the ability to monitor surface water dynamics taking into account water body size has not been performed. Therefore, the aim of this paper is to assess if our MODIS SWF product can effectively capture the temporal changes in surface water extent for a wide range of water body sizes (*i.e.*, 0.01-3100 km²). Our specific objectives are:

- to derive temporal dynamics of surface water extent between 2000 and 2017 for individual water bodies using the previously developed 8-day 500 m resolution MODIS SWF dataset;
- (2) to assess how closely those temporal dynamics match the variability of surface water extent obtained from a Landsat-based dataset and water level data;
- (3) to characterize the water extent time series and to identify different dynamic patterns (*e.g.*, intermittent, seasonal, inter-annual or long-trend dynamics).

5.2 Data

5.2.1 Landsat-based global surface water dataset

The Joint Research Centre's Global Surface Water (JRC-GSW) dataset (Pekel *et al.* 2016) maps the temporal and spatial dynamics of surface water over a 32-year period (from March 1984 to October 2015) at monthly time interval and at 30 m spatial resolution globally. This dataset was produced by analyzing the entire archive of Landsat 5, 7 and 8 using an expert system. The expert system is a sequential decision tree that used both the multi-spectral and multi-temporal attributes of the Landsat archive as well as ancillary data layers. The dataset was validated using more than 40,000 reference points from very high resolution satellite and aerial imagery, which revealed a high accuracy with a commission error less than 1% and omission error less than 5% (Pekel *et al.* 2016). The JRC-GSW dataset includes the monthly water history dataset, which is a set of 380 global maps documenting the water presence for each month from March 1984 to October 2015, and eight additional information layers, documenting different facets of the surface water dynamics. The complete dataset with the described layers is available in Google Earth Engine.

Although JRC-GSW has a high spatial accuracy, it only provides information on changes in surface water at a monthly interval. Furthermore, it contains long temporal gaps due to both the limited number of acquisitions during specific time intervals, and due to regionand month-dependent persistency of cloud cover. These temporal gaps can affect the accuracy of the trend and seasonality information (Li *et al.* 2019; Yamazaki and Trigg 2016)

5.2.2 MODIS-derived surface water fraction dataset

The MODIS-based surface water fraction (SWF) dataset (Li 2019) provides 18 years (2000–2017) of water fraction maps at 500 m spatial resolution and 8-day temporal resolution for the Mediterranean region. We derived this dataset from the MODIS Terra and Aqua Nadir BRDF-Adjusted Reflectance (NBAR) product (MCD43A4, V006) through a global rule-based regression model, which used MODIS-derived spectral information, its temporal characteristics, and topographic information as predictor variables. Reference data were generated from the JRC-GSW monthly water history maps for 7,000 MODIS-scale cells (500x500 m), divided into 3500 training and 3500 validation locations. These reference data areas are sampled across the entire Mediterranean area,

different water permanence types (seasonal and permanent), and different water fraction categories (0%–100%). The correspondence between MODIS SWF and surface water fraction derived from the JRC-GSW product is strong ($R^2 = 0.91$). An advantage of MODIS SWF over JRC-GSW is that it can provide surface water information for the long temporal and spatial data gaps of JRC-GSW. Moreover, the high temporal resolution allows better capturing of rapid changes in surface water extent that cannot be detected by the monthly JRC-GSW product (Li *et al.* 2019).

5.2.3 Water level data

We obtained water level time series measured with spaceborne radar altimetry for 14 large water bodies (>100 km²) (Figure 5.1 and Table 5.3) from three websites: (1) the U.S. Department of Agriculture's Global Reservoir and Lake Monitoring (GRLM) (http://www.pecad.fas.usda.gov/cropexplorer/global reservoir); (2) the Laboratoire d'Etudes en Géophysique et Océanographies Spatiales (LEGOS: http://hydroweb.theialand.fr/) in France (Crétaux et al. 2011), and (3) the Database for Hydrological Time Series of Inland Waters (DAHITI: http://dahiti.dgfi.tum.de/en/) (Schwatke et al. 2015). The GRLM combines water levels obtained from radar altimeters onboard Topex/Poseidon (T/P), Jason-1, Jason-2 and ENVISAT satellites. GRLM merges T/P, Jason-1 and Jason-2 time series of relative water level variation with respect to the nineyear mean T/P level at 10-day intervals (hereafter referred to as GRLM10). An ENVISAT Radar Altimeter 2 derived time series of relative water level variation with respect to the mean level of a given ENVISAT reference cycle at 35-day intervals is also provided. The GRLM offers both the raw and smoothed time series of water level, which is performed with a median-type filter to eliminate outliers and reduce high-frequency noise. We used the smoothed data in this study. LEGOS and DAHITI use additional altimeter missions besides those used by GRLM, including Jason-3, Cryosat-2, and SARAL/AltiKa.

5.3 Methods

5.3.1 Selection of individual water bodies

To compare water extent dynamics from MODIS SWF and JRC-GSW for individual water bodies, we selected and analyzed 340 water bodies. These were all situated in the Mediterranean region as defined by Li *et al.* (2019). These water bodies vary from 0.01 km² to 3100 km² in size (*i.e.*, mean water area according to JRC-GSW time series). Table

5.1 lists the number of water bodies within five size categories, *i.e.*, 0-1 km², 1-5 km², 5-20 km², 20-100 km², and > 100 km², and Figure 5.1 shows their spatial distribution.

To obtain our selection of 340 water bodies, we first selected all Ramsar wetlands of the Mediterranean region excluding the wetlands types that are not characterized by the presence of surface water, such as wet meadows, marshes, and peatlands. This resulted in 279 water bodies. We added 14 large lakes and reservoirs (>100 km²) to our selection (see Figure 5.1 and the list in Table 5.3). These 14 lakes are not Ramsar sites but were included by Khandelwal *et al.* (2017) in their study on MODIS-based surface water extent mapping, and coincided with the lakes for which altimetry water level data were available (Section 5.2.3). We thus included these lakes for comparison. In addition, we included 47 non-Ramsar water bodies that showed large inter-annual variability in surface water extent according to Li *et al.* (2019).

Table 5.1. Number of selected water bodies in different water size categories (total = 340)

Size (km ²)	Number of water bodies
0–1	95
1–5	88
5–20	62
20-100	56
>100	39



Figure 5.1. Spatial distribution of all analyzed water bodies. Different colours represent different water size categories. Stars show the lakes for which water level data are available, and have an overlap with the research of Khandelwal et al. (2017).

5.3.2 Calculation of water extent time series

To identify surface water dynamics per water body, we first identified the spatial extent within which we summarized the surface water presence. For this purpose, we used the JRC-GSW maximum water extent as the initial region of interest (ROI). Around this ROI, we created a buffer of one 500 m resolution MODIS pixel to include nearby locations, because there is a chance that surface water extends beyond what could be identified by JRC-GSW.

We then derived time series of water extent for each buffered ROI from the MODIS SWF dataset, by summing the water fraction of all pixels inside the ROI and multiplying this by the MODIS pixel size (*i.e.*, 250,000 m²). We also derived time series of water extent from the JRC-GSW monthly history dataset by multiplying the number of water pixels inside the ROI with the Landsat pixel size (*i.e.*, 900 m²). Missing data in a ROI can affect the accuracy when calculating the water area. To deal with this problem and ensure the accuracy of the area calculations, we calculated for each ROI and each moment the valid cloud-free data fraction by dividing the valid pixels by all pixels within the ROI. We then only retained an area if for that moment it contained at least 95% of valid data from MODIS SWF and JRC-GSW derived water extent time series.

5.3.3 Comparison of water extent time series

We compared time series of surface water extent derived from MODIS SWF with those from JRC-GSW to see if the temporal dynamics of surface water extent obtained from the two datasets match with each other. Specifically, we calculated the monthly mean water extent from MODIS SWF for the period from February 2000 to October 2015, and compared it with JRC-GSW monthly water history. Two statistical measures were calculated: the Pearson's correlation coefficient (r) and the normalized mean signed deviation (NMSD):

$$NMSD = \frac{MSD}{Area_{mean}}$$
(5.1)

where Area_{mean} is the mean water area according to the JRC-GSW time series.

To assess if the agreement between JRC-GSW and MODIS-derived surface water results is related to the dynamics of the water body, a measure of relative variability (RV) in surface water extent was calculated for each water body as:

$$RV = \frac{Area_{max} - Area_{min}}{Area_{mean}}$$
(5.2)

where Area_{max}, Area_{min} and Area_{mean} are the maximum, minimum and mean water area according to JRC-GSW time series. We then plotted the distribution of *r* as a function of RV for different water size categories. We also combined the effect of both water size categories and RV on *r* in a single analysis. For this, we further divided all water bodies into six RV levels, *i.e.*, 0 < RV < 0.15, $0.15 \le \text{RV} < 0.3$, $0.3 \le \text{RV} < 1.0$, $1 \le \text{RV} < 2$, $2 \le \text{RV} < 5$ and $\text{RV} \ge 5$.

5.3.4 Comparison with water level data

To evaluate the performance of MODIS SWF for monitoring the dynamics of surface water, we also compared the water extent time series with water level for 14 large lakes (>100 km²) for which satellite altimetry data was available (see Section 5.2.3). Specifically, we calculated the Spearman rank correlation (ρ) between monthly mean water level and monthly water extent obtained from JRC-GSW and MODIS SWF, respectively. A rank correlation was used because relationships between water level and extent are non-linear and depend on bathymetry (Hayashi and van der Kamp 2000).

To obtain a temporal match between our MODIS SWF 8-day product and water level, in addition to the monthly comparison we also resampled the water level data (available

every 10 or ~35 days) to 8-day estimates using linear interpolation. There could be long time gaps between two water level observations due to missing data (Ricko *et al.* 2012). We thus further removed those interpolated data if the time gap is longer than three months. The same 14 lakes were included by Khandelwal *et al.* (2017) who mapped surface water area dynamics between 2000 and 2015 based on binary classification of MODIS 8-day 500 m resolution data. To assess if our product offers an improvement with respect to the one of Khandelwal *et al.* (2017) we calculated ρ also between water level and their 8-day time series, available at http://z.umn.edu/monitoringwaterRSE. To make a fair comparison, for both products ρ was calculated for the 2000-2015 time frame, matching the study by Khandelwal *et al.* (2017).

5.3.5 Characterization of the temporal variability of water extent

To characterize the temporal dynamics within the 18-year water extent time series derived from MODIS SWF, we created a subset that only included the dynamic water bodies, which were defined by applying a threshold to RV for each water size category. We then extracted three features including the strength of seasonality, strength of trend, and stationarity. These features were used in previous studies to characterize hydrological time series (Buma *et al.* 2016; Lafare *et al.* 2016). To extract the features, we first decomposed the water extent time series using the seasonal trend decomposition algorithm (STL: Cleveland *et al.* 1990), which is capable of accommodating missing values and is robust to noise. The STL decomposes a time series Y_t into seasonal (S_t), trend (T_t), and remainder (R_t) components, as:

$$Y_t = S_t + T_t + R_t \tag{5.3}$$

After we obtained the seasonal (S_t) and trend (T_t) series, we normalized them to [0, 1] to indicate the strength of seasonality and trend using the following equations:

$$F_{S} = \max(0, \ 1 - \frac{\operatorname{var}(R_{t})}{\operatorname{var}(S_{t} + R_{t})})$$
(5.4)

$$F_T = \max(0, \ 1 - \frac{\operatorname{var}(R_t)}{\operatorname{var}(T_t + R_t)})$$
(5.5)

where var stands for variance. A series with seasonal strength (F_s) close to 0 exhibits almost no seasonality, while a series with strong seasonality will have F_s close to 1 (Wang *et al.* 2006). This is similar for the strength of the trend component (F_T). We calculated the strength of seasonality and trend by using the "stl_features" package, which implements STL algorithm in R statistical software (R Core Team 2013), and used the suggestions of Cleveland *et al.* (1990) to set the parameter values. To calculate stationarity, the water extent time series were first linearly rescaled to [0, 1]. Then we calculated the annual mean from the rescaled water extent time series. The stationarity is calculated as the variance of annual mean (Wang *et al.* 2006).

Finally, based on these three characteristics (*i.e.*, strength of seasonality, strength of trend and stationarity) extracted from water extent time series derived from MODIS SWF, we organized the water bodies into groups with similar dynamic patterns using the K-means clustering algorithm (Aghabozorgi *et al.* 2015).

5.4 Results

5.4.1 Comparison of water extent time series

Based on MODIS SWF, we constructed temporal dynamics of surface water extent at 8day time interval for 340 water bodies spanning a time period from 2000 to 2017. For comparison with JRC-GSW, three water bodies (*i.e.*, Lake Urmia in Iran, Sebkhet Sidi Elhani in Tunisia and Sultan Marshes in Turkey) were removed from further analysis because for less than 10 individual months both datasets had overlap in terms of valid water extent retrievals, which was considered too few for a proper statistical basis.



Figure 5.2. Pearson's correlation (r) between monthly water extent derived from MODIS SWF dataset and from JRC-GSW as a function of water relative variability (RV) for different water size categories. Possible reasons for the low correlation for the points marked with a red number (1-5) are explored in the Discussion section of this paper.

Figure 5.2 plots the distribution of r between monthly water extent derived from MODIS SWF and those from JRC-GSW as a function of RV for different water size categories. There is a general increase in r when the water relative variability increases for all water size categories. This indicates that small variations in water extent are not similarly described by the two products, suggesting a lower accuracy in monitoring fluctuations of more stable water bodies. For water bodies with a larger RV, the temporal variability in monthly water extent obtained from both datasets had a good agreement with r generally above or near 0.8; this includes very small water bodies (< 1 km²). Examples of four very small water bodies (*i.e.*, Laguna Dulce in Spain, Laguna de la Nava de Fuentes in Spain, Complexe du bas Tahaddart in Morocco, and Livanjsko Polje in Bosnia and Herzegovina) are shown in Figure 5.3. These four water bodies are all Ramsar sites with an area less than 1 km² (i.e., mean water area according to JRC-GSW time series). Comparison of monthly water extent derived from both datasets shows a good agreement ($r \ge 0.79$) for these small water bodies. Note that several cases with small value of r are identified in Figure 5.2, each marked with a red number. These cases are caused by a number of factors which are discussed in the Discussion section.



Figure 5.3. Monthly water extent time series derived from MODIS SWF dataset showing a good agreement (r above or near 0.8) with those from JRC-GSW for four very small water bodies (< 1 km² mean JRC-GSW water extent): (a) Laguna Dulce, Spain; (b) Laguna de la Nava de Fuentes, Spain; (c) Complexe du bas Tahaddart, Morocco; (d) Livanjsko Polje, Bosnia and Herzegovina. The black polygon is the buffered region of interest (ROI) used for calculating the surface water area.

Table 5.2 shows that the average correlation increases when water variability increases, with the smallest *r* of 0.17 for static water bodies (RV < 0.15) and the largest *r* of 0.81 for highly dynamic water bodies ($RV \ge 5$). In addition, Table 5.2 shows that the effect of RV on *r* depends on the water size categories considered. This may be expected as small relative changes for larger water bodies (as expressed by RV) still correspond to substantial absolute changes, which can be captured by MODIS. Table 5.2 signals how much variability can be captured by our MODIS SWF product. For example, the reasonable correlation ($r \ge 0.5$) found for water bodies smaller than 1 km² when RV ≥ 1 implies that MODIS SWF resembles JRC-GSW for representing absolute variability of less than 1 km².

Table 5.2. Effects of size categories and water relative variability (*RV*) on the correlation (*r*) between monthly water extent obtained from MODIS SWF dataset and JRC-GSW. Correlation (*r*) is colour-coded with orange indicating values less than 0.5 and blue greater than or equal to 0.5. The number of water bodies is given in brackets.

Water relative		Size ca	ategories (kn	n ²)		Total
variability (RV)	0-1	1–5	5–20	20-100	>100	
RV<0.15		0.17 (8)	0.14 (8)	0.16 (12)	0.22 (11)	0.17 (39)
0.15≤RV<0.3	0.38 (7)	0.17 (8)	0.52 (6)	0.66 (6)	0.63 (14)	0.49 (41)
0.3≤RV<1	0.30 (18)	0.50 (26)	0.60 (26)	0.79 (14)	0.85 (6)	0.55 (90)
$1 \leq RV \leq 2$	0.53 (26)	0.71 (20)	0.90 (7)	0.89 (7)	0.98 (5)	0.70 (65)
$2 \leq RV \leq 5$	0.71 (25)	0.82 (16)	0.84 (7)	0.88 (9)	0.95 (2)	0.79 (59)
RV≥5	0.74 (19)	0.88 (10)	0.88 (7)	0.85 (7)		0.81 (43)
Total	0.56 (95)	0.59 (88)	0.62 (61)	0.67 (55)	0.62 (38)	

Examination of the NMSD compared to different water size categories shows that MODIS SWF and JRC-GSW have the largest difference for the estimation of very small water bodies (< 1 km²) (Figure 5.4). MODIS SWF tends to smaller values of water extent compared to JRC-GSW in most cases. This can be partly attributed to the underestimation of larger water fractions by MODIS SWF (Li *et al.* 2018a). Figure 5.4 also reveals that NMSD decreases when RV incre ases for water bodies larger than 100 km². However, the NMSD was surprisingly large for highly dynamic water bodies (RV \geq 1) for those in the 0–100 km² size categories, such as those cases marked with red numbers in Figure 5.4. These large NMSD can be caused by a number of factors which are discussed in the Discussion section.



Figure 5.4. Boxplot comparing the normalized mean signed deviation (NMSD) as a function of water relative variability (RV) for different water size categories. Possible reasons for the large NMSD for the points marked with a red number (3, 5–10) are explored in the Discussion section of this paper.

5.4.2 Comparison with water level data

Monthly water extent obtained from MODIS SWF shows moderately high correlation with both monthly water area derived from JRC-GSW ($r \ge 0.76$) and with monthly water level data ($\rho \ge 0.63$) for 10 dynamic large lakes (RV > 0.15 and size > 100 km²) (Table 5.3). Correlations were low for the four lakes with a small variation (RV ≤ 0.15) in surface water (both *r* and ρ are less than 0.4: see Table 5.3). MODIS SWF 8-day water extent shows moderate to high correlation ($\rho \ge 0.60$) with water level for these 10 dynamic lakes, suggesting that the MODIS SWF can accurately monitor the seasonal, intra-annual and long-term variability in surface water extent of those dynamic lakes. Table 5.3 also indicates that our MODIS SWF outperformed the Khandelwal *et al.* (2017) results in monitoring the dynamics of water body with larger correlations for the 10 dynamic lakes.

Jor 14 Water Doates. The Water	un sainoo	e oraerea from m r between	e targest to t monthly	<i>time series</i>	auve variaoun 8-dav	y (AV). time series	Source of
Lake Name	RV	MODIS and JRC- GSW	MODIS Vs. water level	JRC-GSW Vs. water level	MODIS Vs. K water level	handelwal <i>et al.</i> water level	Vs. water level
Razazza Lake	1.57	1.00	0.95	06.0	0.95	0.78	DAHITI
Qadisiyah Lake	1.05	0.99	0.96	0.95	0.95	0.78	LEGOS
Mosul Lake	0.53	0.91	0.82	0.97	0.84	0.66	GRLM10
Tshchikskoye Lake	0.50	0.90	0.80	0.80	0.81	0.48	GRLM10
Mingachevir Reservoir	0.31	0.95	0.84	0.94	0.85	0.80	LEGOS
Karakaya Baraji Reservoir	0.30	0.92	0.88	0.98	0.86	0.66	DAHITI
Tharthar Lake	0.29	0.97	0.94	0.96	0.90	0.86	GRLM10
Tsimlyansk Reservoir	0.17	0.86	0.74	0.80	0.71	0.52	DAHITI
Ataturk Lake	0.16	0.78	0.67	0.87	0.60	0.50	DAHITI
Assad Lake	0.16	0.76	0.63	0.76	0.61	0.08	DAHITI
Krasnooskol Lake	0.15	0.31	0.19	0.52	0.12	-0.48	GRLM10
Kremenshugskoye Reservoir	0.06	0.18	0.27	0.26	0.16	-0.24	GRLM10
Kakhovskoye Lake	0.06	0.01	-0.07	-0.08	-0.13	-0.42	GRLM10
Volgogradskoye Lake	0.05	-0.17	0.21	0.31	0.17	-0.36	GRLM10

Evaluation of a New 18-year MODIS-Derived Surface Water Fraction Dataset



Figure 5.5. (a) Google Earth imagery for Karakaya Baraji Reservoir, Turkey. (b) Comparison of monthly water area obtained from JRC-GSW and MODIS SWF with relative water level. (c) Time series of monthly water extent and monthly water level. (d) Comparison of 8-day water extent obtained from MODIS SWF with relative water level. (e) Comparison of 8-day water extent as detected by Khandelwal et al. (2017) with relative water level. (f) Time series of 8-day water extent and water level.



Figure 5.6. Same graphs as provided in Figure 5.5, but here showing the results for *Tshchikskoye Lake, Russia.*

Figure 5.5 and Figure 5.6 show two examples of the water extent and level variability based on the different data sources. The Karakaya Baraji Reservoir in Turkey has a compact shape. Figure 5.5c and Figure 5.5f show that the reservoir has large temporal variations, despite its RV of 0.30. High correlations were achieved for both the MODIS SWF monthly ($\rho = 0.88$) and 8-day water extent ($\rho = 0.86$), suggesting that the temporal variations of this lake were well captured by MODIS SWF, and better than by the Khandelwal *et al.* (2017) study ($\rho = 0.66$). The Tshchikskoye Lake in Russia has strong seasonal variations (Figure 5.6c and f) with a RV of 0.50. The JRC-GSW monthly water extent has a high correlation ($\rho = 0.80$) with water level data, but with many missing data in the time series (Figure 5.6c). The MODIS SWF monthly water extent also achieved a high correlation ($\rho = 0.80$) (Figure 5.6b), with the seasonal variations accurately captured (Figure 5.6c). The MODIS SWF 8-day water extent also greatly improved with respect to the Khandelwal *et al.* (2017) study ($\rho = 0.48$).

5.4.3 Characterization of water bodies by their temporal variability of water extent

We created a subset of 265 water extent time series by only selecting the dynamic water bodies (*i.e.*, $RV \ge 1$ for water bodies smaller than 1 km², $RV \ge 0.3$ for water bodies between 1 and 5 km², and $RV \ge 0.15$ for water bodies larger than 5 km²). Following the feature-based clustering, these 265 water bodies were separated into four different groups with different seasonal, inter-annual, and long-term dynamic patterns:

- (1) Seasonal dynamic pattern (45 water bodies): this group had a clear seasonal dry-wet cycle in the water extent time series, but there was no obvious long-term trend (Figure 5.7a). This type of water bodies is generally located in temperate climate regions where temperature and precipitation vary markedly across the year (Figure 5.7e). The water level in wet- and dry-season can rise and fall in concert with the water extent.
- (2) Seasonal and inter-annual dynamic pattern (105 water bodies): the second group also showed a strong seasonal pattern, but different from the first group it also showed a clear inter-annual variability (Figure 5.7b). This type of water bodies is common in temperate climates and in humid continental areas, typified by large seasonal and inter-annual temperature differences, but sometimes can occur in arid and semi-arid areas that experience regular drought periods.
- (3) Intermittent dynamic pattern (42 water bodies): the third group was intermittently flooded with rapid inundation and drought. Surface water is present for variable periods without detectable seasonal periodicity. Weeks, months, or even years may be between short periods of inundation. There was also no detectable long-term trend (Figure 5.7c). This type of water bodies typically occurs in arid areas (Figure 5.7e), where the periodic cycles of inundation and drought are primarily in response to extreme and irregular rainfall events. This pattern is also found in floodplains, such as those on the bank of the River Tisza in Hungary. They may be mostly dry outside seasonal high-water periods, during which swollen rivers overspill their banks.
- (4) Inter-annual and long-term dynamic pattern (73 water bodies): the fourth group showed an obvious inter-annual and long-trend (either increase or decrease) in surface water extent over long time without seasonal periodicity. Figure 5.7d shows the creation of a new dam in Italy (*i.e.*, the Alaco dam) and has been expanding over time. This type of water bodies is mainly distributed in dry summer temperate climate zones, but can also be found in arid areas. The long-term changes are likely influenced by climate change and human activities.



Figure 5.7. Example of four different dynamic patterns: (a) Seasonally dynamic pattern; (b) Seasonal and inter-annual dynamic pattern; (c) Intermittent dynamic pattern; (d) Inter-annual and long-term dynamic pattern. (e) Spatial distribution of water bodies with different dynamic patterns across different climate zones as mapped by Peel et al. (2007).

5.5 Discussion

This paper demonstrates that the MODIS SWF dataset can provide rich hydrological data for a large number of water bodies across a large spatial scale. The constructed time series provide detailed temporal information on intermittent, seasonal, inter-annual, and long-

Chapter 5

term changes in surface water extent. To our knowledge, this is the first such construction of wetland temporal dynamics for water bodies with varying sizes $(0.01-3100 \text{ km}^2)$ across a large region and at fine temporal resolution. Our work improves on previous efforts on surface water monitoring from coarse spatial resolution imagery (MODIS or similar) in regard to detect and monitor small water bodies. This is because MODIS SWF estimates surface water fraction instead of binary water presence/absence (Kaptue et al. 2013; Khandelwal et al. 2017; Sharma et al. 2015). Our surface extent time series match well with the relative water level variations and improve the correlation between MODISderived water area and relative water level when compared with previous MODIS-based water mapping by Khandelwal et al. (2017). The comparison with high spatial resolution data (i.e., JRC-GSW) showed that the MODIS SWF dataset can accurately construct surface water dynamics for many water bodies including very small ones (< 1 km²), as long as they show sufficient area variations (see Figure 5.2, Figure 5.3 and Table 5.2). MODIS SWF was found to be less effective for more stable water bodies that have limited temporal variation, likely due to the effect of mixed pixels that may be insufficiently captured by MODIS (Klein et al, 2017; Busker et al, 2018). More importantly, our work complements the high spatial resolution data (i.e., JRC-GSW) by allowing for the detection of abrupt changes, inundation and recession processes over short time periods (Figure A5.1–A5.3), and for the better representation of seasonal cycle and long-term variability (Figure 5.8 and Figure 5.9)

The fact that we compared datasets (*i.e.*, MODIS SWF versus Landsat-derived JRC-GSW) implies that errors and uncertainties may occur in either of the two datasets. Table 5.4 gives a summary of all illustrative examples in this paper (including Appendix) together with their likely error sources and uncertainties. These examples represent the water bodies highlighted with red numbers in Figure 5.2 and Figure 5.4.

Tab the	le 5.4. A summary of all illus water bodies highlighted with	trative exam red number	ıples demon in Figure 5	ıstrating th 5.2 and Fig	he sources of erro zure 5.4.	or and uncertainties. These exampl	es represents
Case NO.	Lake Name	Mean water extent (km ²)	Relative variability (RV)	Pearson's correlation (r)	Normalized mean signed deviation (NMSD)	Sources of error and	Figure
1	Sebkhet El Hamiet, Algeria	0.38	12.85	0.41	0.03	Uncertainty in dates of JRC-GSW	Figure A5.1
7	Azraq Oasis, Jordan	0.44	23.40	0.53	0.29	Uncertainty in dates of JRC-GSW	Figure A5.2
б	Chott Oum EI Raneb, Algeria	2.02	3.49	0.48	-0.46	Ommision and commision errors in both JRC-GSW and MODIS SWF	Figure A5.8
4	Albufera de Valencia, Spain	33.92	1.39	0.55	0.16	Uncertainty in dates of JRC-GSW	Figure 5.8
S	Chott El Jerid, Tunisia	26.39	11.72	0.61	0.31	Uncertainty in dates of JRC-GSW	Figure A5.3
9	Meke Maar, Turkey	0.34	1.68	0.87	1.49	Impact of salt crust	Figure 5.9
7	Csongrád-Bokrosi Sóstó sodic- alkaline pans, Hungary	0.01	11.71	0.74	4.97	Impact of emergent vegetation	Figure A5.5
~	Réserve Intégrale du Lac Tonga, Algeria	3.74	5.76	0.81	1.31	Impact of emergent vegetation	Figure A5.4
6	La Réserve Naturelle du Lac des Oiseaux, Algeria	8.30	6.99	0.84	1.04	Impact of saturated soil	Figure A5.6
10	Chott de Zehrez Gharbi, Algeria	37.40	3.63	0.93	-1.08	Ommision in MODIS SWF	Figure A5.7

132
Chapter 5

One source of uncertainty is that the JRC-GSW monthly water dataset does not include exact dates of the Landsat observations used in that month, and may not be best represented by the monthly mean value that we calculated from MODIS SWF. Particularly for lakes with strong dynamics within a month, this uncertainty increases. Figure 5.8 shows an illustrative example of Albufera de Valencia, a Ramsar wetland located in eastern Spain. It is composed of the Albufera Lake and surrounding rice fields with an average water area of 34 km² according to the JRC-GSW time series. The Landsat imagery (Figure 5.8a) revealed that this rice fields flooded twice in 2014; once around mid-May to June during rice sowing and start of growth, then following harvesting in September-October a second flooding took place in November-December for nutrient mineralization of the harvested rice fields (Campos-Taberner et al. 2018; Doña et al. 2015). Comparison of monthly surface area time series derived from JRC-GSW and MODIS SWF showed a moderate level of agreement (r = 0.55) for this wetland. MODIS SWF had a larger water extent than JRC-GSW, as shown as the red boxes in Figure 5.8b and 8c, corresponding to May and June when water area experienced rapid changes. These changes may not well captured by JRC-GSW due to the coarse temporal resolution and limited valid Landsat observations. However, both the monthly (Figure 5.8c) and 8day (Figure 5.8d) time series from the MODIS SWF clearly captured the two flooding times for each year. Similarly in Figure 5.S1, Figure 5.S2 and Figure 5.S3 for Sebkhet El Hamiet in Algeria, Azraq Oasis in Jordan and Lake Chott El Jerid in Tunisian, respectively, strong differences between two datasets correspond to months with rapid water changes in response to erratic rainfall events (Ben Abdallah et al. 2018; Bryant 1999). These examples demonstrate that a low to moderate correlation between both datasets does not necessarily imply a poor accuracy of MODIS SWF, but rather that MODIS allows to capture rapid temporal changes.



Evaluation of a New 18-year MODIS-Derived Surface Water Fraction Dataset

Figure 5.8. (a) One year Landsat imagery (January 2014 to December 2014) for Albufera de Valencia, Spain. (b) Comparison of monthly water extent (km²) derived from JRC-GSW and MODIS SWF. (c) Monthly water extent (km²) time series from 2000 to 2015. (d) 8-day water extent (km²) time series derived from MODIS SWF.



Figure 5.9. Illustration of dynamics of a small water body (i.e., Meke Maar wetland, a Ramsar site in Turkey). (a) Comparison of monthly water extent (km²) time series obtained from MODIS SWF and JRC-GSW from 2000 to 2015. (b) Landsat imagery acquired on April 20, 2001. (c) GSW water maps for April 2001. (d) MODIS SWF maps for April 23, 2001.

There are several sources of commission errors that may not be completely solved by MODIS SWF, predominantly due to salt presence (Figure 5.9) and emergent vegetation (Figure A5.4–A5.5). Such commission errors lead to an overestimation of water extent by MODIS SWF. Figure 5.9 provides an illustrative example of a small water body in Turkey (Kurt *et al.* 2013) where salt crystallization occurred during the dry season. The SWF misclassified salt as water due to spectral confusion (Li *et al.* 2015), resulting in overestimation of the water area by MODIS SWF. For example, The JRC-GSW estimated the surface water extent for this water body (derived from Landsat images in April 2001) to be 0.6 km² against 1.6 km² for MODIS SWF (Figure 5.9c and d). Despite this large bias (NMSD =1.49), MODIS SWF clearly captured the water body's seasonal variations and its reduction in size between 2008 and 2014 (Figure 5.9a).



Figure 5.10. Examples showing JRC-GSW water maps have omission errors over time for Szeged-Fehértó artificial fishponds, Hungary. The top row shows the original Landsat imagery and the bottom row the corresponding JRC-GSW water maps.

Moreover, we found that JRC-GSW also has considerable omission errors for some locations. Figure 5.10 shows an example where omission errors occur for multiple dates (for which we only show four). Moreover, the heavily sediment-laden seasonally occurring waters in Africa were also poorly represented as already indicated by Pekel *et al.* (2016) (see Figure A5.8). These errors result in a large underestimation of water area for these dates by JRC-GSW. The temporal dynamics were thus not accurately captured by JRC-GSW, hence further illustrating that low correlations between both products do not necessarily imply a poor accuracy of MODIS SWF. In this situation, linking the estimates with temporal *in situ* observations would be desirable.

Both JRC-GSW and MODIS SWF contain missing values due to the presence of clouds, cloud shadows, snow, aerosol and sensor-related issues (Klein *et al.* 2017). If the missing data are inside the water-covered area of the ROI, JRC-GSW and MODIS SWF are likely to underestimate the actual water extent. In this study, we thus applied a 5% missing data threshold for the calculation of water area (see Section 5.3.2). Locations with poor quality MODIS pixels (*e.g.*, due to cloud or aerosol effect) will often return no data when exceeding the 5% threshold, resulting in a sparse time series of MODIS SWF. Examples are Sebkhet Sidi Elhani in Tunisia, Lake Urmia in Iran, and the Sultan Marshes in Turkey. Future efforts may consider incorporating temporal interpolation techniques (Klein *et al.* 2017), or combining radar remote sensing which has the advantage of collecting data under poor weather or atmospheric conditions (Brisco 2015; Montgomery *et al.* 2018; O'Grady *et al.* 2014), for areas with missing data to ensure a gap-free reconstruction of inland water variability.

The constructed time series of surface water extent have potential to benefit a large number of applications. For example, the water extent can help to estimate a series of hydrological and climatological parameters. It can be used to estimate river width (Yamazaki et al. 2014), river discharge (Bjerklie et al. 2018; Pan et al. 2016), and water volume (Ogilvie et al. 2018b). This would be particularly useful for data-poor and ungauged wetlands and catchments. Several studies have estimated lake surface temperatures (Kettle et al. 2004) and evaporation using water area (Zhan et al. 2019; Zhang et al. 2017). The time series of water extent may be used as a monitoring tool for analyzing hydrologic extremes such as floods and droughts, detecting abnormal changes of wetland hydrology, capturing short-duration events, and identifying newly-formed and disappearing water bodies. These dynamics may be driven by a wide range of natural (e.g., climate, topography, geology, geomorphology and pedology), and human factors (maintenance, irrigation, wetland conversion). Because the role of a specific factor may differ substantially per location and wetland type, further studies may use our dataset to investigate possible drivers of wetland degradation and other changes. Most of the water bodies analyzed in this study are wetlands of international importance (Ramsar sites). They provide a critical habitat for many sensitive species in Mediterranean environments. The long monitoring information reported here may also help to improve species distribution modeling (Bradley and Fleishman 2008) as well as understanding changes in species numbers and distribution, and thus lead to improved management and conservation of biodiversity.

5.6 Conclusion

This paper highlights that our MODIS-derived surface water dataset is valuable for monitoring the dynamics of Mediterranean water bodies. Based on this dataset, we constructed detailed temporal dynamics of surface water extent for 340 water bodies with varying size (0.01–3100 km²). For large water bodies (>100 km²) for which altimetry water level data was available, time series of water extent derived from MODIS SWF dataset showed good correlation with water levels. Our work also showed that the correlation between MODIS-derived water area and relative water level improved when compared with previous MODIS-based water mapping by Khandelwal et al (2017). For water bodies smaller than 100 km², SWF proved to be able to effectively monitor their dynamics as long as their area variability was sufficient. Our work and dataset can therefore be applied to better understand the pattern of hydrological dynamics and as a monitoring tool to identify abnormal changes to seasonal dry-wet cycles, capturing short-duration events and identifying newly formed water bodies. As such it may prove a useful tool for improved water management and biodiversity conservation. The 8-day

time series of water extent for the 340 water bodies, as derived from MODIS SWF, can be access through https://surfdrive.surf.nl/files/index.php/s/jlVJcDNLFFhIdKU (password: waterextent_ts), and will be open available publicly through <u>https://easy.dans.knaw.nl/ui/home</u> at the time of publication.

Appendix

This Appendix contains additional illustrative examples demonstrating the error sources and uncertainties in the comparison of MODIS SWF and JRC-GSW.

Figure A5.1 shows the water extent time series for Sebkhet El Hamiet, a Ramsar site in Algeria. The water area in Sebkhet El Hamiet is highly dynamic (relative variability of 12.85). Comparison of monthly surface extent time series derived from JRC-GSW and MODIS SWF showed a poor correlation (r = 0.41), which is due to large differences between the two datasets for months with rapid changes in water extent. The red boxes in Figure A5.1a, b correspond to September 2007 and August 2015. No surface water was detected by JRC-GSW for September 2007 because it corresponded to Landsat imagery from 18 September 2007 when the area had not yet flooded. The corresponding MODIS SWF maps for September 2007 clearly show the rapid increase in water area at the end of that month (Figure A5.1c). For August 2015 instead, JRC-GSW had larger water extent values than MODIS SWF. The reason is that JRC-GSW presents the monthly maximum when multiple images are available for a month, while for MODIS SWF we calculated the monthly average.



Figure A5.1. (a) Comparison of monthly water extent (km²) derived from JRC-GSW and MODIS SWF for Sebkhet El Hamiet, Algeria. (b) Monthly water extent (km²) time series from 2000 to 2015. Comparison of Landsat results (Original Landsat imagery and JRC-GSW water maps) and MODIS SWF maps for September 2007 (c) and for August 2015 (d).



Evaluation of a New 18-year MODIS-Derived Surface Water Fraction Dataset

Figure A5.2. (a) Comparison of monthly water extent (km²) derived from JRC-GSW and MODIS SWF for the Azraq Oasis, Jordan. (b) Monthly water extent (km²) time series from 2000 to 2015. Comparison of Landsat results (Original Landsat imagery and JRC-GSW water maps) and MODIS SWF maps for April 2005 (c) and for April 2006 (d).

Figure A5.2 gives an example for the Azraq Oasis, a Ramsar site in Jordan. The water area in the Azraq Oasis is highly dynamic (relative variability of 23.4). Comparison of monthly surface extent time series derived from JRC-GSW and MODIS SWF showed a relatively poor correlation (r = 0.53). Also in this case, the large differences between the two datasets corresponded to months with rapid water changes (*e.g.*, April 2005 and April 2006: highlighted with red boxes in Figure A5.2a, b).



Figure A5.3. (a) Comparison of monthly water area (km²) derived from JRC-GSW and MODIS SWF for Chott El Jerid, Tunisia. (b) Monthly water area (km²) time series from 2000 to 2015. (c) MODIS SWF maps for May 2007, white color represents missing data. (d) Landsat imagery and JRC-GSW water maps (bottom right).

Figure A5.3 shows the same comparison for Chott El Jerid, a Ramsar site in south-central Tunisia, located in the Sahara desert. It is an ephemeral salt lake that floods in response to occasional rainfall (Ben Abdallah *et al.* 2018; Bryant 1999). The water area in Chott El Jerid is highly dynamic (relative variability of 11.7). For this site, JRC-GSW and MODIS SWF showed a moderate level of agreement (r = 0.61). This can be attributed largely to one outlier (highlighted in red in Figure A5.3a), which corresponds to May 2007 (Figure A5.3b). For May 2007, MODIS SWF has four valid maps showing the decrease in water area of Chott El Jerid (Figure A5.3c), while JRC-GSW relies on a single Landsat image from 31 May 2007 when water almost retreated (Figure A5.3d). This resulted in a much larger water extent for MODIS SWF as compared to JRC-GSW (NMSD = 0.31).

Evaluation of a New 18-year MODIS-Derived Surface Water Fraction Dataset



Figure A5.4. (a) Monthly water area (km²) derived from JRC-GSW and MODIS SWF from 2000 to 2015 for Réserve Intégrale du Lac Tonga, Algeria. (b) Landsat imagery acquired on August 11, 2010. (c) JRC-GSW water maps for August 2010. (d) MODIS SWF maps for August 13, 2010.



Figure A5.5. (a) Monthly water extent (km²) time series obtained from MODIS SWF and JRC-GSW from 2000 to 2015 for Csongrád-Bokrosi Sóstó sodic-alkaline pans, Hungary. (b) Landsat imagery acquired on October 11, 2010. (b) JRC-GSW water maps for August 2010. (c) MODIS SWF maps for August 13, 2010.



Figure A5.6. (a) Monthly water extent (km²) time series obtained from MODIS SWF and JRC-GSW from 2000 to 2015 for La Réserve Naturelle du Lac des Oiseaux, Algeria. (b) Landsat imagery acquired on May 10, 2011. (c) JRC-GSW water maps for May 2011. (d) MODIS SWF maps for May 9, 2011.

It can be challenging to map surface water extent when a water body is covered by emergent vegetation or in floodplains where the soil is saturated. Figure A5.4 shows an illustrative example for Réserve Intégrale du Lac Tonga, a Ramsar site in Algeria. This wetland is under anthropogenic threat, resulting in great reductions of surface water area. MODIS SWF tends to give larger values in summer (May to September) when vegetation is emerged (Figure A5.4), which further results in a large NMSD of 1.31. Similarly, MODIS SWF gave larger values for Csongrád-Bokrosi Sóstó sodic-alkaline pans, Hungary (Figure A5.5), and for La Réserve Naturelle du Lac des Oiseaux, Algeria (Figure A5.6) in summer months. Despite the large deviation, MODIS SWF clearly captured the seasonal variations for these water bodies (Figure A5.4a, Figure A5.5a and Figure A5.6a).



Figure A5.7. (a) Monthly water extent (km²) time series obtained from MODIS SWF and JRC-GSW from 2000 to 2015 for Chott de Zehrez Gharbi, Algeria. (b) Landsat imagery acquired on September 22, 2015. (c) JRC-GSW water maps for September 2015. (d) MODIS SWF maps for September 22, 2015.

MODIS SWF might result in smaller water area values due to the underestimation of larger water fractions by MODIS SWF (Li *et al.* 2018a). Figure A5.7 shows an example for Chott de Zehrez Gharbi, a Ramsar wetland in Algeria. This lake was complete covered with surface water in September (Figure A5.7b, c) while MODIS SWF gave an estimation near 80% for most pixels (Figure A5.7d). Further work could consider re-assigning pixels with more than 80% surface water to 100% water.



Figure A5.8. (a) (b) Comparison of monthly surface extent time series derived from JRC-GSW and MODIS SWF for Chott Oum EI Raneb, Algeria. (c) Omission errors at multidates in JRC-GSW. (d) Commission errors in JRC-GSW. (e) Omission errors in MODIS SWF.

Figure A5.8 shows the comparison results for Chott Oum EI Raneb, a Ramsar site in Algeria. The correlation coefficient between monthly surface extent time series derived from JRC-GSW and MODIS SWF is 0.48. The errors and uncertainties are propagated from both datasets. For example, JRC-GSW had omission errors at multi-dates (Figure A5.8c) and commission errors from wet saline soil at some time point (Figure A5.8d). This might due to the lack of training data as already indicated by Pekel *et al.* (2016). MODIS SWF also tends to underestimate the water area (Figure A5.8e), which might due to the fact that the surface water is shallow.

Chapter 6

Synthesis

6.1 Introduction

Terrestrial surface water plays a significant role in biogeochemical cycles, provides valuable freshwater resources, and supports biodiversity (Chahine 1992; Tranvik et al. 2009). Changes in surface water strongly affect the delivery of important ecosystem services such as water supply, carbon storage, biodiversity, and climate regulation. Water extent has been identified as one of the Terrestrial Essential Climate Variables to support climate change assessment and policy development (GCOS 2011), and its changes are also an indicator to assess sustainable development of water resources and biodiversity loss (CBD 2010; Griggs et al. 2013). Remote sensing provides unique opportunities to detect and monitor surface water bodies at different spatial and temporal scales. While much progress has been made recently with global Landsat-based surface water assessment (e.g., Donchyts et al. 2016; Pekel et al. 2016), such assessments can have large spatial and temporal gaps due to both the limited number of acquisitions and persistent cloud cover, preventing accurate assessments of water extent variability. Therefore, the aim of this thesis was to improve long-term mapping and monitoring dynamics of surface water at fine temporal resolution using high-frequency optical remote sensing data provided by the Moderate Resolution Imaging Spectroradiometer (MODIS), while accounting for small (<1 km²) and dynamic water bodies.

In this dissertation, the applicability of high-frequency MODIS was explored for monitoring surface water dynamics at a range of spatial scales, including the local scale (*i.e.*, a single water body in a local region in Chapter 2), the regional scale (*i.e.*, two regions in Spain in Chapter 3), and for the whole Mediterranean region in Chapters 4 and 5. The main contribution of this work relates to four aspects: (1) it showed the potential and limitations of spectral information for monitoring hydrological dynamics; (2) it demonstrated advantages of sub-pixel surface water fraction mapping for quantifying water extent variability in small water bodies; (3) it analysed the robustness of a machine learning approach to application at larger spatial scale; and (4) application of the approach resulted in a new long and dense time series of water extent for the whole Mediterranean area, which can be used for analysing and monitoring surface water dynamics.

The following sections discuss these four aspects in detail. Subsequently, Section 6.6 explores potential applications of the approach and the surface water dataset developed in this thesis, and Section 6.7 provides a future outlook on monitoring surface water dynamics.

6.2 The potential and limitations of spectral information for monitoring hydrological dynamics

Existing methods to extract surface water from optical imagery are mostly based on the fact that water has strong absorption in near infrared (NIR) and shortwave infrared (SWIR) spectral regions. Individual NIR and SWIR bands have consequently been used for water delineation (e.g., Mondejar and Tongco 2019; Wolski et al. 2017). Spectral indices, combining two (e.g., Gao 1996; McFeeters 1996; Xiao et al. 2002a; Xu 2006) or more multispectral bands (e.g., Crist 1985; Feyisa et al. 2014; Fisher et al. 2016; Wang et al. 2018a) from NIR/SWIR and visible spectral regions can enhance the information on surface water presence. These spectral indices provide a simple and effective way for detecting, mapping and monitoring surface water (Borro et al. 2014; Chipman and Lillesand 2007; Huang et al. 2018a; Zhou et al. 2017). In addition, Pekel et al. (2014) introduced a Hue-Saturation-Value (HSV) colour space transformation applied on SWIR, NIR, and Red reflectance to enhance the contrast between water bodies and other land cover types. However, a robust detection of water surface is challenging due to the great spatial and temporal variability in the water spectral signatures caused by variations in vegetation cover, material of the bathymetric surface, dissolved organic matter and suspended sediment (Pekel et al. 2014).

In Chapter 2, several indices for characterizing the temporal variability in the hydrology of a shallow seasonally flooded wetland in Spain were compared. The tested indices included the Normalized Difference Wetness Index (NDWI: McFeeters 1996), the Modified Normalized Difference Wetness Index (MNDWI: Xu 2006), the Normalized Difference Wetness Index (LSWI_{B5}: Gao 1996), the Land Surface Water Index (LSWI_{B6}: Xiao et al. 2002a), the Normalized difference vegetation index (NDVI: Tucker 1979), the Tasseled Cap Wetness Index (TCWI), and the Tasseled Cap Brightness Index (TCBI). Results showed that the hydrological variability (expressed by variations in water level) is strongly negatively correlated to TCBI (r = -0.87) for the 10-year analysis period. It also gave consistently good results for wet and dry years, and for areas characterized by different inundation characteristics (Table 2.4-2.5, Figure 2.5-2.6). This is because water strongly absorbs radiation across the visible and infrared portion of the spectrum, and consequently water presence will cause TCBI to decrease as it is a linear combination of spectral bands with positive coefficients. Therefore, TCBI is sensitive to all types of wetness changes including open water, soil moisture and water content in the vegetation. TCWI's correlation with water level data was much poorer for the wet years (r = 0.53)

Synthesis

than for the dry years (r = 0.92), likely due to larger TCWI values caused by salt deposition on the soil surface in wet years when more water evaporates. The two-band index NDWI, a combination of green and NIR spectral bands, gave a strong correlation (r = 0.84) with water level only for areas that are frequently covered with water, but not for surrounding areas. The two-band indices that use longer wavelengths located in the SWIR (*i.e.*, MNDWI, LSWI_{B5} and LSWI_{B6}) showed stronger water level-water index correlations for areas never covered by surface water (*e.g.*, r = 0.85 for MNDWI, and r = 0.86 for LSWI_{B6}), suggesting these indices are sensitive to water content in soil and vegetation.

In Chapter 3 and 4, the utility of spectral information, including MODIS bands, spectral indices, colour space transformation, and temporal characteristics of spectral indices (*i.e.*, the annual mean, minimum, maximum, standard deviation, and coefficient of variation per year), for estimating and monitoring water area was demonstrated. The spectral information was used as predictor variables within machine learning algorithms in order to estimate surface water area at sub-pixel level. Analysis of the relative importance of these variables showed that the NIR band is the most important predictor for estimating and monitoring sub-pixel surface water fraction, which is consistent with Sun *et al.* (2012). The fact that NIR gave a higher importance than SWIR may be explained because SWIR relates to many types of wetness including open water, and water content in soil and vegetation (findings from Chapter 2), which could result in misclassification while detecting surface water.

Results also showed that the temporal characteristics of water indices are important variables for creating rule conditions that split up the data into various sub-groups when predicting sub-pixel water fraction (Figure 3.6), hence help improve model prediction accuracy (Table 3.4). This can be because temporal characteristics are effective in partitioning the samples into subsets with homogeneous environmental and temporal characteristics, whereby consequently a more effective relationship can be built for each subset.

One limitation of only using spectral information for detecting and monitoring water is that other land surfaces such as volcanic materials, and shadows induced by clouds and terrain and buildings, have a similar spectral signature to water, which can lead to falsely claiming that water is detected (Pekel *et al.* 2016). These errors can be accounted for by using masks derived from auxiliary data (Klein *et al.* 2017; Pekel *et al.* 2016). For

example, in Chapter 4, the Digital elevation model (DEM) data were used to eliminate the confusion caused by terrain shadows, and land use map to eliminate building shadows.

6.3 Advantages of sub-pixel surface water fraction mapping techniques for the quantification of small water bodies

Detection of objects with remote sensing imagery is usually confined to objects with at least the size of the sensor's spatial resolution. Previous studies on surface water mapping with MODIS generally applied a binary land/water classification method, which did not allow for detection of smaller-sized water bodies. For example, Khandelwal *et al.* (2017) used MODIS 500 m time series data to map and monitor surface water extent of lakes across the globe, but restricted their application to water bodies larger than ten MODIS pixels (*i.e.*, > 2.5 km²). However, globally a large fraction of the surface water bodies are of small size (*i.e.*, 0.1-1 km²) (Downing *et al.* 2006; Verpoorter *et al.* 2014). Consequently, binary classification using MODIS may underestimate the global surface water extent due to large omission errors.

To overcome the shortcoming of coarse spatial resolution and incorporate small water bodies, either finer spatial resolution data are needed, or alternatively an effective method is required to effectively estimate sub-pixel water fraction (*i.e.*, the percentage of surface water within a single grid cell). The latter approach was explored in Chapters 3 and 4. This was achieved by the use of machine learning approaches that incorporated MODIS spectral information, its temporal characteristics, and topographic information. Results showed that the sub-pixel water fractions could be accurately derived with a high estimation accuracy ($R^2 > 0.90$) when compared to previous studies on MODIS-derived water fraction estimates, such as by Weiss and Crabtree (2011) ($R^2 = 0.63$), and Guerschmann *et al.* (2011) ($R^2 = 0.70$).

Accurate assessment of sub-pixel surface water fraction further allows to accurately estimate the area and dynamics of small water bodies (Chapter 5). Selected examples (*e.g.*, Figure 3.7 and Figure 5.3) for small water bodies ($< 1 \text{ km}^2$) show that the temporal variability in water extent derived from MODIS has a good agreement (*r* above or near 0.8) with temporal variability obtained from Landsat reference data.

6.4 Robustness of the machine learning approach for application to large regions

Because the spectral characteristics of surface water vary in space and time, machine learning approaches need to learn these spatially and temporally varying relationships between remote sensing derived information and surface water presence for different regions and at different moments in time (Muster et al. 2013). For example, Weiss and Crabtree (2011) derived separate region- and time-specific equations to estimate the percent surface water from MODIS; Guerschmann et al. (2011) developed a logistic regression model for water fraction mapping on the Australian continent; and Muster et al. (2013) used a simple linear regression model to estimate surface water fraction in three Arctic tundra wetlands. Each of those models was different because they were empirically derived for specific study regions. With the growing need for global and real-time monitoring, there is a high demand for efficient, flexible and scalable algorithms that can deal with the spatial and temporal heterogeneity in surface water characteristics, while maintaining the local representativeness and accuracy. Machine learning algorithms such as regression tree models can predict a single continuous response variable from one or more explanatory variables by recursively partitioning a data set into more homogeneous subsets, thus can deal with strong nonlinear situations.

Chapter 3 evaluated the robustness and transferability of machine learning approaches (e.g., random forest and cubist rule-based regression model) to different geographic and climatic zones. This was tested for two regions in Spain with contrasting environments and climatic characteristics. Two types of models were created: (1) unique local models for each region; (2) a global model for both regions. The local models were built using training data collected from a single region while the global model was based on the combined training dataset from both regions. Results showed that the performance of local models was good when applied to the regions for which models were trained (R^2 above 0.81 for all local models). However, when applying the local models developed for one region to the other region and vice versa, the accuracy was greatly reduced (Table 3.4). When combining training data collected from the two regions in a single global model, the accuracies were comparable to the local models (Table 3.4). These findings indicate that a model developed for mapping surface water fraction in one region cannot easily be transferred to other regions. Instead of building a specific model per region, it was shown in this thesis that a single global model can provide accurate results as long as the training data comprise the various environmental conditions encountered across the whole spatial domain. In this sense, the approach for constructing a global model can be expanded effectively to larger spatial scales.

Chapter 4 successfully scaled the machine learning approach developed in Chapter 3 to a large region (*i.e.*, Mediterranean region) while retaining a similar high accuracy. As shown in Figure 3.4 and Table 4.4, the mapping accuracies for the Mediterranean region are similar to the accuracies obtained in Chapter 3 for the two regions in Spain ($R^2 = 0.91$, RMSE = 11.05%, MAE = 7.67% for the Spanish regions vs. $R^2 = 0.91$, RMSE = 11.41%, MAE = 6.39% for the Mediterranean). This indicates that machine learning algorithms are robust tools for mapping water fraction at large spatial scales within heterogeneous environments. A prerequisite for successfully upscaling machine learning approaches to large spatial scales is a high-quality training dataset. Collection of training data is a labour-intensive and time-consuming process, especially when developing a global training dataset. In Chapter 4, the historical monthly global surface water dataset developed by the European Commission Joint Research Centre (Pekel *et al.* 2016) proved to be an efficient basis for building a reliable training dataset. Considering that GSW is globally available, we are confident that our approach can be scaled globally.

6.5 A new long and dense time series of water extent for monitoring surface water dynamics

Surface water extent can exhibit large variability due to natural processes and human interventions (Pekel *et al.* 2016). The changes can affect ecosystem functioning, species distributions and composition (Koning 2005; Robledano *et al.* 2010), and influence climate change (Degu *et al.* 2011; Ekhtiari *et al.* 2017). Waterbodies show a wide range of dynamic patterns in relation to the frequency of inundation and duration of standing water, such as intermittently and temporary dynamic water, seasonal water, semi-permanent and permanent water, and artificially flooded water (Cowardin *et al.* 1979). The 8-day repeat coverage is considered to be a minimum for effectively capturing water bodies with short hydroperiods (Guerschmann *et al.* 2011; Wulder *et al.* 2016). Existing fine spatial resolution datasets, such as developed by Pekel *et al.* (2016), cannot accurately assess seasonal and long-term changes especially in regions with persistent cloud cover. Therefore, frequent and long-term monitoring of surface water is required to fully understand the seasonal and long-term dynamic patterns of the water cycle.

In Chapter 4, a new dense 18-year time series water fraction dataset (SWF) was developed for the whole Mediterranean at 500 m resolution and 8-day interval, which can be

expanded to daily frequency using the model developed in this chapter. Based on MODIS SWF, we constructed a time series of surface water extent with an eight-day time step for hundreds of Mediterranean wetlands in Chapter 5. The derived temporal patterns of surface extent correspond strongly with *in situ* or radar altimetry-derived water level variations and with the temporal dynamics derived from high spatial resolution data. Importantly, these dense time series allow for the detection of abrupt changes and inundation and recession processes over short time periods (*e.g.*, rapid flooding due to precipitation), as well as for an improved representation of seasonal (*e.g.*, cycle rice field flooding and drying), and long-term variability (See Figure 5.7–5.9, Figure A5.1–5.8) as compared to the Landsat-derived GSW dataset. In this sense, the MODIS-based surface water monitoring approach developed in this thesis allows to fill the long temporal gaps caused by persistent cloud cover and limited number of acquisitions of fine spatial resolution water datasets, and can improve our understanding of surface water dynamics.

6.6 Potential applications of the new surface water dataset

This thesis demonstrated an approach for estimating sub-pixel surface water fraction from coarse resolution MODIS data. The robustness and transferability of the regression tree models suggest that the approaches can be applied globally. The approach takes use of spectral information from visible to shortwave infrared spectral regions, which can potentially transferred to other optical remote sensing data with similar spectral resolution such as from NOAA/AVHRR, Suomi-NPP/VIIRS, Landsat/MSS, Landsat/TM, Landsat/ETM+, Landsat/OLI, Sentinel-2/MSI, and Sentinel-3/OLCI etc.

In this thesis, a dense 18-year time series of surface water fraction (SWF) for the Mediterranean region has been developed, with the ability to monitor many small-sized (*e.g.*, smaller than a 500x500 m MODIS cell) and highly dynamic water bodies. This new dataset has potential to benefit a large number of hydrological and ecological applications.

MODIS SWF can be used as a monitoring tool for analyzing hydrologic extremes such as floods and droughts, detecting abnormal changes of wetland hydrology, capturing short-duration events, identifying newly-formed and disappearing water bodies, and estimating global water loss. Assessing the variability in water extent can help to estimate a series of hydrological, climatological and biogeochemical parameters. For example, water extent can be used to estimate river width (Yamazaki *et al.* 2014), river discharge (Bjerklie *et al.* 2018; Pan *et al.* 2016), and water volume (Busker *et al.* 2019; Cael *et al.*

2017; Duan and Bastiaanssen 2013; Tong *et al.* 2016). This is particularly useful for datapoor and ungauged wetlands and catchments. Lake surface temperature and evaporation are key climatological parameters which are closely related to local and regional air temperatures (Livingstone 1999; Straile *et al.* 2003). Several studies have used water extent to estimate lake surface temperatures (Becker and Daw 2005; Kettle *et al.* 2004) and evaporation (Zhan *et al.* 2019). Lakes and reservoirs play an important role in the biogeochemical cycles of carbon as their sediments sequester organic carbon, and as such reduce greenhouse gas emissions (Tranvik *et al.* 2009). Therefore the size and distribution of global lakes and ponds, including very small ponds, is a critical input for developing global estimates of carbon burial in and efflux from aquatic ecosystems (Bastviken *et al.* 2004; Holgerson and Raymond 2016; Tranvik *et al.* 2009; Wik *et al.* 2016).

Such new long-term information on surface water as developed in this thesis can be used for biodiversity and conservation applications. For example, it can be used to improve species distribution modeling (Bradley and Fleishman 2008; Redfern *et al.* 2005; Shannon *et al.* 2009), as well as for mapping and monitoring species population (Mathieu *et al.* 2006). It can also help for understanding changes in water-dependent species distribution, abundance and richness in relation to the spatial and temporal variability of surface water, the inundation frequency, duration and intensity. For example, many migratory species rely on the timing and quantity of available water during their migration and breeding period (Cezilly *et al.* 1995; Hamasaki *et al.* 2013). With this new high frequency time series dataset, the distribution, quantity, timing, and duration of water resources can be accurately estimated, which may further provide an important tool to better explain and predict timing, abundance and patterns of animal migrations.

Closely monitoring hydrological variability is important for understanding how climate change and human actions affect the dynamics of surface water, or may affect these in the future (*e.g.*, Tulbure and Broich 2019; Zhang *et al.* 2019). For example, a tendency towards shorter annual hydroperiods has been observed for seasonal water at local scale due to aquifer exploitation (Diaz-Paniagua and Aragones 2015; Gómez-Rodríguez *et al.* 2010). Climate change (*e.g.*, enhanced evapotranspiration and reduced precipitation) will lead to a reduction of river flow and a decline of water level in lakes and reservoirs in the Mediterranean region (Cramer et al. 2018). The long-term recording of the temporal changes in the number and distribution of surface water bodies, as well as the extent and duration of surface water presence may provide new insights for understanding their effect on climate. For instance, newly built large artificial reservoirs and dams can significantly alter local and regional natural precipitation (Degu *et al.* 2011; Ekhtiari *et*

al. 2017; Hossain et al. 2009). The lake area together with other lake physical properties can be utilised in general circulation models for climate simulation (Subin et al. 2012). Therefore, the importance of surface water dynamics for various processes makes the new dense dataset an important input for a wide variety of applications. For this reason, the 18year time series of surface water dataset developed in this thesis was made publicly available in an online depository, and can he accessed through https://doi.org/10.17026/dans-xrz-y92s.

6.7 Outlook

Big data and cloud computing

Earth Observation data is now accumulated to the petabyte-scale (Gorelick *et al.* 2017) and will increase exponentially with time. Large-scale, frequent, and long-term monitoring of the land surface requires processing and management of large amounts of geospatial data, which is hard to be handled by desktop computing resources alone. As a consequence, new scientific challenges arise related to data storage, processing capacity, as well as development of more efficient, robust and automatic methods to extract information from these datasets.

Cloud-based platforms have opened a new era of Earth Observation data processing. Examples include Google Earth Engine (GEE: https://earthengine.google.com), Amazon Web Services (AWS: https://aws.amazon.com/earth), the Austrian Geoscience data cube (https://www.opendatacube.org), the European Copernicus Data and Information Access Services (DIAS: https://www.copernicus.eu/en/access-data/dias), and the Austrian Earth Data Observation Centre for Water Resources Management (EODC: https://www.eodc.eu). These emerging platforms can provide the required flexibility to manage, for both storage and computation, huge amounts of geospatial data and to efficiently process them (Bucur et al. 2018; Gorelick et al. 2017). This offers new opportunities in the geospatial sciences for applications at larger spatial and finer temporal scales, such as planetary-scale monitoring of land surface components (Donchyts et al. 2016; Hansen et al. 2013; Liu et al. 2018; Pekel et al. 2016; Poortinga et al. 2018).

Recently, various studies have explored the utility of GEE for hydrological applications. This includes paddy rice mapping (Dong *et al.* 2016), near real-time monitoring of ice breakup (Beaton *et al.* 2019), wetland classification and inventory (Hird *et al.* 2017;

Mahdianpari et al. 2018), water level monitoring (Nguy-Robertson et al. 2018), and planetary-scale surface water monitoring (Donchyts et al. 2016; Ji et al. 2018; Pekel et al. 2016). Pekel et al. (2016) exploited 1.8 petabytes of Landsat archive imagery to map global surface water dynamics in 30 m pixels over 1984–2015, taking advantage of the high-performance computing power of GEE. These studies will significantly improve our knowledge and understanding of water resources availability and various dynamic processes associated with water-related ecosystems. GEE allows researchers to disseminate and share their results and datasets, which creates opportunities and a convenient environment for validation, comparison, integration, and fusion. For example, this thesis successfully leveraged the Global Surface Water dataset (GSW) developed by Pekel et al. (2016), freely available in GEE, to train and validate the regression models for deriving surface water fraction. Potentially, the regression models developed in this thesis could be implemented in GEE at the global scale and to daily MODIS data. In this way, fast and timely information on surface water extent and its changes can be delivered. Moreover, GEE also houses many meteorological data and environmental variables (e.g., temperature, precipitation, winds, humidity, moisture content), and socio-economic datasets, many of which are updated on a daily basis (Gorelick et al. 2017). This allows end users to link hydrological dynamics to the climatological or anthropological drivers, and to address major causes of water resource loss, wetland degradation and other changes for different regions.

Opportunities with new and future satellite missions

Following the launch of the Terra satellite in 1999 and the Aqua satellite in 2002, MODIS instruments have successfully operated beyond their 6-year design lifetime. However, with the evident degradation especially of the Terra MODIS sensor (Sayer *et al.* 2015; Wang *et al.* 2012), MODIS is expected to cease operation in the coming years. To guarantee seamless time series of surface water information, it is critical to establish which satellite systems can take over MODIS's role effectively.

The Visible Infrared Imaging Radiometer Suite (VIIRS) sensor aboard the Suomi National Polar-orbiting Partnership (Suomi-NPP) satellite and the newly launched Ocean and Land Color Instrument (OLCI) onboard Sentinel-3, both satellite sensors with a coarse spatial resolution but with a short repeat cycle. These sensors can provide continuity of MODIS for global land science applications including surface water, and will add to the coarse-resolution, high-frequency time series record. Future research should evaluate if the approach for large-scale and sub-pixel surface water monitoring

Synthesis

developed in this thesis could be extended to VIIRS and Sentinel-3 data, and can yield comparable results in order to construct a seamless time series. This can ensure continuous monitoring of global surface water at similar spatial resolution and in near real-time.

The capabilities of new and future optical satellite remote-sensing missions are increasing and these missions are expected to provide better spatial, temporal, spectral, and radiometric resolutions. The USGS/NASA Landsat Science Team recently prioritized the requirements for Landsat-10 and beyond, and assigned the highest priority to increasing the temporal revisit frequency (Wulder *et al.* 2016). Already, the recent Sentinel-2 MultiSpectral Instrument (MSI) contributes to the continuity and improvement of the Landsat series of multispectral observations, offering new opportunities to monitor surface water at both fine spatial and fine temporal resolution. Several studies have verified its ability to map surface water (Du *et al.* 2016; Kaplan and Avdan 2017b; Ogilvie *et al.* 2018a; Yang *et al.* 2018b). As time goes by and its time series grows, more applications will emerge exploring the full potential of Sentinel-2 for surface water monitoring including its ability for identification of small ponds, lakes and narrow rivers, and capturing temporary water bodies.

This thesis focussed on optical multispectral remote sensing data sources, which may have limited discrimination capabilities when other land surface components have similar spectral information with water due to their broad spectral bands. Due to the fine and large number of spectral bands, hyperspectral imaging may offer improved discrimination capabilities (Goetz 2009). For example, hyperspectral imagers have been used for accurate surface water detection in urban areas where building shadows can easily be mistaken as water (Xie et al. 2014; Yang et al. 2018a), and for classifying aquatic vegetation characteristics to delineate wetland areas (Schmidt and Skidmore 2003) at small spatial scales. Newly launched hyperspectral spaceborne missions, such as the PRecursore IperSpettrale della Missione Applicativa with 30 m spatial resolution and 29 days revisit time (Pignatti et al. 2013), and the upcoming missions with a combination of fine spatial and temporal resolution, such as the Environmental Mapping and Analysis Program (30 m, 4-day revisit) (Guanter et al. 2015), the Spaceborne Hyperspectral Applicative Land and Ocean Mission (10 m, 4-day revisit) (Feingersh and Dor 2015) and the Hyperspectral X Imagery (8 m, 3-day revisit), could therefore prove beneficial in more accurate classification and monitoring of hydrological dynamics.

Chapter 6

Optical data is subject to cloud coverage reducing the effective observations of land cover, and has limitations in detecting water bodies beneath vegetation. Although gapfilling techniques such as data fusion (Chen et al. 2018; Ghamisi et al. 2019; Heimhuber et al. 2018; Huang et al. 2016a; Zhu et al. 2018) and data interpolation (Chen et al. 2011; Yan and Roy 2018) can partially address the issue of spatial and temporal gaps, Synthetic Aperture Radar (SAR) also has potential to overcome these issues due to its ability to penetrate clouds and vegetation. While SAR data analysis has been difficult in the past due to image costs and inconsistent observation characteristics, the Sentinel-1 satellites, launched in 2014 and 2016, provide for the first time free, consistent, global observations every 6 days at 5×20 m ground resolution, offering a range of benefits for water body detection including those under vegetation (Amitrano et al. 2014; Bioresita et al. 2018; Clement et al. 2017; Hardy et al. 2019; Xing et al. 2018). The NASA/CNES Surface Water and Ocean Topography (SWOT) mission, to be launched in 2021, is specifically designed to provide fine spatial resolution ($\simeq 10$ m) and good temporal sampling (22 days repeat) of the extent and water level of continental surface water bodies (Biancamaria et al. 2016; Prigent et al. 2016). Microwave data from the Sentinel-1 and SWOT missions would be a good supplement to the surface water monitoring with optical satellites, as performed in this thesis, because of their fine spatial resolution and their insensitivity to cloud cover. In addition, integrating satellite imagery from multiple passive optical, active radar, and LiDAR sensors shows great potential for surface water monitoring, as it can increase observations during cloudy conditions to obtain dense time series data. This will help to improve the accuracy of surface water dynamics especially at critical stages such as flood peaks. Example applications include fast flood water detection and mapping (Donchyts et al. 2017; Irwin et al. 2017), frequent surface water monitoring for cloud prone regions (Markert et al. 2018), and wetland monitoring (Bourgeau-Chavez et al. 2009; Hird et al. 2017; Mahdianpari et al. 2018; Montgomery et al. 2019).

While many opportunities exists for the future research on Earth hydrology with the emergent technologies and missions, this thesis has demonstrated the possibility to accurately and efficiently map the location and extent of surface water, and monitor its temporal changes over time using freely-available remote sensing data. The developed dataset containing two decades of surface water information for the Mediterranean is therefore considered to be a relevant contribution towards an effective analysis of dynamic water resources.

Synthesis

Bibliography

- Adams, J.B., Sabol, D.E., Kapos, V., Almeida, R., Roberts, D.A., Smith, M.O., & Gillespie, A.R. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, 52, 137-154
- Aghabozorgi, S., Seyed Shirkhorshidi, A., & Ying Wah, T. (2015). Time-series clustering – A decade review. *Information Systems*, 53, 16-38
- Aires, F., Miolane, L., Prigent, C., Pham, B., Fluet-Chouinard, E., Lehner, B., & Papa, F. (2017). A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. *Journal of Hydrometeorology*, 18, 1305-1325
- Alsdorf, D., Lettenmaier, D., & Vörösmarty, C. (2003). The need for global, satellitebased observations of terrestrial surface waters. *Eos, Transactions American Geophysical Union, 84*, 269-276
- Alsdorf, D.E., Rodriguez, E., & Lettenmaier, D.P. (2007). Measuring surface water from space. *Reviews of Geophysics*, 45, 1-24
- Amat, J.A., Hortas, F., Arroyo, G.M., Rendon, M.A., Ramirez, J.M., Rendon-Martos, M., Perez-Hurtado, A., & Garrido, A. (2007). Interannual variations in feeding frequencies and food quality of greater flamingo chicks (Phoenicopterus roseus): Evidence from plasma chemistry and effects on body condition. *Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology, 147*, 569-576
- Amitrano, D., Martino, G., Iodice, A., Mitidieri, F., Papa, M., Riccio, D., & Ruello, G. (2014). Sentinel-1 for monitoring reservoirs: A performance analysis. *Remote Sensing*, 6, 10676-10693
- Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., & Weber, J.-L. (2008). GlobCover: The most detailed portrait of earth, ESA Bulletin-European Space Agency
- Arino, O., Gross, D., Ranera, F., Leroy, M., Bicheron, P., Brockman, C., Defourny, P., Vancutsem, C., Achard, F., Durieux, L., Bourg, L., Latham, J., Gregorio, A.D., Witt, R., Herold, M., Sambale, J., Plummer, S., & Weber, J. (2007). GlobCover: ESA service for global land cover from MERIS, in: *International Geoscience and Remote Sensing Symposium*, Barcelona, Spain, 23-27 July 2007, 2412-2415
- Bai, J., Chen, X., Li, J., Yang, L., & Fang, H. (2011). Changes in the area of inland lakes in arid regions of central Asia during the past 30 years. *Environmental Monitoring* andAssessment, 178, 247-256

- Baird, A.J., & Wilby, R.L. (1999). *Eco-hydrology: plants and water in terrestrial and aquatic environments*. London, UK: Routledge
- Baldwin, R.F., Calhoun, A.J.K., & deMaynadier, P.G. (2006). The significance of hydroperiod and stand maturity for pool-breeding amphibians in forested landscapes. *Canadian Journal of Zoology*, 84, 1604-1615
- Bartholomé, E. (2007). Monitoring the environment in Africa: The VGT4Africa and the AMESD projects, in: *International workshop on crop and rangeland monitoring in Eastern Africa*, Nairobi, Kenya, 27-29 March 2007, 303-317
- Bartholomé, E., & Belward, A.S. (2005). GLC2000: A new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, 26, 1959-1977
- Barton, I.J., & Bathols, J.M. (1989). Monitoring floods with AVHRR. Remote Sensing of Environment, 30, 89-94
- Bastviken, D., Cole, J., Pace, M., & Tranvik, L. (2004). Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate. *Global Biogeochemical Cycles, 18*, GB4009
- Batalla, R.J., Gomez, C.M., & Kondolf, G.M. (2004). Reservoir-induced hydrological changes in the Ebro River basin (NE Spain). *Journal of Hydrology, 290*, 117-136
- Batanero, G.L., León-Palmero, E., Li, L., Green, A.J., Rendón-Martos, M., Suttle, C.A.,
 & Reche, I. (2017). Flamingos and drought as drivers of nutrients and microbial dynamics in a saline lake. *Scientific Reports*, 7, 12173
- Beaton, A., Whaley, R., Corston, K., & Kenny, F. (2019). Identifying historic river ice breakup timing using MODIS and Google Earth Engine in support of operational flood monitoring in Northern Ontario. *Remote Sensing of Environment, 224*, 352-364
- Beck, P.S.A., Atzberger, C., Høgda, K.A., Johansen, B., & Skidmore, A.K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, 100, 321-334
- Becker, M.W., & Daw, A. (2005). Influence of lake morphology and clarity on water surface temperature as measured by EOS ASTER. *Remote Sensing of Environment*, 99, 288-294
- Belward, A. (2016). The global observing system for climate: Implementation needs, GCOS-200
- Ben Abdallah, M., Karray, F., Kallel, N., Armougom, F., Mhiri, N., Quemeneur, M., Cayol, J.L., Erauso, G., & Sayadi, S. (2018). Abundance and diversity of prokaryotes

in ephemeral hypersaline lake Chott El Jerid using Illumina Miseq sequencing, DGGE and qPCR assays. *Extremophiles*, *22*, 811-823

- Bertels, L., Smets, B., & Wolfs, D. (2016). Dynamic water surface detection algorithm applied on PROBA-V multispectral data. *Remote Sensing*, *8*, 1010
- Beven, K.J., & Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin, 24*, 43-69
- Biancamaria, S., Lettenmaier, D.P., & Pavelsky, T.M. (2016). The SWOT mission and its capabilities for land hydrology. *Surveys in Geophysics*, 37, 307-337
- Biggs, J., von Fumetti, S., & Kelly-Quinn, M. (2017). The importance of small waterbodies for biodiversity and ecosystem services: implications for policy makers. *Hydrobiologia*, 793, 3-39
- Bioresita, F., Puissant, A., Stumpf, A., & Malet, J.-P. (2018). A method for automatic and rapid mapping of water surfaces from Sentinel-1 imagery. *Remote Sensing*, 10, 217
- Bjerklie, D.M., Birkett, C.M., Jones, J.W., Carabajal, C., Rover, J.A., Fulton, J.W., & Garambois, P.A. (2018). Satellite remote sensing estimation of river discharge: Application to the Yukon River Alaska. *Journal of Hydrology*, 561, 1000-1018
- Bolpagni, R., Poikane, S., Laini, A., Bagella, S., Bartoli, M., & Cantonati, M. (2019). Ecological and conservation value of small standing-water ecosystems: A systematic review of current knowledge and future challenges. *Water*, 11, 402
- Borro, M., Morandeira, N., Salvia, M., Minotti, P., Perna, P., & Kandus, P. (2014). Mapping shallow lakes in a large South American floodplain: A frequency approach on multitemporal Landsat TM/ETM data. *Journal of Hydrology*, *512*, 39-52
- Boschetti, M., Nutini, F., Manfron, G., Brivio, P.A., & Nelson, A. (2014). Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems. *PLOS ONE*, *9*, e88741
- Bourgeau-Chavez, L., Riordan, K., Powell, R., Miller, N., & Nowels, M. (2009). Improving wetland characterization with multi-sensor, multi-temporal SAR and optical/infrared data fusion. In G. Jedlovec (Ed.), *Advances in Geoscience and Remote Sensing* 679-708. India: InTech Publishers
- Bourgeau-Chavez, L.L., Smith, K.B., Brunzell, S.M., Kasischke, E.S., Romanowicz, E.A., & Richardson, C.J. (2005). Remote monitoring of regional inundation patterns and hydroperiod in the greater everglades using synthetic aperture radar. *Wetlands*, 25, 176-191

- Bradley, B.A., & Fleishman, E. (2008). Can remote sensing of land cover improve species distribution modelling? *Journal of Biogeography*, 35, 1158-1159
- Breiman, L. (2001). Random Forests. Machine Learning, 45, 5-32
- Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A. (1984). *Classification and Regression Trees*. Taylor & Francis
- Brisco, B. (2015). Mapping and monitoring surface water and wetlands with synthetic aperture radar. In R.W. Tiner *et al.* (Eds.), *Remote Sensing of Wetlands: Applications and Advances* 119-136: CRC Press
- Bryant, R.G. (1999). Application of AVHRR to monitoring a climatically sensitive playa. case study: Chott El Djerid, southern Tunisia. *Earth Surface Processes and Landforms*, 24, 283-302
- Bucur, A., Wagner, W., Elefante, S., Naeimi, V., & Briese, C. (2018). Development of an earth observation cloud platform in support to water resources monitoring. In P.-P. Mathieu, & C. Aubrecht (Eds.), *Earth Observation Open Science and Innovation* 275-283. Cham: Springer
- Buma, W., Lee, S.-I., & Seo, J. (2016). Hydrological evaluation of Lake Chad basin using space borne and hydrological model observations. *Water*, 8, 205
- Busker, T., de Roo, A., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., Pekel, J.F.,
 & Cottam, A. (2019). A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry. *Hydrology and Earth System Sciences, 23*, 669-690
- Cael, B.B., Heathcote, A.J., & Seekell, D.A. (2017). The volume and mean depth of Earth's lakes. *Geophysical Research Letters*, 44, 209-218
- Calhoun, A.J.K., Mushet, D.M., Bell, K.P., Boix, D., Fitzsimons, J.A., & Isselin-Nondedeu, F. (2017). Temporary wetlands: challenges and solutions to conserving a 'disappearing' ecosystem. *Biological Conservation*, 211, 3-11
- Campbell, J.B. (2002). Introduction to remote sensing. (3 ed.). London: CRC Press
- Campos-Taberner, M., García-Haro, F., Busetto, L., Ranghetti, L., Martínez, B., Gilabert, M.A., Camps-Valls, G., Camacho, F., & Boschetti, M. (2018). A critical comparison of remote sensing Leaf Area Index estimates over rice-cultivated areas: From Sentinel-2 and Landsat-7/8 to MODIS, GEOV1 and EUMETSAT polar system. *Remote Sensing*, 10, 763
- Campos, J.C., Sillero, N., & Brito, J.C. (2012). Normalized difference water indexes have dissimilar performances in detecting seasonal and permanent water in the Sahara– Sahel transition zone. *Journal of Hydrology*, 464, 438-446

- Carroll, M.L., DiMiceli, C.M., Wooten, M.R., Hubbard, A.B., Sohlberg, R.A., & Townshend, J.R.G. (2017). MOD44W MODIS/Terra Land Water Mask Derived from MODIS and SRTM L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC
- Carroll, M.L., Townshend, J.R., DiMiceli, C.M., Noojipady, P., & Sohlberg, R.A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2, 291-308
- Castañeda, C., & Herrero, J. (2008). Assessing the degradation of saline wetlands in an arid agricultural region in Spain. *CATENA*, 72, 205-213
- CBD (2010). Aichi Biodiversity Targets. Convention on Biological Diversity. Retrieved from: http://www.cbd.int/sp/targets/
- Cezilly, F., Boy, V., & Green, R.E. (1995). Interannual variation in greater flamingo breeding success in relation to water levels. *Ecology*, 76, 20-26
- Chahine, M.T. (1992). The hydrological cycle and its influence on climate. *Nature*, *359*, 373-380
- Chen, B., Chen, L., Huang, B., Michishita, R., & Xu, B. (2018). Dynamic monitoring of the Poyang Lake wetland by integrating Landsat and MODIS observations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 139, 75-87
- Chen, D., Huang, J., & Jackson, T.J. (2005). Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing of Environment, 98*, 225-236
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., & Mills, J. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry* and Remote Sensing, 103, 7-27
- Chen, J., Zhu, X., Vogelmann, J., Gao, F., & Jin, S. (2011). A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sensing of Environment*, 115, 1053-1064
- Chen, L., Michishita, R., & Xu, B. (2014). Abrupt spatiotemporal land and water changes and their potential drivers in Poyang Lake, 2000–2012. *ISPRS Journal of Photogrammetry and Remote Sensing*, 98, 85-93
- Chen, Y., Huang, C., Ticehurst, C., Merrin, L., & Thew, P. (2013). An evaluation of MODIS daily and 8-day composite products for floodplain and wetland inundation mapping. *Wetlands*, 33, 823-835

- Chipman, J.W., & Lillesand, T.M. (2007). Satellite-based assessment of the dynamics of new lakes in southern Egypt. *International Journal of Remote Sensing*, 28, 4365-4379
- Cian, F., Marconcini, M., & Ceccato, P. (2018). Normalized Difference Flood Index for rapid flood mapping: Taking advantage of EO big data. *Remote Sensing of Environment, 209*, 712-730
- Clement, M., Kilsby, C., & Moore, P. (2017). *Multi-Temporal SAR Flood Mapping using Change Detection. Journal of Flood Risk Management, 11, 152-168*
- Cleveland, R.B., Cleveland, W.S., McRae, J.E., & Terpenning, I. (1990). STL: a seasonal-trend decomposition. *Journal of Official Statistics*, *6*, 3-73
- Cohen, W.B. (1991). Response of vegetation indexes to changes in 3 measures of leaf water-stress. *Photogrammetric Engineering and Remote Sensing*, *57*, 195-202
- Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., & Melack, J. (2007). Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems*, 10, 172-185
- Collen, B., Whitton, F., Dyer, E.E., Baillie, J.E.M., Cumberlidge, N., Darwall, W.R.T., Pollock, C., Richman, N.I., Soulsby, A.-M., & Böhm, M. (2014). Global patterns of freshwater species diversity, threat and endemism. *Global Ecology and Biogeography*, 23, 40-51
- Conde-Álvarez, R.M., Bañares-España, E., Nieto-Caldera, J.M., Flores-Moya, A., & Figueroa, F.L. (2012). Submerged macrophyte biomass distribution in the shallow saline lake Fuente de Piedra (Spain) as function of environmental variables. *Anales del Jardín Botánico de Madrid*, 69, 119-127
- Costa, L.T., Farinha, J.C., Hecker, N., & P, T.V. (1996). *Mediterranean Wetland Inventory: A Reference Manual*. Lisboa, Portugal: MedWet Publication
- Cowardin, L.M., Carter, V., Golet, F.C., & LaRoe, E.T. (1979). Classification of wetlands and deepwater habitats of the United States, US Department of the Interior, US Fish and Wildlife Service
- Cramer, W., Guiot, J., Fader, M., Garrabou, J., Gattuso, J.P., Iglesias, A., Lange, M.A., Lionello, P., Llasat, M.C., Paz, S., Penuelas, J., Snoussi, M., Toreti, A., Tsimplis, M.N., & Xoplaki, E. (2018). Climate change and interconnected risks to sustainable development in the Mediterranean. *Nature Climate Change*, *8*, 972-980
- Crétaux, J.F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.C., Nino, F., Abarca Del Rio, R., Cazenave, A., & Maisongrande, P.

(2011). SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Advances in Space Research*, *47*, 1497-1507

- Crist, E.P. (1985). A TM Tasseled Cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment, 17*, 301-306
- Crist, E.P., & Cicone, R.C. (1984). A physically-based transformation of thematic mapper data - the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, 22, 256-263
- Crist, E.P., Laurin, R., & Cicone, R.C. (1986). Vegetation and soils information contained in transformed Thematic Mapper data, in: *International Geoscience and Remote Sensing Symposium*, Zurich, Switzerland, 1465-1470
- Davranche, A., Lefebvre, G., & Poulin, B. (2010). Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sensing of Environment*, 114, 552-562
- Davranche, A., Poulin, B., & Lefebvre, G. (2013). Mapping flooding regimes in Camargue wetlands using seasonal multispectral data. *Remote Sensing of Environment*, 138, 165-171
- De Castro, F., & Reinoso, J.C.M. (1997). Model of long-term water-table dynamics at Donana National Park. *Water Research*, *31*, 2586-2596
- De Wit, R., Guerrero, M.C., Legaz, A., Jonkers, H.M., Blocier, L., Gumiaux, C., & Gautret, P. (2013). Conservation of a permanent hypersaline lake: management options evaluated from decadal variability of Coleofasciculus chthonoplastes microbial mats. *Aquatic Conservation-Marine and Freshwater Ecosystems*, 23, 532-545
- Deane, D.C., Fordham, D.A., He, F., & Bradshaw, C.J.A. (2017). Future extinction risk of wetland plants is higher from individual patch loss than total area reduction. *Biological Conservation*, 209, 27-33
- Degu, A.M., Hossain, F., Niyogi, D., Pielke Sr., R., Shepherd, J.M., Voisin, N., & Chronis, T. (2011). The influence of large dams on surrounding climate and precipitation patterns. *Geophysical Research Letters*, 38, L04405
- Díaz-Delgado, R., Aragonés, D., Afán, I., & Bustamante, J. (2016). Long-term monitoring of the flooding regime and hydroperiod of Doñana marshes with Landsat time series (1974–2014). *Remote Sensing*, 8, 775
- Diaz-Paniagua, C., & Aragones, D. (2015). Permanent and temporary ponds in Donana National Park (SW Spain) are threatened by desiccation. *Limnetica*, 34, 407-423

- Dickens, C., Rebelo, L.-M., & Nhamo, L. (2017). Guidelines and indicators for Target6.6 of the SDGs: Change in the extent of water-related ecosystems over time, International Water Management Institute
- Dietz, A.J., Klein, I., Gessner, U., Frey, C.M., Kuenzer, C., & Dech, S. (2017). Detection of water bodies from AVHRR data-A TIMELINE thematic processor. *Remote Sesning*, 9, 57
- Doña, C., Chang, N.-B., Caselles, V., Sánchez, J., Pérez-Planells, L., Bisquert, M., García-Santos, V., Imen, S., & Camacho, A. (2016). Monitoring hydrological patterns of temporary lakes using remote sensing and machine learning models: Case study of la Mancha Húmeda Biosphere Reserve in central Spain. *Remote Sensing*, 8, 618
- Doña, C., Chang, N.-B., Caselles, V., Sánchez, J.M., Camacho, A., Delegido, J., & Vannah, B.W. (2015). Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain. *Journal of Environmental Management*, 151, 416-426
- Donchyts, G., Baart, F., Winsemius, H., Gorelick, N., Kwadijk, J., & van de Giesen, N. (2016). Earth's surface water change over the past 30 years. *Nature Climate Change*, 6, 810-813
- Donchyts, G., Giesen, N.v.d., & Gorelick, N. (2017). Reconstruction of reservoir and lake surface area dynamics from optical and SAR satellite imagery, in: *International Workshop on the Analysis of Multitemporal Remote Sensing Images* (MultiTemp), Bruges, Belgium, 27-29 June 2017
- Dong, J., Xiao, X., Kou, W., Qin, Y., Zhang, G., Li, L., Jin, C., Zhou, Y., Wang, J., Biradar, C., Liu, J., & Moore, B. (2015). Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms. *Remote Sensing of Environment*, 160, 99-113
- Dong, J., Xiao, X., Menarguez, M.A., Zhang, G., Qin, Y., Thau, D., Biradar, C., & Moore III, B. (2016). Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sensing of Environment, 185*, 142-154
- Downing, J. (2008). Emerging global role of small lakes and ponds: little things mean a lot. *Limnetica*, *29*, 9-24
- Downing, J.A., Prairie, Y.T., Cole, J.J., Duarte, C.M., Tranvik, L.J., Striegl, R.G., McDowell, W.H., Kortelainen, P., Caraco, N.F., Melack, J.M., & Middelburg, J.J. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*, 51, 2388-2397
- Drake, J.C., Jenness, J.S., Calvert, J., & Griffis-Kyle, K.L. (2015). Testing a model for the prediction of isolated waters in the Sonoran Desert. *Journal of Arid Environments, 118*, 1-8
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25-36
- Drzewiecki, W. (2016). Comparison of selected machine learning algorithms for sub-Pixel imperviousness change assessment, in: *Baltic Geodetic Congress Geomatics*, Gdansk, Poland, 2-4 June 2016, 67-72
- Du, Y., Zhang, Y.H., Ling, F., Wang, Q.M., Li, W.B., & Li, X.D. (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, 354
- Duan, Z., & Bastiaanssen, W.G.M. (2013). Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sensing of Environment, 134*, 403-416
- Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Leveque,
 C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L.J., & Sullivan, C.A.
 (2006). Freshwater biodiversity: importance, threats, status and conservation
 challenges. *Biological Reviews*, *81*, 163-182
- Eitel, J.U.H., Gessler, P.E., Smith, A.M.S., & Robberecht, R. (2006). Suitability of existing and novel spectral indices to remotely detect water stress in Populus spp. *Forest Ecology and Management*, 229, 170-182
- Ekhtiari, N., Grossman-Clarke, S., Koch, H., Meira de Souza, W., Donner, R.V., & Volkholz, J. (2017). Effects of the lake Sobradinho reservoir (Northeastern Brazil) on the regional climate. *Climate*, *5*, 50
- European Communities (1992). Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. *Official Journal of the European Union, 206,* 7-50
- Feingersh, T., & Dor, E.B. (2015). SHALOM A commercial hyperspectral space mission. In S. Qian (Ed.), Optical Payloads for Space Missions 247-263. Wiley
- Feng, D., Zhao, Y., Yu, L., Li, C., Wang, J., Clinton, N., Bai, Y., Belward, A., Zhu, Z.,& Gong, P. (2016). Circa 2014 African land-cover maps compatible with FROM-

GLC and GLC2000 classification schemes based on multi-seasonal Landsat data. *International Journal of Remote Sensing*, *37*, 4648-4664

- Feng, L., Hu, C.M., Chen, X.L., Cai, X.B., Tian, L.Q., & Gan, W.X. (2012). Assessment of inundation changes of Poyang Lake using MODIS observations between 2000 and 2010. *Remote Sensing of Environment*, 121, 80-92
- Feng, M., Sexton, J.O., Channan, S., & Townshend, J.R. (2015). A global, high-resolution (30-m) inland water body dataset for 2000: first results of a topographic–spectral classification algorithm. *International Journal of Digital Earth*, 9, 113-133
- Feyisa, G.L., Meilby, H., Fensholt, R., & Proud, S.R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140, 23-35
- Finlayson, C., Davidson, N., Spiers, A., & Stevenson, N. (1999). Global wetland inventory-current status and future priorities. *Marine and Freshwater Research*, 50, 717-727
- Fisher, A., & Danaher, T. (2013). A water index for SPOT5 HRG satellite imagery, New South Wales, Australia, determined by linear discriminant analysis. *Remote Sensing*, 5, 5907-5925
- Fisher, A., Flood, N., & Danaher, T. (2016). Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sensing of Environment*, 175, 167-182
- Fluet-Chouinard, E., Lehner, B., Rebelo, L.-M., Papa, F., & Hamilton, S.K. (2015). Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sensing of Environment*, 158, 348-361
- Foley, J.A., Costa, M.H., Delire, C., Ramankutty, N., & Snyder, P. (2003). Green surprise? How terrestrial ecosystems could affect earth's climate. *Frontiers in Ecology and the Environment*, 1, 38-44
- Frazier, P.S., & Page, K.J. (2000). Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering and Remote Sensing*, *66*, 1461-1467
- Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, *83*, 287-302

- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182
- Frohn, R.C., D'Amico, E., Lane, C., Autrey, B., Rhodus, J., & Liu, H. (2012). Multitemporal sub-pixel Landsat ETM+ classification of isolated wetlands in Cuyahoga County, Ohio, USA. *Wetlands*, 32, 289-299
- Galewski, T. (2012). Biodiversity: Status and trends of species in Mediterranean wetlands, Mediterranean Wetlands Observatory
- Gan, M., Deng, J., Zheng, X., Hong, Y., & Wang, K. (2014). Monitoring urban greenness dynamics using multiple endmember spectral mixture analysis. *PLOS ONE*, 9, e112202
- Gao, B.C. (1996). NDWI-A normalized difference water index for remote sensing of vegetation liquid water from spac. *Remote Sensing of Environment*, 58, 257-266
- Gao, H., Birkett, C., & Lettenmaier, D.P. (2012). Global monitoring of large reservoir storage from satellite remote sensing. *Water Resources Research, 48, W0954*
- García, C.M., García-Ruiz, R., Rendón, M., Niell, F.X., & Lucena, J. (1997). Hydrological cycle and interannual variability of the aquatic community in a temporary saline lake (Fuente de Piedra, southern Spain). *Hydrobiologia*, 345, 131-141
- García, C.M., & Niell, F.X. (1993). Seasonal change in a saline temporary lake (Fuente de Piedra, southern Spain). *Hydrobiologia*, 267, 211-223
- Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., & Fernandez, V. (2017). Copernicus Sentinel-2A calibration and products validation status. *Remote Sensing*, *9*, 584
- GCOS (2011). Systematic observation requirements for satellite-based products for climate supplemental details to the satellite-based component of the implementation plan for the global observing system for climate in support of the UNFCCC: 2011 update, World Meteorological Organization (WMO)
- Geraci, J., Bechet, A., Cezilly, F., Ficheux, S., Baccetti, N., Samraoui, B., & Wattier, R. (2012). Greater flamingo colonies around the Mediterranean form a single interbreeding population and share a common history. *Journal of Avian Biology*, 43, 341-354

- Ghamisi, P., Rasti, B., Yokoya, N., Wang, Q.M., Hofle, B., Bruzzone, L., Bovolo, F., Chi, M.M., Anders, K., Gloaguen, R., Atkinson, P.M., & Benediktsson, J.A. (2019).
 Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 7, 6-39
- Gibbs, J.P. (1993). Importance of small wetlands for the persistence of local-populations of wetland-associated animals. *Wetlands*, 13, 25-31
- Goetz, A.F.H. (2009). Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment, 113*, S5-S16
- Gómez-Rodríguez, C., Bustamante, J., & Díaz-Paniagua, C. (2010). Evidence of hydroperiod shortening in a preserved system of temporary ponds. *Remote Sensing*, 2, 1439-1462
- Gond, V., Bartholomé, E., Ouattara, F., Nonguierma, A., & Bado, L. (2004). Surveillance et cartographie des plans d'eau et des zones humides et inondables en régions arides avec l'instrument VEGETATION embarqué sur SPOT-4. *International Journal of Remote Sensing*, 25, 987-1004
- Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., & Chen, J. (2013). Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, *34*, 2607-2654
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27
- Grabs, T., Seibert, J., Bishop, K., & Laudon, H. (2009). Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. *Journal of Hydrology*, 373, 15-23
- Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M.C., Shyamsundar,
 P., Steffen, W., Glaser, G., Kanie, N., & Noble, I. (2013). Sustainable development
 goals for people and planet. *Nature*, 495, 305-307
- Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., Kuester, T., Hollstein, A., Rossner, G., Chlebek, C., Straif, C., Fischer, S., Schrader, S., Storch, T., Heiden, U., Mueller, A., Bachmann, M., Mühle, H., Müller, R., Habermeyer, M.,

Ohndorf, A., Hill, J., Buddenbaum, H., Hostert, P., Van der Linden, S., Leitão, P.J., Rabe, A., Doerffer, R., Krasemann, H., Xi, H., Mauser, W., Hank, T., Locherer, M., Rast, M., Staenz, K., & Sang, B. (2015). The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sensing*, *7*, 8830-8857

- Guerschman, J.P., Scarth, P.F., McVicar, T.R., Renzullo, L.J., Malthus, T.J., Stewart, J.B., Rickards, J.E., & Trevithick, R. (2015). Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. *Remote Sensing of Environment, 161*, 12-26
- Guerschman, J.P., Warren, G., Byrne, G., Lymburner, L., Mueller, N., & Van Dijk, A.I.
 (2011). MODIS-based Standing Water Detection for Flood and Large Reservoir Mapping: Algorithm Development and Applications for the Australian Continent. Canberra, Australia: CSIRO
- Haas, E.M., Bartholomé, E., & Combal, B. (2009). Time series analysis of optical remote sensing data for the mapping of temporary surface water bodies in sub-Saharan western Africa. *Journal of Hydrology*, 370, 52-63
- Halabisky, M., Moskal, L.M., Gillespie, A., & Hannam, M. (2016). Reconstructing semiarid wetland surface water dynamics through spectral mixture analysis of a time series of Landsat satellite images (1984–2011). *Remote Sensing of Environment, 177*, 171-183
- Hails, A.J. (1997). Wetlands, Biodiversity and the Ramsar Convention: The Role of the Convention on Wetlands in the Conservation and Wise Use of Biodiversity. Gland, Switzerland: Ramsar Convention Bureau
- Hamasaki, K., Yamashita, S., Ishiyama, N., & Kitada, S. (2013). Effects of Water Availability and Migration Timing from Sea to Land on Survival and Moulting in Megalopae and Juveniles of the Coconut Crab, Birgus Latro: Implications for Mass Production of Juveniles. *Journal of Crustacean Biology*, 33, 627-632
- Hansen, M.C., Defries, R.S., Townshend, J.R.G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331-1364
- Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment*, 83, 303-319

- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., & Townshend, J.R.G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*, 850-853
- Hardy, A., Ettritch, G., Cross, D.E., Bunting, P., Liywalii, F., Sakala, J., Silumesii, A., Singini, D., Smith, M., Willis, T., & Thomas, C.J. (2019). Automatic detection of open and vegetated water bodies using Sentinel 1 to map African malaria vector mosquito breeding habitats. *Remote Sensing*, 11, 25
- Hayashi, M., & van der Kamp, G. (2000). Simple equations to represent the volume-areadepth relations of shallow wetlands in small topographic depressions. *Journal of Hydrology*, 237, 74-85
- He, Y., Zhang, X., & Hua, L. (2016). Object-based distinction between building shadow and water in high resolution imagery using fuzzy-rule classification and artificial bee colony optimization. *Journal of Sensors, 2016*, 1-10
- Heimhuber, V., Tulbure, M.G., & Broich, M. (2016). Modeling 25 years of spatiotemporal surface water and inundation dynamics on large river basin scale using time series of Earth observation data. *Hydrology and Earth System Sciences, 20*, 2227-2250
- Heimhuber, V., Tulbure, M.G., & Broich, M. (2018). Addressing spatio-temporal resolution constraints in Landsat and MODIS-based mapping of large-scale floodplain inundation dynamics. *Remote Sensing of Environment*, 211, 307-320
- Heredia, J., García de Domingo, A., Ruiz, J.M., & Araguás, L. (2010). Fuente de Piedra Lagoon (Málaga, Spain): a deep Karstic flow discharge point of a regional hydrogeological system. In B. Andreo *et al.* (Eds.), *Advances in Research in Karst Media* 231-236. Berlin Heidelberg: Springer
- Herrero, J., & Castaneda, C. (2009). Delineation and functional status monitoring in small saline wetlands of NE Spain. *Journal of Environmental Management*, 90, 2212-2218
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965-1978
- Hird, J.N., DeLancey, E.R., McDermid, G.J., & Kariyeva, J. (2017). Google Earth Engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. *Remote Sensing*, 9, 1315
- Holgerson, M.A., & Raymond, P.A. (2016). Large contribution to inland water CO2 and CH4 emissions from very small ponds. *Nature Geoscience*, 9, 222

- Hommersom, A., Wernand, M.R., Peters, S., Eleveld, M.A., van der Woerd, H.J., & de Boer, J. (2011). Spectra of a shallow sea-unmixing for class identification and monitoring of coastal waters. *Ocean Dynamics*, 61, 463-480
- Hope, A.S., Coulter, L.L., & Stow, D.A. (1999). Estimating lake area in an Arctic landscape using linear mixture modelling with AVHRR data. *International Journal* of Remote Sensing, 20, 829-835
- Hossain, F., Jeyachandran, I., & Pielke Sr., R. (2009). Have large dams altered extreme precipitation patterns? *Eos, Transactions American Geophysical Union, 90*, 453-454
- Howari, F., Goodell, P., & Miyamoto, S. (2002). Spectral properties of salt crusts formed on saline soils. *Journal of Environmental Quality*, *31*, 1453-1461
- Hu, S., Niu, Z., & Chen, Y. (2017). Global wetland datasets: a review. *Wetlands, 37*, 807-817
- Huang, C., Chen, Y., & Wu, J. (2014a). Mapping spatio-temporal flood inundation dynamics at large river basin scale using time-series flow data and MODIS imagery. *International Journal of Applied Earth Observation and Geoinformation*, 26, 350-362
- Huang, C., Chen, Y., Wu, J.P., Li, L.Y., & Liu, R. (2015). An evaluation of Suomi NPP-VIIRS data for surface water detection. *Remote Sensing Letters*, *6*, 155-164
- Huang, C., Chen, Y., Zhang, S., Li, L., Shi, K., & Liu, R. (2017). Spatial downscaling of Suomi NPP-VIIRS image for lake mapping. *Water*, 9, 834
- Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018a). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Reviews of Geophysics*, 56, 333-360
- Huang, C., Chen, Y., Zhang, S.Q., Li, L.Y., Shi, K.F., & Liu, R. (2016a). Surface water mapping from Suomi NPP-VIIRS imagery at 30 m resolution via blending with Landsat data. *Remote Sensing*, 8, 14
- Huang, C., Peng, Y., Lang, M., Yeo, I.-Y., & McCarty, G. (2014b). Wetland inundation mapping and change monitoring using Landsat and airborne LiDAR data. *Remote Sensing of Environment*, 141, 231-242
- Huang, C., & Townshend, J.R.G. (2003). A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover. *International Journal of Remote Sensing*, 24, 75-90
- Huang, Q., Long, D., Du, M., Zeng, C., Qiao, G., Li, X., Hou, A., & Hong, Y. (2018b). Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River. *Remote Sensing of Environment, 219*, 115-134

- Huang, X., Schneider, A., & Friedl, M.A. (2016b). Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights. *Remote Sensing of Environment*, 175, 92-108
- Irish, R.R. (2000). Landsat 7 science data users handbook, National Aeronautics and Space Administration
- Irwin, K., Beaulne, D., Braun, A., & Fotopoulos, G. (2017). Fusion of SAR, optical imagery and airborne LiDAR for surface water detection. *Remote Sensing*, *9*, 18
- JAES-CC (2010). A Framework for Integrated Wetland Management of the Jabbul Agroecosystem. Aleppo, Syria: International Center for Agricultural Researach in the Dry Areas
- Jain, S.K., Singh, R.D., Jain, M.K., & Lohani, A.K. (2005). Delineation of flood-prone areas using remote sensing techniques. *Water Resources Management*, 19, 333-347
- Jarvis, A., Reuter, H.I., Nelson, A., & Guevara, E. (2008). Hole-filled seamless SRTM data Version 4. International Center for Tropical Agriculture (CIAT). Retrieved from: http://srtm. csi. cgiar. org
- Jensen, J.R. (2009). *Remote sensing of the environment: An earth resource perspective.* (2 ed.) Upper Saddle, US: Pearson Education
- Jenson, S.K., & Domingue, J.O. (1988). Extracting topographic structure from digital elevation data for geographic information-system analysis. *Photogrammetric Engineering and Remote Sensing*, 54, 1593-1600
- Ji, L., Gong, P., Wang, J., Shi, J., & Zhu, Z. (2018). Construction of the 500-m resolution daily global surface water change database (2001–2016). *Water Resources Research*, 54, 270-292
- Ji, L., Zhang, L., & Wylie, B. (2009). Analysis of dynamic thresholds for the Normalized Difference Water Index. *Photogrammetric Engineering and Remote Sensing*, 75, 1307-1317
- Ji, L., Zhang, L., Wylie, B.K., & Rover, J. (2011). On the terminology of the spectral vegetation index (NIR - SWIR)/(NIR + SWIR). *International Journal of Remote Sensing*, 32, 6901-6909
- Jian, J., Yang, W.N., Jiang, H., Wan, X.N., Li, Y.X., & Peng, L. (2012). A model for retrieving soil moisture saturation with Landsat remotely sensed data. *International Journal of Remote Sensing*, 33, 4553-4566
- Jiang, H., Liu, C., Sun, X., Lu, J., Zou, C., Hou, Y., & Lu, X. (2015). Remote sensing reversion of water depths and water management for the stopover site of Siberian Cranes at Momoge, China. *Wetlands*, 35, 369-379

- Jin, H., Huang, C., Lang, M.W., Yeo, I.-Y., & Stehman, S.V. (2017). Monitoring of wetland inundation dynamics in the Delmarva Peninsula using Landsat time-series imagery from 1985 to 2011. *Remote Sensing of Environment*, 190, 26-41
- Jin, S., & Sader, S.A. (2005). Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment*, 94, 364-372
- Jones, J.W. (2015). Efficient wetland surface water detection and monitoring via Landsat: comparison with in situ data from the Everglades Depth Estimation Network. *Remote Sensing*, 7, 12503
- Jung, H.C., Alsdorf, D., Moritz, M., Lee, H., & Vassolo, S. (2011). Analysis of the relationship between flooding area and water height in the Logone floodplain. *Physics* and Chemistry of the Earth, 36, 232-240
- Kabat, P., Claussen, M., Dirmeyer, P., Gash, J.H.C., Bravo de Guenni, L., Meybeck, M., Pielke Sr, R., Vörösmarty, C.J., Hutjes, R., & Lütkemeier, S. (2004). Vegetation, Water, Humans and the Climate: A New Perspective on an Interactive System. Berlin, Germany: Springer
- Kaplan, G., & Avdan, U. (2017a). Mapping and monitoring wetlands using Sentinel-2 satellite imagery, in: *International Workshop on Geoinformation Science: GeoAdvances 2017*, Safranbolu, Turkey, 14-15 October 2017, 271-277
- Kaplan, G., & Avdan, U. (2017b). Object-based water body extraction model using Sentinel-2 satellite imagery. *European Journal of Remote Sensing*, 50, 137-143
- Kaptue, A.T., Hanan, N.P., & Prihodko, L. (2013). Characterization of the spatial and temporal variability of surface water in the Soudan-Sahel region of Africa. *Journal of Geophysical Research-Biogeosciences*, 118, 1472-1483
- Karpatne, A., Khandelwal, A., Chen, X., Mithal, V., Faghmous, J., & Kumar, V. (2016). Global monitoring of inland water dynamics: State-of-the-art, challenges, and opportunities. In J. Lässig *et al.* (Eds.), *Computational Sustainability* 121-147. Cham: Springer
- Kauth, R.J., & Thomas, G.S. (1976). The tasselled cap-a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat, in: *Machine Processing of Remote Sensing Data Symposium*, LARS, Purdue, 41-51
- Kettle, H., Thompson, R., Anderson, N.J., & Livingstone, D.M. (2004). Empirical modeling of summer lake surface temperatures in southwest Greenland. *Limnology* and Oceanography, 49, 271-282

- Khandelwal, A., Karpatne, A., Marlier, M.E., Kim, J., Lettenmaier, D.P., & Kumar, V. (2017). An approach for global monitoring of surface water extent variations in reservoirs using MODIS data. *Remote Sensing of Environment*, 202, 113-128
- Kim, J.-W., Lu, Z., Jones, J.W., Shum, C.K., Lee, H., & Jia, Y. (2014). Monitoring Everglades freshwater marsh water level using L-band synthetic aperture radar backscatter. *Remote Sensing of Environment*, 150, 66-81
- Klein, I., Dietz, A.J., Gessner, U., Galayeva, A., Myrzakhmetov, A., & Kuenzer, C. (2014). Evaluation of seasonal water body extents in Central Asia over the past 27 years derived from medium-resolution remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*, 26, 335-349
- Klein, I., Gessner, U., Dietz, A.J., & Kuenzer, C. (2017). Global WaterPack A 250m resolution dataset revealing the daily dynamics of global inland water bodies. *Remote Sensing of Environment*, 198, 345-362
- Kobayashi, T., Tsend-Ayush, J., & Tateishi, R. (2014). A new tree cover percentage map in Eurasia at 500 m resolution using MODIS data. *Remote Sensing*, 6, 209-232
- Kohfahl, C., Rodriguez, M., Fenk, C., Menz, C., Benavente, J., Hubberten, H., Meyer, H., Paul, L., Knappe, A., López-Geta, J.A., & Pekdeger, A. (2008). Characterising flow regime and interrelation between surface-water and ground-water in the Fuente de Piedra salt lake basin by means of stable isotopes, hydrogeochemical and hydraulic data. *Journal of Hydrology*, 351, 170-187
- Koning, C.O. (2005). Vegetation patterns resulting from spatial and temporal variability in hydrology, soils, and trampling in an isolated basin marsh, New Hampshire, USA. *Wetlands*, 25, 239-251
- Koshal, A.K. (2012). Spectral characteristics of soil salinity areas in parts of South-West Punjab through Remote Sensing and GIS. *International Journal of Remote Sensing* and GIS, 1, 84-89
- Kuhn, M., Weston, S., Keefer, C., & Coulter, N. (2012). Cubist models for regression. *R* package Vignette R Package Version 0.0.8
- Kuhn, M., Weston;, S., Keefer;, C., Coulter;, N., & Quinlan, R. (2016). Rule- and instance-based regression modeling, *R Package Version 0.0.19*
- Kurt, L., Özdeniz, E., & Özbey, B. (2013). Analysis of Floristic Diversity and Plant Dynamism at Meke Maar Wetland (Karapınar/Konya/Turkey), *in: International Conference on Environmental Science and Technology*, Nevsehir, Turkey, 18-21 June 2013, 47-59

- Lacava, T., Ciancia, E., Faruolo, M., Pergola, N., Satriano, V., & Tramutoli, V. (2019). On the potential of RST-FLOOD on Visible Infrared Imaging Radiometer Suite data for flooded areas detection. *Remote Sensing*, 11, 598
- Lafare, A.E.A., Peach, D.W., & Hughes, A.G. (2016). Use of seasonal trend decomposition to understand groundwater behaviour in the Permo-Triassic Sandstone aquifer, Eden Valley, UK. *Hydrogeology Journal*, 24, 141-158
- Lecina, S., Isidoro, D., Playán, E., & Aragüés, R. (2010). Irrigation modernization and water conservation in Spain: The case of Riegos del Alto Aragón. *Agricultural Water Management*, 97, 1663-1675
- Lehner, B., & Doll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology, 296*, 1-22
- Leibowitz, S.G. (2003). Isolated wetlands and their functions: An ecological perspective. *Wetlands, 23*, 517-531
- Li, L. (2019). Development and validation of a dense 18-year time series of surface water fraction estimates from MODIS for the Mediterranean region. In: DANS. Retrieved from: https://doi.org/10.17026/dans-xrz-y92s
- Li, L., Skidmore, A., Vrieling, A., & Wang, T. (2019). A new dense 18-year time series of surface water fraction estimates from MODIS for the Mediterranean region. *Hydrology and Earth System Sciences*, *23*, 3037-3056
- Li, L., Vrieling, A., Skidmore, A., Wang, T., Muñoz, A.-R., & Turak, E. (2015). Evaluation of MODIS spectral indices for monitoring hydrological dynamics of a small, seasonally-flooded wetland in southern spain. *Wetlands*, 35, 851-864
- Li, L., Vrieling, A., Skidmore, A., Wang, T., & Turak, E. (2018a). Monitoring the dynamics of surface water fraction from MODIS time series in a Mediterranean environment. *International Journal of Applied Earth Observation and Geoinformation*, 66, 135-145
- Li, R.R., Kaufman, Y.J., Gao, B.C., & Davis, C.O. (2003). Remote sensing of suspended sediments and shallow coastal waters. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 559-566
- Li, S., Sun, D., & Yu, Y. (2013a). Automatic cloud-shadow removal from flood/standing water maps using MSG/SEVIRI imagery. *International Journal of Remote Sensing*, 34, 5487-5502
- Li, S., Sun, D.L., Goldberg, M.D., Sjoberg, B., Santek, D., Hoffman, J.P., DeWeese, M., Restrepo, P., Lindsey, S., & Holloway, E. (2018b). Automatic near real-time flood

detection using Suomi-NPP/VIIRS data. *Remote Sensing of Environment, 204*, 672-689

- Li, S.M., Sun, D.L., Yu, Y.Y., Csiszar, I., Stefanidis, A., & Goldberg, M.D. (2013b). A new shortwave infrared (SWIR) method for quantitative water fraction derivation and evaluation with EOS/MODIS and Landsat/TM data. *IEEE Transactions on Geoscience and Remote Sensing*, 51, 1852-1862
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2, 18-22
- Lin, L., Di, L.P., Tang, J.M., Yu, E., Zhang, C., Rahman, M.S., Shrestha, R., & Kang, L.J. (2019). Improvement and validation of NASA/MODIS NRT global flood mapping. *Remote Sensing*, 11, 18
- Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., Pei, F., & Wang, S. (2018). Highresolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sensing of Environment, 209*, 227-239
- Liu, X., Zhang, S., & Li, X. (1983). The application of Landsat image in the surveying of water resources of Dongting Lake, in: *General Assembly of the International Union* of Geodesy and Geophysics, Hamburg, Germany, August 1983, 483-489
- Liu, Y. (2012). Why NDWI threshold varies in delineating water body from multitemporal images? in: *International Geoscience and Remote Sensing Symposium*, Munich, Germany, 22-27 July 2012, 4375-4378
- Livingstone, D.M. (1999). Ice break-up on southern Lake Baikal and its relationship to local and regional air temperatures in Siberia and to the North Atlantic Oscillation. *Limnology and Oceanography*, 44, 1486-1497
- Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., & Merchant, J.W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21, 1303-1330
- Lu, D.S., Moran, E., & Batistella, M. (2003). Linear mixture model applied to Amazonian vegetation classification. *Remote Sensing of Environment*, 87, 456-469
- Lutz, A.F., Immerzeel, W.W., Shrestha, A.B., & Bierkens, M.F.P. (2014). Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nature Climate Change*, 4, 587-589
- Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Homayouni, S., & Gill, E. (2018). The first wetland inventory map of newfoundland at a spatial resolution of 10 m using

sentinel-1 and Sentinel-2 data on the Google Earth Engine cloud computing platform. *Remote Sensing*, *11*, 43

- Mahe, G., Mariko, A., & Orange, D. (2013). Relationships between water level at hydrological stations and inundated area in the River Niger Inner Delta, Mali. In G. Young, & G.M. Perillo (Eds.), *Deltas: Landforms, Ecosystems and Human Activities* 110-115. Wallingford: International Association of Hydrological Sciences
- Mahe, G., Orange, D., Mariko, A., & Bricquet, J.P. (2011). Estimation of the flooded area of the Inner Delta of the River Niger in Mali by hydrological balance and satellite data, In Stewart W. Franks*et at al.* (Eds.), *Hydro-Climatology: Variability and Change* 138-143. Melbourne, Australia: International Association of Hydrological Sciences
- Margono, B.A., Bwangoy, J.-R.B., Potapov, P.V., & Hansen, M.C. (2014). Mapping wetlands in Indonesia using Landsat and PALSAR data-sets and derived topographical indices. *Geo-spatial Information Science*, 17, 60-71
- Markert, K.N., Chishtie, F., Anderson, E.R., Saah, D., & Griffin, R.E. (2018). On the merging of optical and SAR satellite imagery for surface water mapping applications. *Results in Physics*, 9, 275-277
- Marti-Cardona, B., Lopez-Martinez, C., Dolz-Ripolles, J., & Bladè-Castellet, E. (2010). ASAR polarimetric, multi-incidence angle and multitemporal characterization of Doñana wetlands for flood extent monitoring. *Remote Sensing of Environment, 114*, 2802-2815
- Mathieu, R., Seddon, P., & Leiendecker, J. (2006). Predicting the distribution of raptors using remote sensing techniques and Geographic Information Systems: a case study with the Eastern New Zealand falcon (Falco novaeseelandiae). New Zealand Journal of Zoology, 33, 73-84
- Matthews, G.V.T. (1993). *The Ramsar Convention on Wetlands: Its History and Development*. Gland, Switzerland: Ramsar Convention Bureau
- McCarthy, J.M., Gumbricht, T., McCarthy, T., Frost, P., Wessels, K., & Seidel, F. (2003). Flooding patterns of the Okavango wetland in Botswana between 1972 and 2000. *Ambio: A journal of the human environment, 32*, 453-457
- McCulloch, G., Aebischer, A., & Irvine, K. (2003). Satellite tracking of flamingos in southern Africa: the importance of small wetlands for management and conservation. *Oryx*, 37, 480-483
- 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, 1425-1432

- Medina, C., Gomez-Enri, J., Alonso, J.J., & Villares, P. (2010). Water volume variations in Lake Izabal (Guatemala) from in situ measurements and ENVISAT Radar Altimeter (RA-2) and Advanced Synthetic Aperture Radar (ASAR) data products. *Journal of Hydrology*, 382, 34-48
- Messager, M.L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*, 13603
- Ministry of Agriculture Fisheries and Environment (2013). Por el que se declara la Zona especial de conservación Laguna de Fuente de Piedra (eS0000033) y se aprueba el Plan de ordenación de los recursos Naturales de la reserva Natural Laguna de Fuente de Piedra, Boletín Oficial de la Junta de Andalucía
- Mohammadi, A., Costelloe, J.F., & Ryu, D. (2017). Application of time series of remotely sensed normalized difference water, vegetation and moisture indices in characterizing flood dynamics of large-scale arid zone floodplains. *Remote Sensing of Environment*, 190, 70-82
- Mondejar, J.P., & Tongco, A.F. (2019). Near infrared band of Landsat 8 as water index: a case study around Cordova and Lapu-Lapu City, Cebu, Philippines. Sustainable Environment Research, 29, 16
- Montgomery, J., Brisco, B., Chasmer, L., Devito, K., Cobbaert, D., & Hopkinson, C. (2019). SAR and LiDAR temporal data fusion approaches to boreal wetland ecosystem monitoring. *Remote Sensing*, *11*, 24
- Montgomery, J.S., Hopkinson, C., Brisco, B., Patterson, S., & Rood, S.B. (2018). Wetland hydroperiod classification in the western prairies using multitemporal synthetic aperture radar. *Hydrological Processes*, 32, 1476-1490
- Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., & Ip, A. (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341-352
- Muster, S., Heim, B., Abnizova, A., & Boike, J. (2013). Water body distributions across scales: a remote sensing based comparison of three arctic tundra wetlands. *Remote Sensing*, 5, 1498-1523
- Najibi, N., & Devineni, N. (2018). Recent trends in the frequency and duration of global floods. *Earth System Dynamics*, 9, 757-783
- Nakaegawa, T. (2012). Comparison of water-related land cover types in six 1-km global land cover datasets. *Journal of Hydrometeorology*, *13*, 649-664

- Nath, R.K., & K Deb, S. (2010). Water-Body Area Extraction from High Resolution Satellite Images-An Introduction, Review, and Comparison. *International Journal of Image Processing*, 3, 353-372
- Nguy-Robertson, A., May, J., Dartevelle, S., Birkett, C., Lucero, E., Russo, T., Griffin, S., Miller, J., Tetrault, R., & Zentner, M. (2018). Inferring elevation variation of lakes and reservoirs from areal extents: Calibrating with altimeter and in situ data. *Remote Sensing Applications: Society and Environment*, 9, 116-125
- O'Grady, D., & Leblanc, M. (2014). Radar mapping of broad-scale inundation: challenges and opportunities in Australia. *Stochastic Environmental Research and Risk Assessment, 28, 29-38*
- O'Grady, D., Leblanc, M., & Bass, A. (2014). The use of radar satellite data from multiple incidence angles improves surface water mapping. *Remote Sensing of Environment*, 140, 652-664
- Ogilvie, A., Belaud, G., Massuel, S., Mulligan, M., Le Goulven, P., & Calvez, R. (2018a). Surface water monitoring in small water bodies: potential and limits of multi-sensor Landsat time series. *Hydrology and Earth System Sciences*, 22, 4349-4380
- Ogilvie, A., Belaud, G., Massuel, S., Mulligan, M., Le Goulven, P., Malaterre, P.O., & Calvez, R. (2018b). Combining Landsat observations with hydrological modelling for improved surface water monitoring of small lakes. *Journal of Hydrology*, 566, 109-121
- Olsen, J.L., Ceccato, P., Proud, S.R., Fensholt, R., Grippa, M., Mougin, E., Ardo, J., & Sandholt, I. (2013). Relation between seasonally detrended shortwave infrared reflectance data and land surface moisture in semi-arid Sahel. *Remote Sensing*, 5, 2898-2927
- Olthof, I., Fraser, R.H., & Schmitt, C. (2015). Landsat-based mapping of thermokarst lake dynamics on the Tuktoyaktuk Coastal Plain, Northwest Territories, Canada since 1985. *Remote Sensing of Environment*, 168, 194-204
- Ordoyne, C., & Friedl, M.A. (2008). Using MODIS data to characterize seasonal inundation patterns in the Florida Everglades. *Remote Sensing of Environment*, 112, 4107-4119
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics*, 9, 62-66
- Ovakoglou, G., Alexandridis, T.K., Crisman, T.L., Skoulikaris, C., & Vergos, G.S. (2016). Use of MODIS satellite images for detailed lake morphometry: Application to basins

with large water level fluctuations. *International Journal of Applied Earth Observation and Geoinformation*, 51, 37-46

- Ozesmi, S.L., & Bauer, M.E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10, 381-402
- Painter, T.H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R.E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, 113, 868-879
- Pan, F.F., Wang, C., & Xi, X.H. (2016). Constructing river stage-discharge rating curves using remotely sensed river cross-sectional inundation areas and river bathymetry. *Journal of Hydrology*, 540, 670-687
- Papa, F., Prigent, C., Aires, F., Jimenez, C., Rossow, W.B., & Matthews, E. (2010). Interannual variability of surface water extent at the global scale, 1993–2004. *Journal* of Geophysical Research: Atmospheres, 115, D12111
- Parrens, M., Al Bitar, A., Frappart, F., Papa, F., Calmant, S., Crétaux, J.-F., Wigneron, J.-P., & Kerr, Y. (2017). Mapping dynamic water fraction under the tropical rain forests of the amazonian basin from SMOS brightness temperatures. *Water*, 9, 350
- Peel, M.C., Finlayson, B.L., & McMahon, T.A. (2007). Updated world map of the Koppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11, 1633-1644
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A.S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*, 418-422
- Pekel, J.F., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., Bartholomé, E., & Defourny, P. (2014). A near real-time water surface detection method based on HSV transformation of MODIS multi-spectral time series data. *Remote Sensing of Environment, 140*, 704-716
- Pesaresi, M., Ehrlich, D., Florczyk, A.J., Freire, S., Julea, A., Kemper, T., & Syrris, V. (2016). The global human settlement layer from Landsat imagery, in: *International Geoscience and Remote Sensing Symposium*, Beijing, China, 10-15 July 2016, 7276-7279
- Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J., & Stenseth, N.C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20, 503-510
- Petus, C., Lewis, M., & White, D. (2013). Monitoring temporal dynamics of Great Artesian Basin wetland vegetation, Australia, using MODIS NDVI. *Ecological Indicators*, 34, 41-52

- Pham-Duc, B., Prigent, C., Aires, F., & Papa, F. (2017). Comparisons of global terrestrial surface water datasets over 15 years. *Journal of Hydrometeorology*, *18*, 993-1007
- Pignatti, S., Palombo, A., Pascucci, S., Romano, F., Santini, F., Simoniello, T., Umberto, A., Vincenzo, C., Acito, N., Diani, M., Matteoli, S., Corsini, G., Casa, R., Bonis, R.D., Laneve, G., & Ananasso, C. (2013). The PRISMA hyperspectral mission: Science activities and opportunities for agriculture and land monitoring, in: *International Geoscience and Remote Sensing Symposium*, Melbourne, Australia, 21-26 July 2013, 4558-4561
- Poortinga, A., Clinton, N., Saah, D., Cutter, P., Chishtie, F., Markert, K.N., Anderson, E.R., Troy, A., Fenn, M., Tran, L.H., Bean, B., Nguyen, Q., Bhandari, B., Johnson, G., & Towashiraporn, P. (2018). An operational Before-After-Control-Impact (BACI) designed platform for vegetation monitoring at planetary scale. *Remote Sensing*, 10, 760
- Prigent, C., Lettenmaier, D.P., Aires, F., & Papa, F. (2016). Toward a high-resolution monitoring of continental surface water extent and dynamics, at global scale: from GIEMS (Global Inundation Extent from Multi-Satellites) to SWOT (Surface Water Ocean Topography). *Surveys in Geophysics*, 37, 339-355
- Prigent, C., Papa, F., Aires, F., Jimenez, C., Rossow, W.B., & Matthews, E. (2012). Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophysical Research Letters*, 39, L08403
- Prigent, C., Papa, F., Aires, F., Rossow, W.B., & Matthews, E. (2007). Global inundation dynamics inferred from multiple satellite observations, 1993–2000. *Journal of Geophysical Research-Atmospheres*, 112, D12107
- Quinlan, J.R. (1993). Combining instance-based and model-based learning, in: International Conference on Machine Learning, Amherst, USA, 27-29 July 1993, 236-243
- R Core Team (2013). R: A language and environment for statistical computing, in: Vienna, Austria, R foundation for statistical computing
- Ramsar Convention on Wetlands (2002). Wetlands: water, life and culture, in: The Conference of the Contracting Parties to the Convention on Wetlands, Valencia, Spain, 18-26 November 2002
- Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., & Guth, P. (2013). Global carbon dioxide emissions from inland waters. *Nature*, 503, 355-359

- Redfern, J.V., Grant, C.C., Gaylard, A., & Getz, W.M. (2005). Surface water availability and the management of herbivore distributions in an African savanna ecosystem. *Journal of Arid Environments*, 63, 406-424
- Rendón-Martos, M. (1996). La laguna de Fuente de Piedra en la dinámica de la población de flamencos (Phoenicopterus ruber roseus) del Mediterráneo occidental. Doctoral Dissertation, University of Málaga
- Rendón, M.A., Green, A.J., Aguilera, E., & Almaraz, P. (2008). Status, distribution and long-term changes in the waterbird community wintering in Doñana, south-west Spain. *Biological Conservation*, 141, 1371-1388
- Reschke, J., & Huttich, C. (2014). Continuous field mapping of Mediterranean wetlands using sub-pixel spectral signatures and multi-temporal Landsat data. *International Journal of Applied Earth Observation and Geoinformation*, 28, 220-229
- Richards, J.A., Woodgate, P.W., & Skidmore, A.K. (1987). An explanation of enhanced radar backscattering from flooded forests. *International Journal of Remote Sensing*, 8, 1093-1100
- Ricko, M., Carton, J.A., Birkett, C.M., & Cretaux, J.-F. (2012). Intercomparison and validation of continental water level products derived from satellite radar altimetry. *Journal of Applied Remote Sensing*, 6, 061710
- Robledano, F., Esteve, M.A., Farinos, P., Carreno, M.F., & Martinez-Fernandez, J. (2010). Terrestrial birds as indicators of agricultural-induced changes and associated loss in conservation value of Mediterranean wetlands. *Ecological Indicators*, 10, 274-286
- Rodriguez-Rodriguez, M., Martos-Rosillo, S., & Pedrera, A. (2016). Hydrogeological behaviour of the Fuente-de-Piedra playa lake and tectonic origin of its basin (Malaga, southern Spain). *Journal of Hydrology*, 543, 462-476
- Rodriguez, Y.C., El Anjoumi, A., Gomez, J.A.D., Perez, D.R., & Rico, E. (2014). Using Landsat image time series to study a small water body in Northern Spain. *Environmental Monitoring and Assessment, 186*, 3511-3522
- Rokni, K., Ahmad, A., Selamat, A., & Hazini, S. (2014). Water feature extraction and change detection using multitemporal Landsat imagery. *Remote Sensing*, 6, 4173-4189
- Roshier, D.A., Robertson, A.I., & Kingsford, R.T. (2002). Responses of waterbirds to flooding in an arid region of Australia and implications for conservation. *Biological Conservation*, 106, 399-411

- Roshier, D.A., Whetton, P.H., Allan, R.J., & Robertson, A.I. (2001). Distribution and persistence of temporary wetland habitats in arid Australia in relation to climate. *Austral Ecology*, *26*, 371-384
- Rover, J., Wylie, B.K., & Ji, L. (2010). A self-trained classification technique for producing 30 m percent-water maps from Landsat data. *International Journal of Remote Sensing*, 31, 2197-2203
- Ruiz, E. (2008). Management of Natura 2000 habitats. 3170* Mediterranean temporary ponds, European Commission
- Ryu, J.-H., Won, J.-S., & Min, K.D. (2002). Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea. *Remote Sensing of Environment*, 83, 442-456
- Salomon, J., Hodges, J.C.F., Friedl, M., Schaaf, C., Strahler, A., Gao, F., Schneider, A., Zhang, X., Saleous, N.E., & Wolfe, R.E. (2004). Global land-water mask derived from MODIS Nadir BRDF-adjusted reflectances (NBAR) and the MODIS land cover algorithm, in: *International Geoscience and Remote Sensing Symposium*, Anchorage, USA, 20-24 September 2004, 239-241
- Santoro, M., Lamarche, C., Bontemps, S., Wegmüller, U., Kalogirou, V., Arino, O., & Defourny, P. (2013). Introducing a global dataset of open permanent water bodies, in: *ESA Living Planet Symposium*, Edinburgh, UK, 9-13 September 2013
- Sayer, A.M., Hsu, N.C., Bettenhausen, C., Jeong, M.J., & Meister, G. (2015). Effect of MODIS Terra radiometric calibration improvements on Collection 6 Deep Blue aerosol products: validation and Terra/Aqua consistency. *Journal of Geophysical Research-Atmospheres*, 120, 157-174
- Schaaf, C., & Wang, Z. (2015). MCD43A2 MODIS/Terra+Aqua BRDF/Albedo Quality Daily L3 Global - 500m V006. NASA EOSDIS Land Processes DAAC
- Schaaf, C., & Wang., Z. (2015). MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006. NASA EOSDIS Land Processes DAAC
- Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B., Liang, S., Privette, J.L., & Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*, 135-148

- Schaffer-Smith, D., Swenson, J.J., Barbaree, B., & Reiter, M.E. (2017). Three decades of Landsat-derived spring surface water dynamics in an agricultural wetland mosaic; Implications for migratory shorebirds. *Remote Sensing of Environment*, 193, 180-192
- Schmidt, K.S., & Skidmore, A.K. (2003). Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, 85, 92-108
- Schroeder, R., McDonald, K.C., Chapman, B.D., Jensen, K., Podest, E., Tessler, Z.D., Bohn, T.J., & Zimmermann, R. (2015). Development and evaluation of a multi-year fractional surface water data set derived from active/passive microwave remote sensing data. *Remote Sensing*, 7, 16688-16732
- Schumann, G.J.P., & Moller, D.K. (2015). Microwave remote sensing of flood inundation. *Physics and Chemistry of the Earth*, 83, 84-95
- Schwatke, C., Dettmering, D., Bosch, W., & Seitz, F. (2015). DAHITI an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrology and Earth System Sciences*, *19*, 4345-4364
- Šefferová Stanová, V., Janák, M., & Ripka, J. (2008). Management of Natura 2000 habitats. 1530 * Pannonic salt steppes and salt marshes, European Commission
- Semlitsch, R.D., & Bodie, J.R. (1998). Are small, isolated wetlands expendable? Conservation Biology, 12, 1129-1133
- Shannon, G., Matthews, W.S., Page, B.R., Parker, G.E., & Smith, R.J. (2009). The affects of artificial water availability on large herbivore ranging patterns in savanna habitats: a new approach based on modelling elephant path distributions. *Diversity and Distributions*, 15, 776-783
- Sharma, R.C., Tateishi, R., Hara, K., & Nguyen, L.V. (2015). Developing Superfine Water Index (SWI) for global water cover mapping using MODIS data. *Remote Sensing*, 7, 807-841
- Sheng, Y., Gong, P., & Xiao, Q. (2001). Quantitative dynamic flood monitoring with NOAA AVHRR. International Journal of Remote Sensing, 22, 1709-1724
- Shiklomanov, A.I., Lammers, R.B., & Vörösmarty, C.J. (2002). Widespread decline in hydrological monitoring threatens Pan-Arctic Research. *Eos, Transactions American Geophysical Union*, 83, 13-17
- Sim, L.L., Davis, J.A., Strehlow, K., McGuire, M., Trayler, K.M., Wild, S., Papas, P.J.,
 & O'Connor, J. (2013). The influence of changing hydroregime on the invertebrate communities of temporary seasonal wetlands. *Freshwater Science*, *32*, 327-342

- Sippel, S.J., Hamilton, S.K., Melack, J.M., & Novo, E.M.M. (1998). Passive microwave observations of inundation area and the area/stage relation in the Amazon River floodplain. *International Journal of Remote Sensing*, 19, 3055-3074
- Smith, L.C. (1997). Satellite remote sensing of river inundation area, stage, and discharge: A review. *Hydrological Processes*, *11*, 1427-1439
- Song, C.H. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment*, 95, 248-263
- Stefan, S., Fionnuala, H.O.N., Marianna, B., Christian, D., Viktor, G., Robert, K., Theo van der, S., Andreas, K., Sophie, G.L., Zita, S., Martin, P., Boris, B., Thomas, E., Bernd, N., James, R.M., Katrin, E., Volker, M., & Thomas, W. (2016). Multifunctional floodplain management and biodiversity effects: a knowledge synthesis for six European countries. *Biodiversity and Conservation, 25*, 1349-1382
- Straile, D., Jöhnk, K., & Henno, R. (2003). Complex effects of winter warming on the physicochemical characteristics of a deep lake. *Limnology and Oceanography*, 48, 1432-1438
- Street, F.A., & Grove, A. (1976). Environmental and climatic implications of late Quaternary lake-level fluctuations in Africa. *Nature*, 261, 385-390
- Subin, Z.M., Riley, W.J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *Journal* of Advances in Modeling Earth Systems, 4, M02001
- Sun, D.L., Yu, Y.Y., & Goldberg, M.D. (2011). Deriving water fraction and flood maps from MODIS images using a decision tree approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4, 814-825
- Sun, D.L., Yu, Y.Y., Zhang, R., Li, S.M., & Goldberg, M.D. (2012). Towards operational automatic flood detection using EOS/MODIS data. *Photogrammetric Engineering* and Remote Sensing, 78, 637-646
- SWBD (2005). Shuttle Radar Topography Mission Water Body Data set. Retrieved from: https://www2.jpl.nasa.gov/srtm/index.html
- Tan, Z.Q., Li, Y.L., Xu, X.L., Yao, J., & Zhang, Q. (2019). Mapping inundation dynamics in a heterogeneous floodplain: Insights from integrating observations and modeling approach. *Journal of Hydrology*, 572, 148-159
- Ticehurst, C., Guerschman, J.P., & Chen, Y. (2014). The strengths and limitations in using the daily MODIS open water likelihood algorithm for identifying flood events. *Remote Sensing*, 6, 11791-11809

- Tong, X., Pan, H., Xie, H., Xu, X., Li, F., Chen, L., Luo, X., Liu, S., Chen, P., & Jin, Y. (2016). Estimating water volume variations in Lake Victoria over the past 22years using multi-mission altimetry and remotely sensed images. *Remote Sensing of Environment*, 187, 400-413
- Toomey, M., & Vierling, L.A. (2005). Multispectral remote sensing of landscape level foliar moisture: techniques and applications for forest ecosystem monitoring. *Canadian Journal of Forest Research*, 35, 1087-1097
- Tornos, L., Huesca, M., Dominguez, J.A., Moyano, M.C., Cicuendez, V., Recuero, L., & Palacios-Orueta, A. (2015). Assessment of MODIS spectral indices for determining rice paddy agricultural practices and hydroperiod. *ISPRS Journal of Photogrammetry* and Remote Sensing, 101, 110-124
- Townsend, P.A. (2001). Mapping seasonal flooding in forested wetlands using multitemporal radarsat SAR. *Photogrammetric Engineering and Remote Sensing*, 67, 857-864
- Tranvik, L.J., Downing, J.A., Cotner, J.B., Loiselle, S.A., Striegl, R.G., Ballatore, T.J., Dillon, P., Finlay, K., Fortino, K., Knoll, L.B., Kortelainen, P.L., Kutser, T., Larsen, S., Laurion, I., Leech, D.M., McCallister, S.L., McKnight, D.M., Melack, J.M., Overholt, E., Porter, J.A., Prairie, Y., Renwick, W.H., Roland, F., Sherman, B.S., Schindler, D.W., Sobek, S., Tremblay, A., Vanni, M.J., Verschoor, A.M., von Wachenfeldt, E., & Weyhenmeyer, G.A. (2009). Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, *54*, 2298-2314
- Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*, 127-150
- Tulbure, M.G., & Broich, M. (2019). Spatiotemporal patterns and effects of climate and land use on surface water extent dynamics in a dryland region with three decades of Landsat satellite data. *Science of The Total Environment*, 658, 1574-1585
- Tulbure, M.G., Broich, M., Stehman, S.V., & Kommareddy, A. (2016). Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sensing of Environment*, 178, 142-157
- Turak, E., Harrison, I., Dudgeon, D., Abell, R., Bush, A., Darwall, W., Finlayson, C.M., Ferrier, S., Freyhof, J., Hermoso, V., Juffe-Bignoli, D., Linke, S., Nel, J., Patricio, H.C., Pittock, J., Raghavan, R., Revenga, C., Simaika, J.P., & De Wever, A. (2017). Essential biodiversity variables for measuring change in global freshwater biodiversity. *Biological Conservation*, 213, 272-279

- van Dijk, A.I.J.M., Beck, H.E., Crosbie, R.S., de Jeu, R.A.M., Liu, Y.Y., Podger, G.M., Timbal, B., & Viney, N.R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, 49, 1040-1057
- Van Trung, N., Choi, J.H., & Won, J.S. (2013). A land cover variation model of water level for the floodplain of Tonle Sap, Cambodia, derived from ALOS PALSAR and MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6, 2238-2253
- Verpoorter, C., Kutser, T., Seekell, D.A., & Tranvik, L.J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41, 6396-6402
- Vikhamar, D., & Solberg, R. (2003). Snow-cover mapping in forests by constrained linear spectral unmixing of MODIS data. *Remote Sensing of Environment*, 88, 309-323
- Vörösmarty, C.J., Green, P., Salisbury, J., & Lammers, R.B. (2000). Global water resources: vulnerability from climate change and population growth. *Science*, 289, 284-288
- Vrieling, A., de Beurs, K.M., & Brown, M.E. (2011). Variability of African farming systems from phenological analysis of NDVI time series. *Climatic Change*, 109, 455-477
- Vrieling, A., de Leeuw, J., & Said, M.Y. (2013). Length of growing period over Africa: variability and trends from 30 years of NDVI time series. *Remote Sensing*, 5, 982-1000
- Walton, J.T. (2008). Subpixel urban land cover estimation: comparing cubist, random forests, and support vector regression. *Photogrammetric Engineering and Remote Sensing*, 74, 1213-1222
- Wang, C. (2008). Detecting the effect of water regime on waterbirds population using remote sensing. Master Dissertation, University of Twente
- Wang, D.D., Morton, D., Masek, J., Wu, A.S., Nagol, J., Xiong, X.X., Levy, R., Vermote, E., & Wolfe, R. (2012). Impact of sensor degradation on the MODIS NDVI time series. *Remote Sensing of Environment*, 119, 55-61
- Wang, J., Zhao, Y., Li, C., Yu, L., Liu, D., & Gong, P. (2015a). Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250m resolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 38-47

- Wang, L.L., Qu, J.J., Hao, X.J., & Zhu, Q.P. (2008). Sensitivity studies of the moisture effects on MODIS SWIR reflectance and vegetation water indices. *International Journal of Remote Sensing*, 29, 7065-7075
- Wang, P., Huang, C., & Brown de Colstoun, E.C. (2017). Mapping 2000–2010 impervious surface change in India using global land survey landsat data. *Remote Sensing*, 9, 366
- Wang, S., Baig, M.H.A., Zhang, L., Jiang, H., Ji, Y., Zhao, H., & Tian, J. (2015b). A simple Enhanced Water Index (EWI) for percent surface water estimation using Landsat data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 90-97
- Wang, X., Ling, F., Yao, H., Liu, Y., & Xu, S. (2019). Unsupervised sub-pixel water body mapping with Sentinel-3 OLCI image. *Remote Sensing*, 11, 327
- Wang, X., Smith, K., & Hyndman, R. (2006). Characteristic-based clustering for time series data. *Data Mining and Knowledge Discovery*, 13, 335-364
- Wang, X., Xie, S., Zhang, X., Chen, C., Guo, H., Du, J., & Duan, Z. (2018a). A robust Multi-Band Water Index (MBWI) for automated extraction of surface water from Landsat 8 OLI imagery. *International Journal of Applied Earth Observation and Geoinformation*, 68, 73-91
- Wang, Z., Liu, J., Li, J., & Zhang, D.D. (2018b). Multi-Spectral Water Index (MuWI): A native 10-m multi-spectral water index for accurate water mapping on Sentinel-2. *Remote Sensing*, 10, 1643
- Waterkeyn, A., Grillas, P., Vanschoenwinkel, B., & Brendonck, L. (2008). Invertebrate community patterns in Mediterranean temporary wetlands along hydroperiod and salinity gradients. *Freshwater Biology*, 53, 1808-1822
- Watts, J.D., Kimball, J.S., Jones, L.A., Schroeder, R., & McDonald, K.C. (2012). Satellite Microwave remote sensing of contrasting surface water inundation changes within the Arctic–Boreal Region. *Remote Sensing of Environment*, 127, 223-236
- Wdowinski, S., Kim, S.W., Amelung, F., Dixon, T.H., Miralles-Wilhelm, F., & Sonenshein, R. (2008). Space-based detection of wetlands' surface water level changes from L-band SAR interferometry. *Remote Sensing of Environment*, 112, 681-696
- Weiss, D.J., & Crabtree, R.L. (2011). Percent surface water estimation from MODIS BRDF 16-day image composites. *Remote Sensing of Environment*, 115, 2035-2046
- Weng, Q.H. (2012). Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. *Remote Sensing of Environment*, 117, 34-49

- Wiesnet, D., McGinnis, D., & Pritchard, J. (1974). Mapping of the 1973 Mississippi river floods by NOAA-2 satellite. JAWRA Journal of the American Water Resources Association, 10, 1040-1049
- Wik, M., Varner, R.K., Anthony, K.W., MacIntyre, S., & Bastviken, D. (2016). Climatesensitive northern lakes and ponds are critical components of methane release. *Nature Geoscience*, 9, 99
- Wolski, P., Murray-Hudson, M., Thito, K., & Cassidy, L. (2017). Keeping it simple: Monitoring flood extent in large data-poor wetlands using MODIS SWIR data. *International Journal of Applied Earth Observation and Geoinformation*, 57, 224-234
- Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., & Roy, D.P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, 185, 271-283
- Xia, H., Zhao, W., Li, A., Bian, J., & Zhang, Z. (2017). Subpixel inundation mapping using Landsat-8 OLI and UAV Data for a wetland region on the Zoige Plateau, China. *Remote Sensing*, 9, 31
- Xian, G., Homer, C., Demitz, J., Fry, J., Hossain, N., & Wickham, J. (2011). Change of impervioussurface area between 2001 and 2006 in the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 77, 758-762
- Xiao, J.F., & Moody, A. (2005). A comparison of methods for estimating fractional green vegetation cover within a desert-to-upland transition zone in central New Mexico, USA. *Remote Sensing of Environment*, 98, 237-250
- Xiao, X., Boles, S., Frolking, S., Salas, W., Moore, B., Li, C., He, L., & Zhao, R. (2002a). Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using VEGETATION sensor data. *International Journal of Remote Sensing*, 23, 3009-3022
- Xiao, X.M., Boles, S., Frolking, S., Li, C.S., Babu, J.Y., Salas, W., & Moore, B. (2006).
 Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment*, 100, 95-113
- Xiao, X.M., Boles, S., Liu, J.Y., Zhuang, D.F., Frolking, S., Li, C.S., Salas, W., & Moore,
 B. (2005). Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, 95, 480-492
- Xiao, X.M., Boles, S., Liu, J.Y., Zhuang, D.F., & Liu, M.L. (2002b). Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. *Remote Sensing of Environment*, 82, 335-348

- Xie, H., Luo, X., Xu, X., Tong, X.H., Jin, Y.M., Pan, H.Y., & Zhou, B.Z. (2014). New hyperspectral difference water index for the extraction of urban water bodies by the use of airborne hyperspectral images. *Journal of Applied Remote Sensing*, *8*, 15
- Xing, L., Tang, X., Wang, H., Fan, W., & Wang, G. (2018). Monitoring monthly surface water dynamics of Dongting Lake using Sentinel-1 data at 10 m. *PeerJ*, 6, E4992
- Xu, H.Q. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27, 3025-3033
- Xu, K., Zhang, J., Watanabe, M., & Sun, C. (2004). Estimating river discharge from very high-resolution satellite data: a case study in the Yangtze River, China. *Hydrological Processes*, 18, 1927-1939
- Yamazaki, D., O'Loughlin, F., Trigg, M.A., Miller, Z.F., Pavelsky, T.M., & Bates, P.D. (2014). Development of the global width database for large rivers. *Water Resources Research*, 50, 3467-3480
- Yamazaki, D., & Trigg, M.A. (2016). The dynamics of Earth's surface water. *Nature*, 540, 348-349
- Yamazaki, D., Trigg, M.A., & Ikeshima, D. (2015). Development of a global ~90m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, 171, 337-351
- Yan, L., & Roy, D.P. (2018). Large-area gap filling of Landsat reflectance time series by spectral-angle-mapper based spatio-temporal similarity (SAMSTS). *Remote Sensing*, 10, 609
- Yan, Y.E., Ouyang, Z.T., Guo, H.Q., Jin, S.S., & Zhao, B. (2010). Detecting the spatiotemporal changes of tidal flood in the estuarine wetland by using MODIS time series data. *Journal of Hydrology*, 384, 156-163
- Yang, J., He, Y., & Oguchi, T. (2014). An endmember optimization approach for linear spectral unmixing of fine-scale urban imagery. *International Journal of Applied Earth Observation and Geoinformation, 27, Part B*, 137-146
- Yang, J., Liu, C., Shu, R., & Xie, F. (2018a). The extraction of urban surface water from hyperspectral data based on spectral indices. *Journal of the Indian Society of Remote Sensing*, 46, 1749-1759
- Yang, S., Sun, X., Lv, K., Wang, W., Xiao, W., & Hu, Y. (2011). Mapping the variation of Poyang Lake area based on multi-temporal MODIS imagery, in: *International Conference on Multimedia Technology*, Hangzhou, China, 26-28 July 2011, 4970-4973

- Yang, X., Qin, Q., Grussenmeyer, P., & Koehl, M. (2018b). Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. *Remote Sensing of Environment*, 219, 259-270
- Yang, Y., Liu, Y., Zhou, M., Zhang, S., Zhan, W., Sun, C., & Duan, Y. (2015). Landsat 8 OLI image based terrestrial water extraction from heterogeneous backgrounds using a reflectance homogenization approach. *Remote Sensing of Environment*, 171, 14-32
- Yu, L., Wang, J., & Gong, P. (2013). Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach. *International Journal of Remote Sensing*, 34, 5851-5867
- Zacharias, I., Dimitriou, E., Dekker, A., & Dorsman, E. (2007). Overview of temporary ponds in the Mediterranean region: Threats, management and conservation issues. *Journal of Environmental Biology*, 28, 1-9
- Zacharias, I., & Zamparas, M. (2010). Mediterranean temporary ponds. A disappearing ecosystem. *Biodiversity and Conservation*, *19*, 3827-3834
- Zedler, J.B., & Kercher, S. (2005). Wetland resources: Status, trends, ecosystem services, and restorability. *Annual Review of Environment and Resources*, *30*, 39-74
- Zhan, S., Song, C.Q., Wang, J.D., Sheng, Y.W., & Quan, J.P. (2019). A global assessment of terrestrial evapotranspiration increase due to surface water area change. *Earths Future*, 7, 266-282
- Zhang, G.Q., Yao, T.D., Chen, W.F., Zheng, G.X., Shum, C.K., Yang, K., Piao, S.L., Sheng, Y.W., Yi, S., Li, J.L., O'Reilly, C.M., Qi, S.H., Shen, S.S.P., Zhang, H.B., & Jia, Y.Y. (2019). Regional differences of lake evolution across China during 1960s-2015 and its natural and anthropogenic causes. *Remote Sensing of Environment, 221*, 386-404
- Zhang, H., Gorelick, S.M., Zimba, P.V., & Zhang, X.D. (2017). A remote sensing method for estimating regional reservoir area and evaporative loss. *Journal of Hydrology*, 555, 213-227
- Zhang, L., Ji, L., & Wylie, B.K. (2011). Response of spectral vegetation indices to soil moisture in grasslands and shrublands. *International Journal of Remote Sensing*, 32, 5267-5286
- Zhang, X., Schaaf, C.B., Friedl, M.A., Strahler, A.H., Gao, F., & Hodges, J.C.F. (2002). MODIS tasseled cap transformation and its utility, in: *Geoscience and Remote Sensing Symposium*, Toronto, Canada, 24-28 June 2002, 1063-1065

- Zhou, Y., Dong, J., Xiao, X., Xiao, T., Yang, Z., Zhao, G., Zou, Z., & Qin, Y. (2017). Open surface water mapping algorithms: a comparison of water-related spectral indices and sensors. *Water*, 9, 256
- Zhu, X., Cai, F., Tian, J., & Williams, T.K.-A. (2018). Spatiotemporal fusion of multisource remote sensing data: literature survey, taxonomy, principles, applications, and future directions. *Remote Sensing*, 10, 527
- Zhu, Z., Wang, S., & Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269-277
- Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment, 118*, 83-94

Summary

Terrestrial surface water plays an important role in the global hydrological cycle, biodiversity conservation, and climate processes. Changes in surface water, caused by both natural and human-induced factors, strongly affect socioeconomic development, ecosystem functioning, species distributions and composition, and further influence climate change. Recently, water extent has been identified as an Essential Climate Variable (ECV) for assessing progress towards the Aichi targets for 2020 of the Convention of Biological Diversity. The changes in the extent of water-related ecosystems over time is also an indicator of the Sustainable Development Goals (SDGs). Therefore, knowledge about the spatial and temporal distribution of surface water is needed to support sustainable development and climate change assessment.

Remote sensing provides an effective way to monitor surface water in space and time. Many approaches and datasets have been developed for this purpose. However, measuring long-term changes at fine spatial and temporal resolution remains a challenge due to the trade-off between spatial and temporal resolution of remotely sensed imagery. The main objective of this thesis is to improve long-term mapping and monitoring of surface water extent at fine temporal resolution using high-frequency optical remote sensing data provided by the Moderate Resolution Imaging Spectroradiometer (MODIS), in a way that effectively accounts for small-sized (*e.g.*, smaller than a 500x500 m MODIS cell) and dynamic water bodies. To achieve this goal, this thesis evaluates options to estimate sub-pixel surface water fraction, *i.e.*, the percentage of surface water within a single grid cell. Several machine learning approaches that incorporate MODIS spectral information, temporal characteristics of spectral information, and topographic information were evaluated for accurately mapping and monitoring sub-pixel surface water fraction. This was explored at spatial scales ranging from a small individual wetland to the entire Mediterranean region.

In this thesis, the robustness of machine learning algorithms for mapping and monitoring sub-pixel surface water fraction at large spatial scales was demonstrated. It revealed that a single model could accurately assess water body extent and dynamics in different environmental and climatic conditions, as long as good-quality training data were collected that represent the various environmental conditions. Therefore, there is potential to scale the water fraction mapping approach to larger spatial regions, such as for the globe.

Summary

The need for high-frequency monitoring of surface water is highlighted in the thesis. While much progress has been made recently with global Landsat-based surface water products, these can have large spatial and temporal gaps due to both the limited number of acquisitions and persistent cloud cover, preventing an accurate assessment of water resources variability. Here, using the high-frequency MODIS data and the approach developed in the thesis, a new dense 18-year surface water fraction (SWF) dataset was produced for the Mediterranean region at 500 m resolution and 8-day interval. This MODIS SWF dataset documents the long-term (2000–2017) status of surface water bodies, their location, extent, and change. MODIS SWF complements existing fine spatial resolution water datasets, especially by offering better temporal information for areas suffering from persistent cloud cover during part of the year. Moreover, it allows accurate assessment of surface water bodies, including short-duration surface water that could not be captured by existing datasets of lower temporal resolution. The dataset also accurately detects small water bodies (less than one MODIS pixel) and narrow rivers.

Samenvatting

Oppervlaktewater speelt een belangrijke rol in de globale hydrologische cyclus, het behoud van biodiversiteit en klimaatprocessen. Veranderingen in oppervlaktewater, veroorzaakt door zowel natuurlijke als door de mens veroorzaakte factoren, hebben een sterke invloed op sociaaleconomische ontwikkelingen, het functioneren van ecosystemen, soortverdelingen en -samenstelling en op klimaatverandering. Onlangs is de waterstand geduid als een essentiële klimaatvariabele (ECV) voor het beoordelen van de voortgang op weg naar de Aichi-doelstellingen voor 2020 van het Verdrag inzake biologische diversiteit. De temporele veranderingen in de omvang van watergerelateerde ecosystemen is ook een indicator van de Duurzame Ontwikkelingsdoelstellingen (SDGs). Een betere kennis van de ruimtelijke en temporele verdeling van oppervlaktewater is daarom nodig om duurzame ontwikkeling en klimaatveranderingsanalyses te ondersteunen.

Remote sensing biedt effectieve manieren om de veranderingen in oppervlaktewater in ruimte en tijd te volgen. Hiervoor zijn verschillende datasets en benaderingen ontwikkeld. Het meten van langetermijnveranderingen met fijne ruimtelijke en temporele resolutie blijft echter een uitdaging. Deze dissertatie heeft tot doel de langetermijnkartering en monitoring van het areaal aan oppervlaktewater met fijne temporele resolutie te verbeteren met behulp van hoogfrequente optische satellietbeelden van de Moderate Resolution Imaging Spectroradiometer (MODIS), op een manier die effectief rekening houdt met kleine (bijv. kleiner dan een MODIS-cel) en dynamische wateroppervlaktewater te schatten, d.w.z. het percentage oppervlaktewater binnen een enkele rastercel. Verschillende benaderingen van machinaal leren die gebruik maken van MODIS spectrale informatie, temporele kenmerken van spectrale informatie en topografische informatie, werden geëvalueerd voor het nauwkeurig in kaart brengen en volgen van de oppervlaktewaterfractie. Dit werd onderzocht op ruimtelijke schaal, variërend van een klein individueel drasland tot het hele Middellandse-Zeegebied.

Dit proefschrift demonstreert de robuustheid van algoritmen voor het machinaal leren voor het in kaart brengen en monitoren van oppervlaktewater op grote ruimtelijke schaalniveaus. Er werd aangetoond dat een enkel model een nauwkeurige schatting kon maken van de grootte en dynamiek van waterlichamen in verschillende omgevings- en klimatologische omstandigheden, op voorwaarde dat er trainingsgegevens werden verzameld die de verschillende omstandigheden goed vertegenwoordigen. Daarom is er

Samenvatting

potentie om deze benadering op te schalen naar grotere ruimtelijke regio's, zoals voor de hele wereld.

Dit proefschrift benadrukt de noodzaak van hoogfrequente monitoring van oppervlaktewater. Hoewel er onlangs veel vooruitgang is geboekt met de wereldwijde op Landsat-satellieten gebaseerde gegevens over het oppervlaktewater, kunnen deze gegevens soms grote ruimtelijke en temporele hiaten bevatten vanwege zowel het beperkte aantal beeldopnames alsook de aanhoudende bewolking in de lucht, waardoor nauwkeurige schattingen van veranderingen in oppervlaktewater niet mogelijk zijn. Met behulp van de hoogfrequente MODIS-gegevens en de ontwikkelde benadering van machinaal leren, is in deze dissertatie een nieuwe 18-jarige oppervlaktewaterfractie (SWF) dataset voor het Middellandse-Zeegebied geproduceerd met een ruimtelijke resolutie van 500m en 8-daags interval. Deze MODIS SWF-dataset documenteert de lange termijn (2000–2017) status van oppervlaktewaterlichamen, hun locatie, omvang en veranderingen. MODIS SWF is een aanvulling op bestaande fijne ruimtelijke resolutiewater datasets, met name door het aanbieden van betere temporele informatie voor gebieden waar het gedurende een deel van het jaar vaak bewolkt is. Bovendien maakt het een nauwkeurige beoordeling van de seizoensinvloeden van het oppervlaktewater mogelijk, door een betere representatie van snelle en kortstondige veranderingen in grootte van waterlichamen, zoals voor gebieden die slechts voor korte duur met oppervlaktewater bedekt zijn en daardoor niet voorkomen in bestaande datasets met lagere temporele resolutie. De dataset detecteert ook nauwkeurig de kleinere waterlichamen (minder dan één MODIS pixel) en smalle rivieren.

Biography



Linlin Li was born on 30 April, 1986 in Hebei, China. In 2009, she obtained the BSc degree in Geographic Information System from Lanzhou University, China. She then was recommended for admission to be a postgraduate without examination at Lanzhou University. She received her MSc degree in Cartography and Geographic Information System in June 2012. In the same year, she was awarded a

four-year doctorial scholarship from the China Scholarship Council (CSC), and started to pursue her PhD at the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente. Her research interest lies in mapping and monitoring of surface water and wetlands, time series analysis, and change detection.